Vision-Based Two-Factor Authentication & Localization Scheme for Autonomous Vehicles

Anas Alsoliman
University of California, Irvine
aalsolim@uci.edu

Marco Levorato
University of California, Irvine
levorato@uci.edu

Qi Alfred Chen
University of California, Irvine
alfchen@uci.edu

***Abstract*—In autonomous vehicle systems – whether ground or aerial – vehicles and infrastructure-level units communicate among each other continually to ensure safe and efficient autonomous operations. However, different attack scenarios might arise in such environments when a device in the network cannot physically pinpoint the actual transmitter of a certain message. For example, a compromised or a malicious vehicle could send a message with a fabricated location to appear as if it is in the location of another legitimate vehicle, or fabricate multiple messages with fake identities to alter the behavior of other vehicles/infrastructure units and cause traffic congestion or accidents. In this paper, we propose a Vision-Based Two-Factor Authentication and Localization Scheme for Autonomous Vehicles. The scheme leverages the vehicles' light sources and cameras to establish an "Optical Camera Communication (OCC)" channel providing an auxiliary channel between vehicles to visually authenticate and localize the transmitter of messages that are sent over Radio Frequency (RF) channels. Additionally, we identify possible attacks against the proposed scheme as well as mitigation strategies.**

I. Introduction

Autonomous vehicles such as self-operated drones and cars will soon enable new applications in different domains, such as autonomous transportation and drone delivery. However, safety-critical operations, such as object-detection and avoidance, are integral components of these applications and it would make them appealing targets for cybercriminals [1]. In the United States, various efforts from government agencies have been directed to set up the rules, regulations, and policies for managing the secure operations of future autonomous systems (i.e., vehicles and Roadside Units, or RSUs). For connected and self-driving cars, the U.S. Department of Transportation (USDOT) has adopted the Security Credential Management System (SCMS) [2] for handling secure vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications. For Unmanned Aerial Vehicles (UAVs), the Federal Aviation Administration (FAA) and National Aeronautics and Space Administration (NASA) have jointly developed the Remote ID framework [3] which would set the foundation for the anticipated Unmanned Aircraft System (UAS) Traffic Management System (UTM) [4]. Unfortunately, these systems still lack effective security measures to defend against existing attacks [5][6].

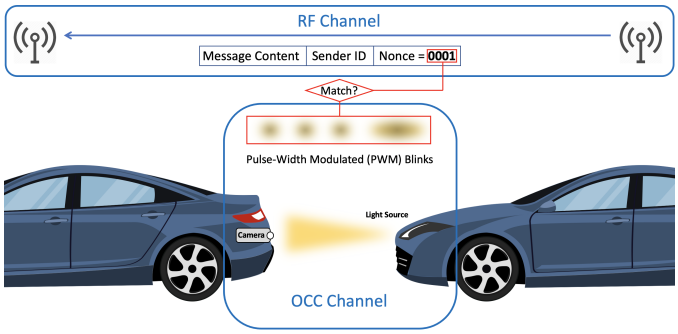


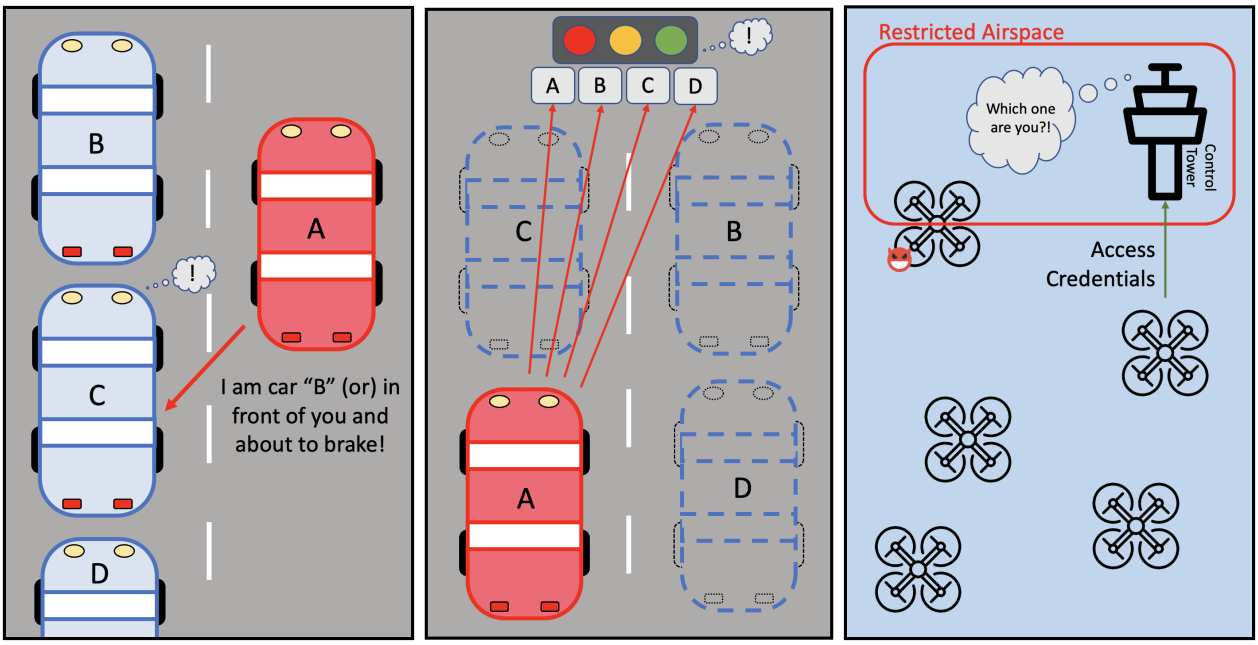
Fig. 1: Scheme Overview: a vehicle sends an RF message and a visual nonce as optical Pulse-Width Modulated (PWM) symbols using its headlights

In the academic community, several attacks have been demonstrated against autonomous systems such as GPS spoofing attacks [7], V2I attacks [8], and Sybil (vehicle ID duplication) attacks [9]. These attacks have different threat models and attack methodologies, and find different mitigation approaches. However, we contend that all these attacks stem from the same root: *the decoupling between the sender of a message and its transmitting location*. In other words, if receiver of a message cannot locate its transmitter, multiple threats are exposed.

In this paper, we propose a Vision-Based Two-Factor Authentication & Localization Scheme for Autonomous Vehicles to pinpoint and authenticate the location of a message transmitter. In this scheme, we leverage the lighting source of the vehicle to create an Optical Camera Communication (OCC) channel to send a random *nonce* – a randomly generated number that to be used only once – encoded as flashing blinks from the light source, while at the same time the sending vehicle includes the same nonce into the message to be sent over the Radio Frequency (RF) channel (figure 1). In this scheme, the receiver is equipped with an RF interface and a camera, so that it can receive the message via the RF channel while simultaneously use its camera to record the sender's blinks and decode the visually modulated nonce using computer-vision algorithms. The receiver then matches the

Acknowledgement:
This work was partially supported by *CNS-1850533, CNS-1929771, and USDOT UTC Grant 69A3552047138.*
This work was partially supported by the *NSF grant IIS-1724331.*

Network and Distributed Systems Security (NDSS) Symposium 2021
21-24 February 2021
ISBN 1-891562-66-5
https://dx.doi.org/10.14722/autosec.2021.23021
www.ndss-symposium.org



(a) C cannot tell if the message came from A or B

(b) RSU cannot tell who actually sent the four messages

(c) Control tower cannot tell which one is the legitimate UAV

Fig. 2: Aerial view of three different attack scenarios

nonces received on the two channels (RF and OCC). If the nonces match, then the vehicle is visually authenticated and localized and is indeed identified as the transmitter of the message. Note that our scheme can be used to authenticate each RF message as well as it can be used as a bootstrap to authenticate a vehicle's location once then establish a secure channel with the localized vehicle.

We summarize our contributions as follows:

- We propose a vision-based authentication scheme that pinpoints (authenticates and localizes) the actual transmitter of a message sent over the RF channel by utilizing cameras and light sources to create an OCC channel.
- We identify a possible attack (named a Copycat Attack) with different variants against the proposed scheme, where a malicious party mimics another legitimate transmitter's visual blinks to spoof its identity or location.
- We present mitigation approaches that circumvent all variants of the copycat attack.

In the next section, we present our threat model with three different attack scenarios. In section III, we give a brief background about wireless communications that utilize optical techniques. In section IV, we provide an overview of our proposed scheme. In section V, we evaluate the attack scenarios discussed in section II against our scheme presented in section IV. Section VI demonstrates a new adaptive attack (named Copycat Attack) under multiple scenarios as well as mitigation approaches for each scenario. Related work is discussed in section VII and section VIII concludes the paper.

II. Threat Model

In our threat model (figure 2), a malicious party exploits the decoupling of a sent message from the location of its transmitter. Therefore, our main objective in this threat model is to identify and authenticate the physical location of a message's transmitter, whether the transmitter is an attacker or a legitimate vehicle. To illustrate further, three different attack scenarios are demonstrated as examples:

Location Spoofing: In this scenario, a compromised or a malicious vehicle fabricates fake GPS coordinates about its current location to cause accidents, alter a system's behavior, or earn unlawful privileges. For example in figure 2a, a platoon of vehicles (cars $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$) is already formed and a malicious adversarial vehicle $\mathcal{A}$ sends an emergency-braking message to vehicle $\mathcal{C}$ while pretending to be in the location of vehicle $\mathcal{B}$. Or if vehicle $\mathcal{A}$ is capable of compromising the identity of vehicle $\mathcal{B}$, $\mathcal{A}$ can pretend to be $\mathcal{B}$ itself without the need of fabricating any coordinates. Vehicle $\mathcal{C}$ would react to $\mathcal{A}$'s fabricated message and apply its brakes to avoid crashing into $\mathcal{B}$ but also it would cause $\mathcal{D}$ to crash into $\mathcal{C}$ if there are no implemented precautions against this scenario. In this scenario, the malicious vehicle does not need to wait for an opening to physically drive up in front of the targeted vehicle. Instead, the attack can be launched from faraway.

Identity Duplication: Identity Duplication attack, also known as Sybil Attack, is an attack carried out when a single malicious vehicle sends out multiple messages with multiple identities to give an impression of a congested road and alter the behavior of the infrastructure and/or other vehicles. For example, in figure 2b, the additional fake vehicles (ghost cars

$\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$) generated by $\mathcal{A}$ would cause the traffic light to turn green sooner in the attacker's lane and longer for other lanes in order to clear the congested lane.

Identity Confusion: In this scenario (figure 2c), a swarm of unmanned aerial vehicles (UAVs) are flying in close-proximity of each other. A legitimate UAV is supplementing its credentials to the control tower for accessing the nearby restricted airspace. During that time, another intruding UAV enters the restricted airspace. Here the control tower cannot enforce the given access (e.g., take down the intruding UAV) since it cannot physically differentiate between the legitimate and the intruding UAV.

We can observe that each scenario represents different attacking capabilities. In the first scenario, the attacker has the ability to either *fabricate the content of its messages or use the identity of another legitimate vehicle*. In the second scenario, the attacker has the ability to *generate multiple identities of non-existing vehicles*. In the third scenario, the attacker is able to *passively listen to messages* in order to make an opportunistic attack (e.g. invasive access towards a restricted airspace). However, all three scenarios can be exploited due to the same reason; *the inability of the receiver to localize the transmitter of the messages*.

III. Background on Optical Wireless Communication

A. Overview

Optical Wireless Communication (OWC) is any communication channel that utilizes the terahertz band of the electromagnetic spectrum which includes the infrared and ultraviolet frequencies. Furthermore, an OWC channel utilizing the visible-light portion of the terahertz band is referred to as Visible Light Communication (VLC).

Optical communications require an imaging sensor for receiving the incoming light photons. That sensor is commonly known as a photodetector. A camera sensor such as Complementary Metal Oxide Semiconductor (CMOS) consists of a matrix of photodetectors, each represents a pixel which gets its color by measuring the intensity and frequency of the recorded photon by its corresponding photodetector. The main objective of a camera sensor is to create an image out of the CMOS output as a grid of colored pixels, and the rate of images created by the sensor is known as frame per second (FPS). However, since cameras have become a commodity hardware, their sensors have been repurposed for various applications such as decoding messages that are digitally modulated into images. This type of optical communication technique is commonly known as Optical Camera Communication (OCC).

B. OCC Channel Model

Our channel model is based on two assumptions: **(i)** The modulation scheme used over the OCC channel is assumed to be optical On-Off Keying (OOK) since it only requires a single narrowband frequency (i.e., single color) which makes it applicable to any light sources. Note that visual symbols in figure 1 and 3 are pulse-width modulated (PWM) for ease of illustration only. **(ii)** We assumes the use of cameras that are operated using the global shutter mode where all camera pixels are scanned simultaneously. The rolling shutter mode (its counterpart) scans rows of pixel independently one after another. However, the security intuitions in this paper still holds true for both assumptions and should be applicable to different imaging techniques.

In OCC systems, a communication channel using OOK modulation can be modeled as a rectangular waveform. Let T_e denote to the camera exposure time which defines how long a camera shutter stays open to capture a single frame. Therefore, T_e is the inverse of the camera sampling rate FPS such that $T_e = \frac{1}{FPS}$. Let PW_s denote the symbol pulse-width which defines how long a light source stays on during the T_e time frame. Intuitively, we have $PW_s < T_e$. However, there is a minimum duty cycle $DC_{min} = \frac{PW_s}{T_e}$ for a given OCC system under certain conditions and SINR requirements. Furthermore, a guard width PW_g must be left unoccupied at the beginning and the end of the T_e time frame to prevent symbols from leaking into adjacent frames (similar to the guard bands that are added between RF channels to prevent inter-symbol interference). Finally, in our scheme, the OCC can be modeled as $PW_s + PW_g < T_e$ for $PW_s \geq DC_{min} \times T_e$. Note that the propagation delay is omitted from the channel model since it has a negligible effect especially for close-range communications such as in V2V and V2I. Furthermore, the guard band PW_g should mitigate the effects of propagation delay whenever it becomes critically high.

C. Challenges

There are several challenges when working with OCC systems such as: **(i)** unidirectional link - a lighting source can only be used for transmitting while a camera is only for receiving, **(ii)** highly directional communications - the camera and the lighting source need to face each other in a clear line-of-sight (LOS), and **(iii)** low bitrate - common camera sensors can sample up to 30 FPS (stand-alone photodetectors are capable of higher sampling rates with more complex modulation schemes, but cannot construct an image out of the received photons).

The requirements of our scheme are not affected by these shortcomings. For instance, the receiver in our scheme does not require a feedback channel for locating the sender, and whenever a feedback is required (e.g. loss in clock synchronization) the RF channel can be used. Also based on the threat model discussed in Section II, the sender is expected to be in a clear view of the receiver which means that clear LOS is an inherent requirement for all the three attack scenarios explained in Section II. Concerning the bitrate, the sender encodes only a short nonce into the OCC channel whereas the main message is sent over the RF channel. Therefore, a low OCC bitrate is sufficient to execute the authentication scheme. Furthermore, lower bitrates can be modulated using longer symbol periods which inherently make it cover larger distances in the optical domain.

IV. Scheme Overview

As discussed in section I, our scheme utilizes two different communication channels, an RF channel and an OCC channel. Both the sender and the receiver are equipped with an RF interface to implement an RF channel while the OCC channel requires the sender to be equipped with lighting source such as a light-emitting diode (LED), and the receiver to be equipped with a camera. We assume that these equipment are already available in most autonomous vehicles since we expect modern autonomous systems to have the minimum set of requirements to carry out safe and secure autonomous operations, which include: an RF interface (for wireless communications), a camera (for object detection and navigation), and a lighting source (for safety, illumination, and identification purposes) such as car's headlights/taillights and drone's anti-collision strobe lights.

Authentication & Localization Process: When a sender transmits a message over the RF channel, it includes a nonce in the message, while at the same time it encodes the same nonce as modulated blinks into the OCC channel using its light source. On the other hand, the receiver will continuously scan every captured camera frame for possible OOK modulated symbols. Here the receiver employs computer-vision algorithms and look at every two consecutive frames for a sudden change in pixel intensity. When a pixel intensity change is found, the algorithm locks on the object emitting the light source that caused the intensity change and extract the symbols encoded into each subsequent frame. The area of the locked object is called Region of Interest (ROI). Now whenever the receiver receives a message over the RF channel, it cross-references its nonce with the demodulated nonces emitted by the current ROI over the OCC channel at the time of message reception. To illustrate, to transmit bit_i = "1" as an OOK modulated symbol, the sender sets its light source on high for PW_s seconds during the camera exposure time frame T_{e_i}. To transmit the next bit_{i+1} = "0", the sender sets its light source to off-state during the next transmission window $T_{e_{i+1}}$. The receiver would detect a transition period from T_{e_i} to $T_{e_{i+1}}$ time frames as a change in pixel intensity caused by bit_i and bit_{i+1} respectively. As a result, the receiver will lock on the ROI region of the captured picture that has a pixel intensity change. This process will be triggered from the very first transmitted bit_1 where a timer T_{OCC} will be set starting from the beginning of T_{e_1}. When the receiver receives a message from the RF channel that has the same nonce which is received from the OCC channel, the time difference between the two messages must satisfy the following condition: $|T_{OCC} - T_{RF}| < \theta$ where T_{RF} denote to the time where the first bit of the message was received over the RF channel and θ denote to a minimum threshold that can be used as an attack window. More details on the aforementioned attack will be discussed in Section VI-A.

Technical Considerations: To prevent accidental ROI locks that are caused by pixel intensity change due to random environmental factors (such as a vehicle turning on its headlights or an object moving in front of a light source which would be interpreted as an OOK modulated symbol), OCC systems would use a preamble with an alternating 0s and 1s to synchronize the sender with the receiver before locking on the ROI and recording the nonce. Another point to consider is that during an OCC transmission, the nonce encoded by the sender might include long runs of 0s or 1s (continuous repetitions of 0s or 1s) which would cause multiple consecutive frames to not have any intensity change across their transitions which in turn cause a loss in clock synchronization on receiver side (cannot distinguish whether the transmission has ended or a long run is being transmitted). Therefore, the sender would use Run Length Limited (RLL) codes such as Manchester encoding where a "1" is represented as "01" and "0" is represented as "10". In this case, the receiver would have at most two consecutive frames with the same modulated symbol (maximum possible run is two frames). Finally, the change in OCC channel state (i.e. blinking) should be *faster than the perception of the human eye* to prevent causing confusions to surrounding human drivers/pedestrians and also physiological effects such as nausea. The IEEE 802.15.7 standard [10] for Short-Range Optical Wireless Communications recommends the use of at least 200 Hz in light flickering frequency.

V. Scheme Defense Analysis

In this section, we evaluate our scheme (section IV) against the threat model (section II).

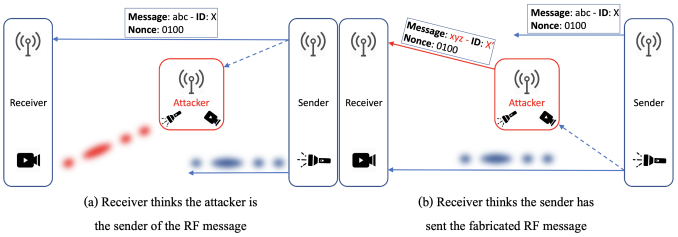
Scenario A: in figure 2a, the malicious vehicle $\mathcal{A}$ attempts to deceive vehicle $\mathcal{C}$ by pretending to be vehicle $\mathcal{B}$ or by fabricating its location to appear as if $\mathcal{A}$ is in front of $\mathcal{C}$. With our scheme, whenever $\mathcal{C}$ receives a message pretending to be from $\mathcal{B}$, it checks the message's nonce with the visual nonce emitted by $\mathcal{B}$ which is physically in front of $\mathcal{C}$. If the two nonces from the two channels (RF and OCC) match, the message is indeed from $\mathcal{B}$. Otherwise, the message will be flagged as a spoofing attempt.

Scenario B: in figure 2b, the malicious vehicle attempts to deceive the traffic light into believing that the lane is crowded by broadcasting multiple messages with different IDs. Here the traffic light will use the OCC channel to match each message with its sender. Since the attacker is the sender of all the messages, all message IDs will be mapped to the same vehicle which will trigger a spoofing attempt.

Scenario C: in figure 2c, a malicious unmanned aerial vehicle (UAV) attempts to access a restricted airspace by opportunistically waiting for a legitimate UAV to present its access credentials to the control tower. When the credentials are broadcasted, the malicious UAV flies to the restricted airspace knowing that the control tower cannot physically distinguish between the legitimate UAV from the malicious one. With our scheme, the legitimate UAV can physically present itself using OCC channel to distinguish itself from all other nearby UAVs.

VI. Copycat Attack & Mitigation Methods

The threat model presented in section II represents three different attack scenarios that an attacker might attempt under



(a) Receiver thinks the attacker is the sender of the RF message

(b) Receiver thinks the sender has sent the fabricated RF message

Fig. 3: RF/OCC Channel Synchronization Problem

different capabilities. However, we identified a possible attack against our authentication and localization scheme where the attacker adapts to our authentication strategy by recording visual nonces using its own camera then plays them back to mimic the legitimate sender (hence, the name Copycat), or fabricate new messages with someone else's OCC nonce.

Copycat attack can be executed under different scenarios. To mitigate its impact, we present the scenarios where this vulnerability might exist and discuss how to overcome it.

A. RF/OCC Channel Synchronization

In this scenario, the attacker attempts to exploit the time difference between the transmission time over the RF channel and the transmission time over the OCC channel in order to inject a spoofed message/nonce. To illustrate, let us consider the scenario in figure 3a. The legitimate vehicle transmits a message over the RF channel then shortly later transmits the nonce over the OCC channel. Assuming the attacker is capable of reading the RF message, the attacker extracts the nonce from the message then reproduces it over the OCC channel using its own lighting source. Now the receiver will believe that the RF message was actually sent by the attacker since the receiver correctly received the attacker's OCC nonce first.

Similarly, in the scenario demonstrated in figure 3b, the legitimate vehicle first transmits the nonce over the OCC channel then afterwards it transmits the message over the RF channel. By assuming the attacker is capable of fabricating the legitimate vehicle's identity, the attacker can read the legitimate vehicle's visual nonce and then fabricate an RF message using that nonce. Now any receiver would believe the attacker's message was actually transmitted by the legitimate vehicle since the two nonces match (attacker's RF nonce and legitimate vehicle's OCC nonce).

To understand how to mitigate this attack, we need first to define the notion of "attacker response time". Let $AttRes$ denote to the attacker response time which defines the minimum time the attacker needs to react to a message transmission. In other words, $AttRes$ is the elapsed time from the moment the attacker detects a transmitted message to the time the attacker reacts to that message by sending a spoofed message on the opposite channel. By knowing or estimating the $AttRes$ time, we can formulate a minimum acceptable RF/OCC transmission time difference such that: $|T_{OCC} - T_{RF}| < AttRes$. In other words, when one of the channels (either RF or OCC) starts transmitting, the other channel should start transmitting no later than $AttRes$ seconds. Note that since we omitted the propagation delay in section III-B, we re-used T_{RF} and T_{OCC} from section IV to denote to transmitting time instead of receiving time.

B. Sender/Receiver Channel Synchronization

In the previous scenario, the attacker either listen to the RF channel then transmit over the OCC channel, or listen to the OCC channel then transmit over the RF channel. In this scenario, the attacker listen to the OCC channel then transmit over the OCC channel as well to mimic the legitimate vehicle's visual nonce and cause confusion at the receiver side. In other words, the receiver would receive a single RF message from the legitimate vehicle but at the same time it will receive the same visual nonce from two different vehicles. Therefore, the receiver cannot distinguish who is the actual transmitter since there are two different vehicles emitting the same visual nonce.

To mitigate the impact of this attack scenario, the camera frame captured by the receiver must only include the OCC symbol produced by the sender. To achieve this property, the sender needs to delay its OCC transmission to the end of the receiver's frame exposure time window T_e. This would cause the attacker's symbol to appear in the subsequent/different frame at the receiver side. To accomplish this, the sender needs to know the sampling rate of the receiver's camera (i.e. its FPS) as well as the receiver's clock time at which the camera starts and ends its shutter exposure period for each frame. Let st_i and st_{i+1} denote to the the start time of $frame_i$ and $frame_{i+1}$ respectively. If the sender knows st_i, then it can synchronize its clock with the receiver for all subsequent frame start times such that $st_{i+1} = st_i + T_e$. Now for the sender to transmit an OCC $symbol_i$ to be captured during $frame_i$, the time to transmit $symbol_i$ is at time $T_i = st_{i+1} - (PW_s + PW_g)$. This would cause $symbol_i$ to be transmitted as close as possible to the end of $frame_i$ but not too close to risk an inter-symbol interference. Furthermore, in section III-B we discussed that the symbol pulse-width must satisfy $PW_s \geq DC_{min} \times T_e$. In this section after defining the Sender/Receiver Channel Synchronization problem, we set $PW_s = DC_{min} \times T_e$ to minimize the attack window as much as possible. Therefore, for the OCC symbol transmission to be safely and successfully captured, the transmission time for $symbol_i$ is $T_i = st_{i+1} - (PW_s + PW_g)$ and the total pulse-width is $AttRes > (PW_s + PW_g) < T_e$ for $PW_s = DC_{min} \times T_e$.

VII. Related Work

Most previous efforts related to vehicle authentication are directed toward authenticating their identities (e.g. through the use of digital certificates) [2]. However, such conventional authentication cannot be directly applied to many threat models as we have seen in section II where authenticating the location (i.e. "where you are") factor is also required. Mainly there are two localization techniques in the literature, RF-based and optical-based. Note that since we are interested in locating transmitters, techniques for localizing passive devices/objects

such as computer-vision object detection and radar will not be discussed.

RF-Based Localization has abundant number publications. It mainly relies on Time/Difference of Arrival (ToA/TDoA), Angle of Arrival (AoA) and RSSI measurements techniques. For example [9] proposed an RSSI-based localization algorithm that depends on vehicles or RSUs as observers to triangulate the target vehicle. However, the dependence on other trusted vehicles sets a limitation to the threat model. Furthermore, the location of a target vehicle is approximated as an area (a common limitation for most RF-based localization schemes). That means a group of vehicles in close-proximity with each other (a typical road scenario) would have very similar RF transmissions in terms of timing and angle which makes it a challenging task to localize one of the vehicles among the other group members. Even a single attacker can manipulate its transmission power and timing to deviate the observers. In general, RF-based localization schemes tend to be complex, require special antenna designs, and suffer from multipath and interference phenomena [11].

Vision-Based Localization is a well studied research field but mostly in the context of using visual cues as beacons for enabling devices to localize themselves. For example, [11] uses projectors and photodetectors while [12] uses LEDs and cameras as indoor localization solutions. However, such solutions cannot be repurposed to our threat model since it either require the cooperation of the localized device or require a map reconstruction phase (also found in RF-based techniques) prior to rolling out the localization system. In vehicle network, the use of VLC channels is gaining a great attention lately but mainly in the context of communication rather than localization which might make it inapplicable to our threat model. For example to increase the bitrate of the VLC channel, [13] sampled each row of a rolling shutter independently while [14] suggested the use of a standalone photodiode but both techniques make it impossible (or at least extremely difficult) to construct an image from the received VLC signal. [15] and [16] attempted to increase the VLC bitrate with a fully constructed image by utilizing visual-MIMO (Multiple Input Multiple Output) where multiple light sources act as multiple output and the individual pixels of the receiver's camera are considered the multiple inputs. However, visual-MIMO via OCC requires additional hardware and its bit error rate drastically increases with increased distances [17]. In an alternative approach, [18] implemented a visual localization scheme for drone swarms where a locator commands the target drone that need to be localized to perform a short flight maneuver such as a quick turn which would act as a visual cue for the locator who is equipped with a camera to distinguish the target drone among the other swarm members. However, the maneuver command cannot be performed until a safe distance from all surrounding objects is established. Furthermore, it is difficult for stationary vehicles to perform maneuvers such as cars stopped at a red light. Finally, accepting maneuver requests from other devices adds an additional attack surface to the autonomous system.

VIII. Conclusion & Future Work

We proposed a vision-based authentication and localization scheme for autonomous vehicles that employ localization as a form of authentication where we use visual nonces as a proof of RF message transmission. We also introduced a Copycat Attack that exploits our scheme's synchronization vulnerabilities as well as mitigation approaches for each vulnerability. In future work, we will design testbed to **(i)** numerically define $AttRes$ and PW_g then evaluate if $AttRes > (DC_{min} \times T_e) + PW_g$ can be applied using commodity cameras with an acceptable SINR, **(ii)** implement an encoding scheme that utilizes a low frame rate (e.g. 30FPS) as a receiver and high flickering frequency (e.g. 200Hz) as a transmitter without introducing the sender/receiver sync problem, and **(iii)** demonstrate a dual-channel protocol that sync between the low bitrate of the OCC channel and high bitrate of the RF channel without introducing the OCC/RF sync problem.

References

[1] Y. Cao *et al.*, "Adversarial sensor attack on lidar-based perception in autonomous driving," in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, 2019.

[2] USDOT, *Security Credential Management System (SCMS)*, 2021 (accessed January 5, 2020). https://www.its.dot.gov/resources/scms.htm.

[3] FAA, *UAS Remote Identification*, 2021 (accessed January 5, 2020). https://www.faa.gov/uas/research$_d evelopment/remote_i d/$.

[4] FAA, *Unmanned Aircraft System Traffic Management (UTM)'*, 2021 (accessed January 5, 2020). https://www.faa.gov/uas/research$_d evelopment/traffic_m anagement/$.

[5] M. D. Furtado *et al.*, "Threat analysis of the security credential management system for vehicular communications," in *2018 IEEE International Symposium on Technologies for Homeland Security (HST)*, IEEE, 2018.

[6] A. Alsoliman *et al.*, "Privacy-preserving authentication framework for uas traffic management systems," in *2020 4th Cyber Security in Networking Conference (CSNet)*, pp. 1–8, IEEE, 2020.

[7] X. Huang *et al.*, "Exposing spoofing attack on flocking-based unmanned aerial vehicle cluster: A threat to swarm intelligence," *Security and Communication Networks*, vol. 2020, 2020.

[8] S. E. Huang *et al.*, "Impact evaluation of falsified data attacks on connected vehicle based traffic signal control," *arXiv preprint arXiv:2010.04753*, 2020.

[9] M. T. Garip *et al.*, "Interloc: An interference-aware rssi-based localization and sybil attack detection mechanism for vehicular ad hoc networks," in *2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, pp. 1–6, IEEE, 2017.

[10] *IEEE Standard for Local and metropolitan area networks*. IEEE 802.15.7-2018.

[11] S. Ma *et al.*, "Foglight: Visible light-enabled indoor localization system for low-power iot devices," *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 175–185, 2017.

[12] P. Chavez-Burbano *et al.*, "Optical camera communication system for three-dimensional indoor localization," *Optik*, vol. 192, p. 162870, 2019.

[13] T. Nguyen *et al.*, "High-speed asynchronous optical camera communication using led and rolling shutter camera," in *2015 Seventh International Conference on Ubiquitous and Future Networks*, IEEE, 2015.

[14] P. H. Pathak *et al.*, "Visible light communication, networking, and sensing: A survey, potential and challenges," *IEEE communications surveys & tutorials*, vol. 17, no. 4, pp. 2047–2077, 2015.

[15] J.-E. Kim *et al.*, "Color-space-based visual-mimo for v2x communication," *sensors*, vol. 16, no. 4, p. 591, 2016.

[16] W. Yuan *et al.*, "Computer vision methods for visual mimo optical system," in *CVPR 2011 workshops*, pp. 37–43, IEEE, 2011.

[17] N. T. Le *et al.*, "A survey of design and implementation for optical camera communication," *Signal Processing: Image Communication*, vol. 53, pp. 95–109, 2017.

[18] C. Ruiz *et al.*, "Idrone: Robust drone identification through motion actuation feedback," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 2, pp. 1–22, 2018.