Transport Policy 122 (2022) 1–10

Contents lists available at ScienceDirect

Transport Policy

journal homepage: www.elsevier.com/locate/tranpol

The deceitful Connected and Autonomous Vehicle: Defining the concept, contextualising its dimensions and proposing mitigation policies

Alexandros Nikitas [a,*], Simon Parkinson [b], Mauro Vallati [b]

[a] Huddersfield Business School, University of Huddersfield, Queensgate, HD1 3DH, Huddersfield, UK
[b] School of Computing and Engineering, University of Huddersfield, Queensgate, HD1 3DH, Huddersfield, UK

ARTICLE INFO

Keywords:
Connected and autonomous vehicles
Deceitful connected and autonomous vehicles
Artificial intelligence and deceitful behaviour
Urban traffic control
Future mobility disruption

ABSTRACT

The Connected and Autonomous Vehicle (CAV) is an emerging mobility technology that may hold a paradigm-changing potential for the future of transport policy and planning. Despite a wealth of likely benefits that have made their eventual launch inescapable, CAVs may also be a source of unprecedented disruption for tomorrow's travel eco-systems because of their vulnerability to cyber-threats, hacking and misinformation. CAVs manipulated by users, traffic controllers or third parties may act in deceitful ways. This scene-setting work introduces the *deceitful CAV*, a vehicle that operates in a deceitful manner towards routing and control functionality for 'selfish' or malicious purposes and contextualises its diverse expressions and dimensions. It specifically offers a systematic taxonomy of eight distinctive deceitful behaviours namely: suppression/camouflage, overloading, mistake, substitution, target conditioning, repackaging capability signatures, amplification and reinforcing impression. These as exemplified by their most common attack forms (i.e., starvation, denial-of-service, session hijacking, man-in-the-middle, poisoning, masquerading, flooding and spoofing) are then benchmarked against five key dimensions referring to time frame (short to long duration), engagement (localised to systemic), urban traffic controller infrastructure (single to multiple components), scale (low to high), and impact (low to high). We then suggest mitigation strategies to protect CAV technology against these dangers. These span from purely technological measures referring to the machine-centric triad of vehicles, communication, and control system including adversarial training, heuristic decision algorithms and weighted voting mechanisms to human factor measures that focus on education, training, awareness enhancement, licensing and legislation initiatives that will enable users and controllers to prevent, control or report deceitful activities.

1. Introduction

Connected and Autonomous Vehicles (CAVs) are an Artificial Intelligence (AI)-based technology with the potential to change the way transportation is perceived, mobility is serviced, travel eco-systems 'behave', and cities and societies as a whole function (Fagnant and Kockelman, 2015; Kassens-Noor et la., 2020; Milakis et al., 2017; Nikitas et al., 2019; Thomopoulos et al., 2015). There are many foreseen safety, accessibility and sustainability benefits resulting from the adoption of CAVs (Xie et al., 2017), because of their ability in theory to operate human error-free (May et al., 2020) and collaboratively (Paddeu et la., 2020), ranging from accident prevention, congestion reduction and decreased carbon emissions to time savings, increased social inclusion, optimised routing, and better traffic control (Campisi et al., 2021; Kopelias et al., 2020; Nikitas et al., 2021a). However, as typical with all new technology innovations, there is the unprecedented potential of new threats (Luo et al., 2019). It is often the case with new and emerging technologies that advances in functionality are at the centre of their development, and it is only once fully adopted that the entire extent of what threats exist is known (Nikitas et al., 2020).

The severity of the threats facing CAVs has resulted in an increased effort to deepen their understanding, for instance by focusing on techniques to validate their behaviour (Ebert and Weyrich, 2019), alongside the development of new functionalities. This is also because CAVs can be seen as the synthesis of two separated complex concepts, namely connectivity and autonomy, whose fusion leads to significant synergies, but also to potentially catastrophic vulnerabilities and previously unseen threats and hazards (Ha et al., 2020). CAVs are still in their embryonic form and there are many aspects that are yet to be considered even though they have already captivated the public and media interest (Lee

* Corresponding author.
E-mail address: a.nikitas@hud.ac.uk (A. Nikitas).

https://doi.org/10.1016/j.tranpol.2022.04.011
Received 14 January 2022; Received in revised form 14 April 2022; Accepted 16 April 2022
Available online 21 April 2022

and Hess, 2020) and have turned into a competitive multi-billion business arena (Talebian and Mishra, 2018) labelled as the 'next big thing' for the future mobility paradigm (Narayanan et al., 2020).

In this work, high-level threats associated with the CAV acting in a deceitful manner towards routing and control functionality for malicious purposes is explored, with an emphasis on understanding how both technology and policy solutions are necessary to tackle this important challenge. Our particular traffic control angle is vital for planners and decision-makers because deceit can be a threat for a deterioration of the overall traffic conditions if CAVs are merely serving their individual user's aims in a myopic or selfish way (Diakaki et al., 2015) or worse serve broader malicious system-threatening purposes (Török et al., 2020).

It is widely understood that the complex machine-led and interconnected dynamics of CAVs make them more prone to data exploitation and vulnerable to cyber-attacks than any of their predecessors, increasing the risks of privacy breaches and cyber security violations for their users (Liu et al., 2020; Khan et al., 2020; Parkinson et al., 2017). This jeopardises the enhanced safety promise, that is for now possibly the most critical advantage and prime reason to prioritise CAVs over human-driven vehicles (Kim et al., 2019), and could potentially generate new layers of mistrust and non-acceptance of connected and automated mobility by the society (Chikaraishi et al., 2020; Ekman et al., 2021; Nikitas et al., 2021b).

More specifically, this new-found potential for manipulating motor traffic could give rise to adversaries seeking to deliberately disrupt the network of CAVs for malicious purposes (Qayyum et al., 2021). These adversaries could be hackers, system users, and even technology components within the urban traffic control (UTC) system. The spectrum of different adversary capabilities can range from a single individual to an organised group of individuals, and their capabilities can range from the use of existing 'off-the-shelf' tools (e.g., the script kiddy) to those willing to invest significant time and resource to expose and exploit new vulnerabilities (e.g., state sponsored). The impacts can range from those causing nuisance to existential, but irrespective they must be considered (Edgar and Manz, 2017). Security is a great concern for CAVs, largely due to the necessity to maintain safety, and how with seemingly little security exploitation, any safety property can quickly be compromised presenting a potential harm to life (Katrakazas et al., 2020; Pham and Xiong, 2021).

This article aims to set the scene and generate an interdisciplinary scientific dialogue by introducing the concept of the deceitful CAV and discussing systematically its adverse repercussions. This is a CAV that, alone or in coordination with others, behaves in an irregular way and tries to exploit the transport system for malicious purposes or for the user's (or even in some cases for the traffic controller's) benefit. This work will be one of the very first studies to introduce a conceptual framework mapping in detail the diverse dimensions of the deceitful CAV phenomenon, predicting direct and indirect effects on the city level and beyond, and proposing possible solutions technical and policy-based ones. These forecasts correspond to an accurate reading of the present state of the art and could possibly need to be revisited as the CAV reign unfolds in real time in the decades to come.

The remainder of this paper is organised as follows. Section 2 provides an overview of work on integration of AI and smart control in CAVs. Following on, the section also discusses the key vulnerabilities and weaknesses of CAVs for the UTC agenda. Section 3 defines and contextualises the concept of the deceitful CAV, while Section 4 offers a taxonomy of deceitful behaviours and relevant attacks benchmarked against key indicators like time, event scale, engagement, impact, and complexity. Section 5 comments on mitigation measures and 'safeguarding' policies going forward. Finally, Section 6 concludes this paper highlighting its key messages and setting up future research directions.

2. Research context

This section provides an encompassing overview of the field of integration of AI and smart control for CAVs. First, we discuss approaches leveraging AI to perform smart urban traffic control. Second, we review identified vulnerabilities of CAVs and smart traffic control systems.

2.1. AI and smart control in UTC in the presence of CAVs

There is a plethora of studies about the use of AI for controlling AVs in general, and about the important role that communication and collaboration play in such environment, in particular (Vallati and Chrpa, 2018). An extensive overview of the interplay between AI, CAVs and UTC is out of the scope of this paper, and the interested reader is referred to Ma et al. (2020) for an all-encompassing overview of the field. Instead, here we focus on how AI approaches, through the lens of CAVs, have been exploited to improve urban traffic management and control (UTMC), which according to Antoniou et al. (2019) is a primary concern for any city with an aspiration to be smart. A first line of work investigated techniques to perform urban traffic assignment and routing by leveraging the communication capabilities of CAVs and the presence of a centralised AI-based traffic controller. Vallati et al. (2021) exploited AI planning techniques to perform real-time traffic re-routing, while Vallati and Chrpa (2020) offered an empirical comparison of different approaches based on the same underlying idea. Finally, Motallebi et al. (2019) proposed a set of algorithms to support route assignment for CAVs.

A different body of the literature considers how junction management can be enhanced by CAVs. Several approaches investigated if the advent of CAVs is going to make traffic lights and other means for controlling traffic at junctions obsolete: CAVs can communicate and take decisions quickly and effectively, so they are in the best position to autonomously manage traffic at junctions. The interested reader is referred to a recent survey by Khayatian et al. (2020) for a comprehensive overview of this area. A less radical approach focuses on techniques for improving traffic light management by considering data provided by CAVs navigating the network. A recent survey of this type of approaches is provided by Jing et al. (2017). Examples of such techniques include Deep Reinforcement Learning (DRL) to optimise traffic flow on road intersections coupled with Smart Re-routing (SR) for the traffic approaching intersections (Mushtaq, 2021).

There is also a growing interest in improving emergency vehicle response by traffic light prioritisation to reduce delay, that can be performed in the presence of CAVs (see, for instance, Beg et al., 2021). Finally, Guo et al. (2019) studied approaches to enhance traffic signal control when both CAVs and non-CAVs are navigating the controlled region.

A third line of research, that is central for the present work, focuses on the cooperative driving of CAVs. This area concentrates on designing approaches that allow nearby vehicles to cooperate and organise themselves, with the aim of maximising the capacity of the network and reduce delays. Cooperative driving includes aspects such as platooning (Rezgui et al., 2020), lane changes (Liu et al., 2019), manoeuvring (Pirri et al., 2021), and coordinated routing (Mostafizi et al., 2021).

These areas of on-going research all focus on providing improvements in vehicle travel, and on developing technologies capable of handling a wide range of different traffic scenarios and different geographies. To achieve that, there is a need to develop a more in-depth empirical understanding of utilised, generalised, and flexible AI methods. These generalised techniques are capable of learning and dealing with a diverse and somewhat unpredictable traffic scenario portfolio. This is achieved by acquiring large volumes of data from all aspects of the infrastructure to understand the current situation, as well as make predictions and decisions to influence future control decisions. However, as with all AI mechanisms operating in diverse environments,

their flexibility results in them being susceptible to abuse, either intentionally or as unforeseen result (Parkinson et al., 2017).

2.2. Vulnerabilities of CAVs & smart UTC

The increased use of digital technologies in smart vehicles has resulted in an increased exposure to digital security threats. CAVs and smart UTC are complex technologies and as such are facing high threat levels. As presented in a recent survey, vulnerabilities arising from within the vehicle and infrastructure can be grouped into three categories: sensors, engine control units (ECUs), and communication mechanisms (Khan et al., 2020; Parkinson et al., 2017). Although this categorisation is useful when considering the point-of-origin of a vulnerability, a successful attack will likely involve a diverse range of vehicle or infrastructure components to successfully impact on vehicle control and routing. The multiplicity of attack origin makes it challenging to predict and examine the impact of a cyber-attack on all the aspects of vehicle and infrastructure functionality.

There has been significant recent interest in understanding how the specific functionality of vehicle platooning can be affected by a cyber-attack. These works often focus on defining malicious types of communication behaviour and on how they can impact on a CAV's ability to perform safe platooning. Mitigation mechanisms are often proposed in the form of understanding and monitoring a vehicle's behaviour against a known model of platooning (Wang et al., 2020). This ability to define and monitor expected behaviour is why platooning has received great attention. In Taylor et al. (2021), the authors review cybersecurity threats and open challenges specifically in platooning communication. The key confirmatory finding is that the variety and arrival of new attacks is the most significant challenge, and even with adopting a 'security by design' paradigm, unforeseen vulnerabilities will be left exposed. This problem is exacerbated with more complex control and routing operations, as their increased cooperation with a larger and more diverse range of technologies increases the potential for new vulnerabilities to be discovered. It is also the case that such functionality is less predictable and more challenging to model as the function involves control decisions that are reactive, safely handling diverse and unforeseen circumstances. In this section, we focus on reviewing and discussing known vulnerabilities reported in literature arising from inside or outside the vehicle and highlight how they may impact on vehicle control and routing within the UTC network.

Low-level communication channels, such as those controlling electro-mechanical systems, can be compromised to have deliberate impact on high-level communication channels and functionality (Amoozadeh, 2015). The authors focus on cooperative adaptive cruise control (CACC), exploring how message modification or masquerading can affect the vehicle communication stream. Here a vehicle communication stream is used to refer to network communication required for cooperative driving. The paper demonstrates that the stream manipulation can have adverse effects, and even result in collision. It has been demonstrated how positioning technologies such as RADAR, LiDAR, and GPS can all be compromised, adversely impacting on vehicle cooperation and control (van der Heijden et al., 2017). These technologies are commonplace in smart transportation systems, exposing a significant and large attack surface.

It is not just sensing technologies that are open to attack. There is a high degree of connectivity in the CAV infrastructure, creating the potential to introduce malicious technologies or compromise existing to infrastructure to achieve desired behaviour. For example, malware executing on the vehicles electronic control units (Zhang et al., 2014). Malware has great potential to compromise control components of the vehicle and infrastructure and developing malware resilient systems is essential for long-term protection (Al-Sabaawi, 2021). A central part to CAV protection and prevention against malware is the control of network communication to detect and mitigate any unwanted commands that maybe trying to exploit a vulnerability. A significant challenge with protecting a vehicle is that of considering all components to ensure that there are no weak points. For this reason, researchers have developed intrusion detection network monitoring tools that can detect unusual and potentially dangerous network activity (Loukas, 2019). Most intrusion detection techniques running on vehicle infrastructure have adopted either parameter monitoring or machine learning (Wu, 2019). However, as those conducting the survey identify, validation is limited to detecting known attack instances.

The scope of this challenge is evolving with the introduction of new communication technologies due to quickly evolving functionality and new low latency and high bandwidth network communication, placing increased demands on the network. New technologies are adopted to improve functionality, such as 5G communication mechanisms and the use of edge computing architecture to prevent centralised computing bottlenecks (Shrestha, 2020). The volume and velocity of the data presents processing challenges to maintain functionality, without considering additional security requirements/challenges. Petrillo et al. (2020) focus on the application layer where cyber-attacks modify information shared among vehicles in platooning. They investigate the impact of using messages to adversely impact on vehicle control in the response to sensory input to detect traffic behaviours. In their work, a technique is presented to protect and mitigate against network level denial of survey and bust attacks, which are primitive attacks aiming to overwhelm the network. The presented approach does not consider more sophisticated attacks whereby information is deliberately falsified to induce behaviours of a deceitful nature (Jo et al., 2019).

It is evident that there is significant body of focused research work considering the identification and mitigation focused security challenges, but technology alone will not solve problems. A combination of understanding the security threats, the exploitation potential of CAVs and infrastructure components, as well as policy interventions are required.

3. Defining the deceitful CAV

Cybersecurity is an increasingly emerging research and development agenda in future mobility (Sun et al., 2021). There is a plethora of papers, discussing ways to deceive, fool or manipulate CAVs, for example via 'toxic' signage (Sitawarin et al., 2018), purposefully crafted alterations of the environment and the sensory measurements (Modas et al., 2020) or through deceiving a vehicle's perception system (Chavarro, 2020), object detectors (Li and Velipasalar, 2020), machine learning algorithms (Porte, 2020) and GPS receivers (Sheehan et al., 2019). Nevertheless, there are still serious gaps in terms of understanding in a holistic way the full spectrum of potential CAV deception specifics, so that we can pro-actively help addressing problems that could jeopardise the safety-enhancing nature of these vehicles.

In this paper we propose the term *deceitful* CAV to encompass a vehicle that is deliberately trying to deceive the smart traffic network based on an ulterior motive. This is a relatively novel approach in contextualising a concept that is still severely understudied. Literature has briefly considered similar but less holistic terms like the rogue vehicle (i.e., a vehicle which is not behaving correctly due to faulty or compromised sensors - for example, following too closely with the vehicle in front (Abegaz et al., 2020; Gupta et al., 2020), or the deceiving vehicle (i.e., an agent aiming to reduce its own travel time by lying to other agents to gain a priority advantage especially in intersections as per Modas et al. (2020)).

The concept of the deceitful CAV goes a step further and takes under consideration the three elements that can be manipulated by hacking adversaries, anti-social or 'me-first' drivers and 'villain' or compromised transport system controllers: *vehicles*, *communication*, and *control system*. Such elements are considered by a deceitful CAV in a holistic way, that goes beyond the neighbouring vehicles or the closest junction, but instead aims at gaining advantages – or generating disruptions – at a larger scale. The deceitful CAV also incorporates the human factor (i.e.,

the user and the controller but also a third party) that could generate, help facilitating or prevent deception or rogue behaviour.

Deceitful CAVs, in principle, are those vehicles which:

(a) Exhibit rogue behaviour by intentionally misreporting or avoiding revealing their position in the road network, something that jeopardises the integrity and accuracy of the control system and creates traffic distribution imbalance causing delays and time costs for other travellers.
(b) Hide, ignore, manipulate or falsify deliberately information about their travel intentions, path choices and real-time driving decisions to get an unfair advantage over the other vehicles, travel modes and road users of the transport system (i.e., gain a faster route or get priority in an intersection).
(c) Generate and disseminate false communication or report a rogue position to trigger certain responses by the traffic controller or the other actors of the travel eco-system to create traffic disruption, distributional effects and in extreme cases chaos.

There are also complex scenarios of collective, third-party or controller-based deceiving performance, which go beyond the unitary vehicle level, and can affect in a larger scale the CAV eco-system that should be also reported.

First, a fleet of deceitful vehicles aiming to influence decision-making to their desired outcome. These CAVs provide together a sufficient volume of false information to directly influence decision-making to their advantage. For instance, a group of CAVs may decide to cooperatively deceive the controller, by following a pre-agreed path that is not the one provided by the traffic control system (TCS). In an urban region, even a small number of vehicles acting together can lead to significant congestion (maybe aimed at disrupting traffic or slowing the movement of emergency response vehicles) and can, at the same time, leave other corridors much freer to be travelled through. As an example, Vivek et al. (2019) demonstrates that a very limited number of connected vehicles disabled via internet-based attacks can severely disrupt traffic in Manhattan; if vehicles are not disabled but are cooperating for achieving a hidden agenda, the potential impact on an urban network would be dramatic. Parallels can be drawn here with crypto currencies, where the '50%' attack aims to bring about consensus in the majority involving false data, therefore verifying the false data.

Second, the man-in-the-middle attack where data is intercepted and deliberately modified between the vehicles and the controller. This could be performed by a third party to directly interfere with another vehicle's routing. For instance, an attacker can forge messages so that traffic is redirected towards some regions of the controlled network, to generate congestion, disrupt traffic, or to clear some other areas. An extreme example can be to force traffic to go near schools at start time to maximise disruption and increase the probability of accidents. However, it is a reasonable assumption that the communication channel can be easily secured. There is still, however, the need to restrict the observability of the data, as it could become a privacy concern if an adversary can inspect a vehicle's routing information. Keeping this information private between only the controller and vehicle can minimise 'costs'.

Third, the scenario of a compromised or 'villain' centralised controller where information sent to the vehicles is incorrect and has been influenced by adversary behaviour or even software error within the controller or from another controller. A controller can also be programmed to optimise some unfair metrics, for instance forcing heavy goods vehicles (HGVs) and highly polluting vehicles through a specific zone to reduce air quality levels over a sustained period, or to maintain high traffic levels in some areas to drive properties price down.

These illegitimate driving-related behaviour phenomena can have significant direct and indirect impacts on the travel eco-system, the transport infrastructure network and the city hosting these vehicles. These adverse economic, social, and environmental impacts can refer among others to:

(a) excessive congestion, traffic delays, time losses, arrival time uncertainty, bottlenecks, in certain parts of the network that could generally lead to business challenges and financial losses,
(b) environmental degradation, increased greenhouse gas emissions, local air pollution and noise nuisance in certain parts of the city,
(c) road severance, social exclusion, land devaluation, land use restructure and property market demand decline in parts of the city where traffic is diverted,
(d) reduced traffic safety for the CAV eco-system and the whole transport network and accident prevention optimisation reduction,
(e) wear and tear in motorways and road infrastructure than are used more because of this behaviour and thus extra maintenance and repair costs.

4. The dimensions and forms of the deceitful CAV behaviour

As discussed in the previous Section, the deceitful CAV is a complex and multidimensional proposition with genuinely disruptive or even hazardous potential. It reflects several different behaviours, which may have multiple distributional impacts and could apply to diverse scenarios affecting a single user, a few vehicles or a whole travel eco-system. What is lacking from the literature is an evidence-based taxonomy recognising the dimensions and expressions of deceitful behaviour characteristics and their potential severity to a transport system. Table 1 presents eight deceitful behaviours that could exist. These types are inspired by the study and classification of deceitful behaviours in military operations and provides a useful framework to classify deceitful behaviours (US Army, 2019); this is ultimately a technological

Table 1
Types of deceitful behaviour.

Deceitful Behaviour	Deception Created	Example Security Attack
Suppression/ camouflage	To make the transport network appear less congested than it really is, resulting in re-routing or false information being relayed to the user.	Starvation (Han et al., 2014; Xie et al., 2021).
Overloading	To confuse or corrupt information systems by providing repetitive data reading.	Denial-of-Service (Han et al., 2014; Xie et al., 2021).
Mistake	Transmitting erroneous data that has been generated through breach of security, negligence, or inefficiency.	Session Hijacking (Jeevitha et al., 2019)
Substitution	Covertly substituting transmitted data with false or real values, ensuring the change goes undetected.	Man-in-the-middle (Chan et al., 2014)
Target Conditioning	Conditioning the data recipient through repetitive behaviour to ensure they believe a normal course of action is being prepared, when in fact a different course of action is being prepared.	Poisoning (Wang et al., 2021)
Repackaging capability signatures	Generate new or deceptive profiles that increase or decrease the range of devices/behaviour profiles in the network, increasing ambiguity in trust.	Masquerading (Ahmed et al., 2021; Jo et al., 2019)
Amplification	To make the transport network appear more congested that it really is, resulting in re-routing or false information being relayed to the user.	Flooding (Ji et al., 2018; Müter and Asaj, 2011).
Reinforcing the impression	To make the traffic controller believe that one or more vehicles are taking one course of actions, when a different one is taken.	Spoofing (Sanders and Wang, 2020)

forecasting exercise using knowledge transfer from a more mature industry.

This classification is useful in the context of CAVs as it helps to understand the different deceitful behaviours and how the technology could be manipulated to have the desired impact. Table 1 provides an example of a specific security attack that could take place to achieve each type of deceitful behaviour. An example of *suppression* behaviour could be a *starvation* attack, whereby the aim is to prevent legitimate communication from being processed. For instance, distributing messages on a vehicle's Controller Area Network (CAN) so that Engine Control Units (ECUs) cannot receive legitimate messages by establishing communication channels that consume and block the target resource from receiving any other communication, and therefore starving it of information necessary to make informed decisions. This could, for example, take place where control units are only expecting one communication channel from a device, meaning that the real device could be blocked if the attacker can take over the communication channel (Han et al., 2014; Xie et al., 2021).

An *overloading* attack can take place through the adversary launching a Denial-of-Service attack. This is where aspects of the infrastructure are overloaded with communication to prevent any legitimate communication from being correctly interpreted. For example, communication protocols can be attacked through the repeated and high-frequency distribution of messages, to the point whereby the recipient cannot process them quickly enough and they become unresponsive to any legitimate messages (Blum et al., 2008; Hu et al., 2021).

A *mistake* could occur where a security vulnerability exists in the UTC that is a result of an incorrect configuration (Liu et al., 2020), which could create the possibility for a session hijacking attack whereby the adversary takes over a valid and authenticated communication channel (Jeevitha et al., 2019). The similar *man-in-the-middle* cyber security example is provided for the *substitution* deceitful behaviour. Here the aim is to not take over an authenticated session, rather modify network traffic to substitute data that will continue to be distributed and processed. One such example is modifying encryption keys used in the challenge-response protocol when authenticating electric vehicles at charging stations (Chan et al., 2014).

Target conditioning is a behaviour with the potential to have far reaching impacts for the control of CAVs. For example, adaptive control mechanisms might change their predictive capabilities through learning patterns, and if an adversary can deliberately introduce false patterns, they could *poison* the autonomous mechanisms to result in undesirable decisions being made. This is a particularly challenging attack to detect as it would take place over a lengthy duration to ensure that the malicious data (the 'poison') is not detected. In one recent paper, the authors discuss, and present challenges focussed on preventing poisoning attacks in electric vehicle energy prediction when using Federated Learning mechanisms to share learnings between charging stations and providers (Wang et al., 2021).

A *masquerading* attack presents a form of *repackaging capability signatures* to deceive other parts of the system. For example, in one recent publication, the authors present how ECU components can take on a disguise of another component to deliberately sabotage other systems (Ahmed et al., 2021). The prevention of such attacks is an on-going area of research, and recent efforts have been focusing on the development of robust communication protocols to prevent the attack from occurring (Jo et al., 2019).

An example of an *amplification* deceitful behaviour could originate from a *flooding* attack, whereby the attack process is like a Denial-of-Service by distributing excessive amounts of communication; however, the key difference is that there is no single target recipient and instead the communication channel is flooded to deliver adverse impacts. One example presented in research is the flooding of the CAN to the point where the rogue messages take priority (Ji et al., 2018; Müter and Asaj, 2011).

In the UTC infrastructure, the ability to deceive the controller as to the intentions or whereabouts of the vehicles could result in decisions being made that are not appropriate for the real situation. This deceitful behaviour is categorised as *reinforcing the impression* and *spoofing* is an example in terms of cyber security attack. One example in the UTC system is the potential to spoof the GPS device so that the controller believes the vehicles are in a different position (Sanders and Wang, 2020). This could allow the attacker to influence routing decisions of other vehicles to their benefit, minimising their own journey time.

As evident, based on the examples provided in Table 1 and discussed above, there is great diversity in the different deceitful behaviours and types of cyber-attacks. This results in difficulty when considering the severity and impact of each for prioritisation and comparison purposes. For better defining present and future deceitful CAV behaviours, we propose a set of dimensions that can be used to model, categorise, and quantify such deceitful behaviours.

(a) **Time Frame**: This dimension indicates the duration of the deception. A short-term deception could last for seconds or minutes, while a long-lasting deception can last for hours or even days – if performed for instance by the UTC centralised controller.
(b) **Engagement:** The deception can require the active action of an agent, for instance a vehicle deliberately providing false information to a decision-maker, or passive such as a vehicle ignoring instructions provided by a controller without actively falsifying them or hiding its own position. Further, a deception may be performed by a single agent, or it may require a cohort of agents that collaborate. Intuitively, the complexity of the attack increases with the number of vehicles that needs to be actively involved.
(c) **UTC infrastructure**: This dimension aims at defining the part of the UTC CAV-based infrastructure that is attacked and/or compromised by the deception. There are three main components that can be considered: vehicles, the controller infrastructure (the centralised controller, traffic lights, cameras, etc.), and the communication channel. One or more of the mentioned components can be exploited by a deceitful CAV.
(d) **Scale**: This dimension aims at capturing the size of the area targeted by the deceitful behaviour. This can range from a very localised disruption, such as reducing the capabilities of a sensor or of a traffic light, to more holistic disruptions that aim at modifying the way in which traffic is moving in the larger urban network.
(e) **Impact:** This dimension captures and informs the level of impact that each type of deceitful behaviour could generate. Effects generated as part of an attack can be reviewed in terms of their impact potential which can be used to prioritise any mitigation.

The proposed dimensions can provide a suitable ground to categorise and analyse deceptive behaviour, and they can also be considered to analyse different aspects of the high-level behaviours presented in Table 1. Fig. 1 presents the relationship between the example security attacks provided in Table 1 and the five dimensions. The classification has been performed by the authors based on examples identified in literature. It is highly likely that variations on attack types and their perceived importance will influence how they are placed in these dimensions; however, the use of these dimensions enables the analysis of each attack and its consideration against how it may manifest deceitful behaviours in the urban traffic infrastructure. The following discussion focusses on the scales of four dimensions (*scale, time frame, engagement, UTC infrastructure*) and how they relate to the *impact* scale.

In terms of the example attacks, each can easily be placed differently against the behavioural dimensions. For example, in terms of *scale*, it is evident that a *flooding* attack has the potential to deliver widespread systemic impact, whereas a *starvation* attack will likely only impact a small sub-component, such as ECU functionality. A large-scale attack, such as *flooding*, which is achieved by generating and transmitting high

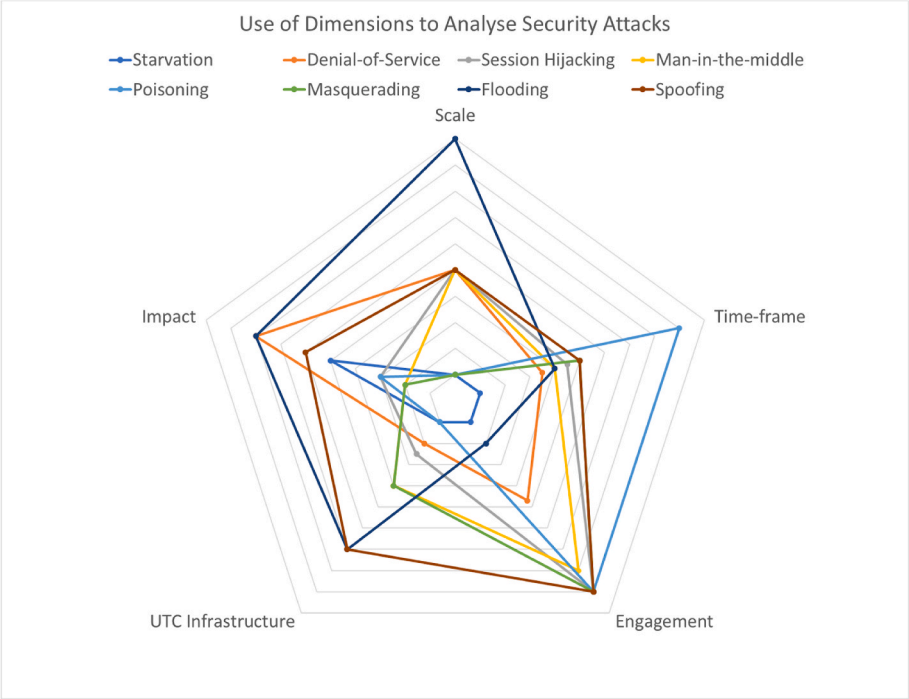

Fig. 1. Classification of example security attacks in the five dimensions. Scale: UTC infrastructure (single to multiple component involvement), scale (localised to systemic), time frame (short to long duration), engagement (low to high), and impact (low to high).

volumes of data to simulate the road network being heavily congested, has the potential to impact many vehicles that may be re-routed due to decisions taken by the controller to alleviate congestion. Although this could clear the way for the attacker to travel on the falsely reported congested route, it would inevitably result in inconvenience for many people, through extended travel times, increasing travel costs, and negative environmental factors such as increased pollution and transport noise. On the contrary, a smaller and more focused attack, such as *session hijacking*, may only affect a single vehicle and individual; however, the significance of the impact for that one user may be high. For example, rendering the vehicle completely unusable, or worse, turning into a significant risk to personal safety.

In terms of *time frame*, a *poisoning* attack will likely require considerable time to gradually introduce false information into the network to influence the learning and adaptation of autonomous capabilities, whereas a *denial-of-service* attack will take place over a shorter term as the result and impact of the attack will be quickly discovered and rectified. Understanding the relationship between *time frame* and *impact* is highly beneficial as it can help to understand which of the attacks might be more likely. For example, a *poisoning* attack is going to take considerable time to perform and may have a limited impact. This is because the attacker is going to need to input enough poisoned data to influence any learning process whilst avoiding detection. This will often result in the attacker focusing on influencing small-scale learning functions with localised impact, affecting individual infrastructure components. An attack of this nature could require significant effort to achieve an impact smaller than an easier to perform an attack, such as a *denial-of-service* attack.

Similarly, there is large variation in *engagement* whereby *spoofing* and *masquerading* attacks are likely to require in-depth knowledge of the system, whereas *starvation* and *denial-of-service* can be executed using primitive means, such as automatically distributing messages in high-volume. Attacks requiring high levels of engagement can have varying levels of impact. For example, a *spoofing* attack, where the attacker manipulates data being transferred throughout the network requires high levels of engagement as methods of data capture, modification and retransmission will need to be developed and used. This attack would likely have a focussed and high-impacting effect. On the contrary, a *masquerading* attack would also require high engagement where the attacker will be generating fake profiles in the network; however, this may only lead to a small impact whereby the infrastructure is making decisions including the disguised device, and to have a more significant *impact*, the attacker would need to create a high number of disguised devices.

The example attacks also demonstrate a large difference in how much of the *UTC infrastructure* they impact. *Flooding* could have widespread effects throughout all network communication mechanisms, requiring great effort to defend against, whereas *denial-of-service* might only impact on a subset of systems and can easily be mitigated through restricting communication flow. In terms of *impact*, there is a general observation that those involving more components of the *UTC infrastructure* will have a greater overall impact. For example, *flooding* the *UTC infrastructure* to overload the network will require overloading many *UTC infrastructure* components. This will require the attacker to be highly engaged, creating high volumes of data to be directed to many individual *UTC infrastructure* components. On the other end of the scale, an attack such as a *man-in-the-middle* attack may be focussed on only a few elements of the UTC infrastructure through intercepting and utilising communication. Further, as it has a narrow focus, the overall *impact* is likely to be lower.

5. Mitigation and policy implications

The rise of mobility automation in general, and deceitful CAVs in particular, will create a layer of unprecedented disruption concerns and problems for authorities and transport providers as discussed. These may have major societal implications and adverse repercussions for the context of cities that could dictate appropriate policy reform. Identifying, contextualising and proposing a diverse policy guide of mitigation approaches to correspond to the deception behaviours we mapped out, is the pro-active step forward for ensuring that the possible transition to driverless urban futures is less challenging. This anti-deception policy

map should include techno-centric measures referring to vehicles, communication and control systems but also socio-technical courses of action reflecting the human parameter of the CAV eco-system. This is important because on the one hand effective human-vehicle collaboration, as noted in Xing et al. (2021), is vital for the success of this technology and on the other hand, CAV eco-systems should be ultimately prioritising humans over vehicles (Nikitas et al., 2021a; Thomopoulos and Nikitas, 2019) meaning that the human factor should be explored and managed adequately. Fig. 2 provides an illustration of the mitigation measures and their relationship to the five dimensions and example attacks provided in Fig. 1.

5.1. Human factor externality measures

Many of the key anti-deception measures should be therefore education-based and information provision-centric. Even when the transition to a machine-led vehicle and infrastructure landscape is achieved the human factor will remain crucial; people, both users and controllers, intentionally or unintentionally will be able, to lesser or higher degrees, to generate, help facilitating or prevent deception or rogue behaviour. So, while Sheehan et al. (2019) justifiably proposes '*it is humans who pose* (deliberately or not) *the greatest threat to CAVs*' humans could at the same time and at any level (i.e., as a user, controller or even observer) be capable to detect, identify, prevent, fix or report CAV deceit.

People should be aware of the cyber-threat exposure they might be subjected to by a deceitful CAV and need to have some understanding and expertise about how to treat some of the challenges arising (Nagarajan et al., 2012). Therefore, anti-deception measures should include cybersecurity ***education and training*** (Beuran et al., 2018; Furman et al., 2011) and even ***user-licenses*** that could be attained after a cybersecurity training course (Liu et al., 2020). Easy to follow human-machine interaction interface (Bishop et al., 2020; Lavrov et al., 2018) and ***two-way dynamic communication*** channels to report failures and problems to the controller are also of critical importance (Liu et al., 2020). Controllers should be adequately trained to start with, completing courses that teach them how to operate according to strict ethical standards and supervised (or working in teams) at all times to avoid rogue individuals getting advantage of a 'trusting' system.

Legislation and ***enforcement*** are equally important. Legislative systems should regulate, making mandatory or highly recommended, this awareness-elevating education provision processes described above but at the same time should create 'sticks' making illegal this deception. New laws will be needed designed to predict, identify, describe and penalise appropriately all the deceitful behaviours that manipulate the transport system and have the capacity to 'harm' the individual user, the society, the built environment or the economy. Regulations built to prevent and condemn the actions of deceitful hackers, users and controllers linked with enforcement and prosecution measures spanning from imprisonment for severe actions to internet or CAV access denial should be introduced. There is the need for a principled way to verify and test CAVs and the automated traffic control infrastructure, to help minimise risks of vulnerabilities. Finally, a ***code of ethical CAV behaviour*** is needed (Nikitas et al., 2019).

5.2. Techno-centric externality measures

Machine learning and AI-based techniques that are network specific could help reducing cybersecurity risks and eliminate (or at least minimise) deceit in CAVs. ***Adversarial training***, in which a network is

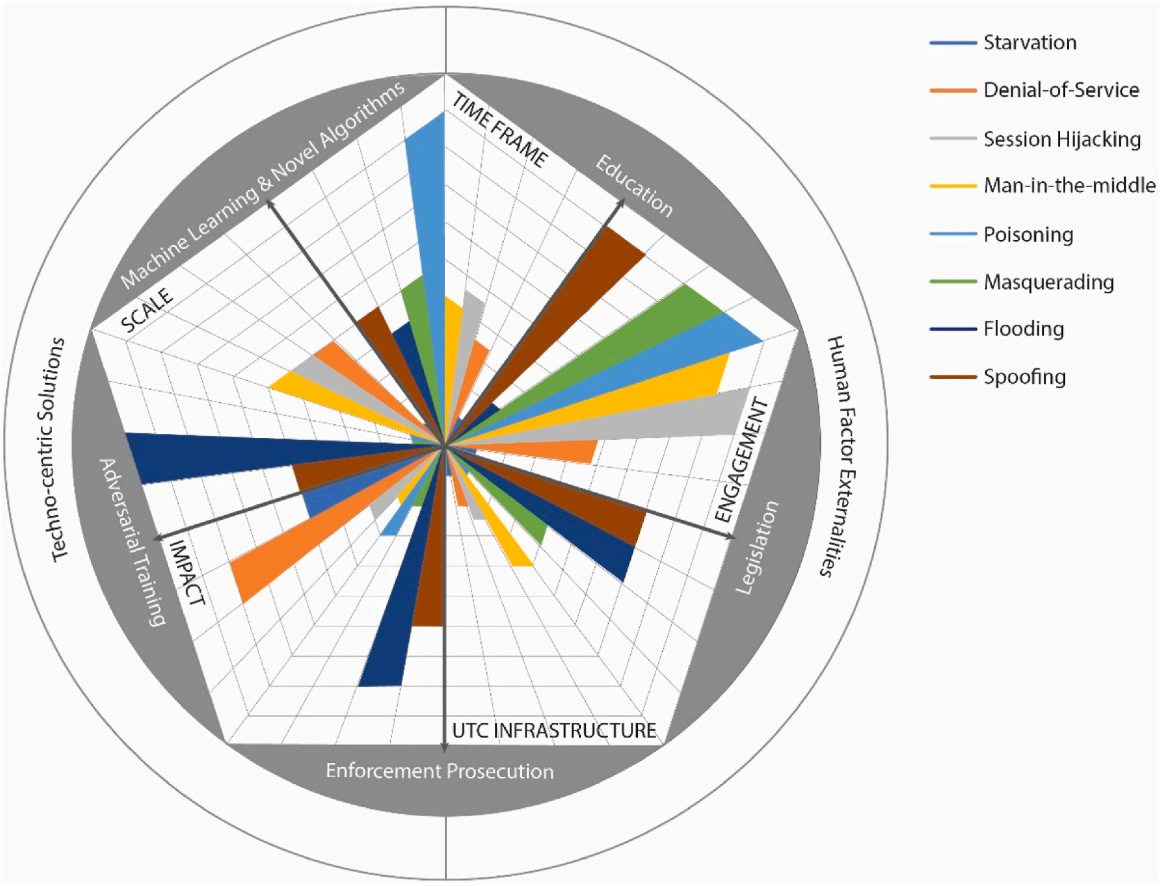

Fig. 2. A solution-based thematic map for the deceitful CAV. The eight example attack types on the five different dimensions, in addition to demonstrating their relationship with different mitigation strategies that are either of a techno-centric or human factor nature.

trained on adversarial examples by learning to operate in the presence of a destabilising adversary that applies disturbance forces to the system (Shafahi et al., 2019), is one of the few defences against adversarial attacks that has the ability to withstand strong attacks (Nagarajan et al., 2012). Event-triggered secure path tracking control for CAVs under deception attacks is another safeguard mechanism that could be employed by relevant stakeholders. This would help relieve the burden of the shareable vehicle communication network and to improve the tracking performance in the presence of deception attacks (Gu et al., 2021).

A commonly encountered attack on position sensors is one that injects false data into the position measurements; for example, GPS spoofing is a typical security concern in position measurements exploiting GPS signal, which is a type of deception attacks (Ju et al., 2020). ***Anti-spoofing technology*** is therefore needed.

For secure message dissemination in vehicular networks against insider attackers, who may tamper the content of the disseminated messages, solutions have to go beyond traditional ***message encryption*** and key management-based approaches since these might fall short to deliver (Chen et al., 2019). To help a vehicle distinguish which message is true when the messages received from multiple nearby vehicles are conflicting, Chen et al. (2019) proposes the employment of ***novel heuristic decision algorithms***, while Huang et al. (2014) suggests a solution based on weighted ***voting algorithms***. This holistic and interactive approach is based on having every CAV or a significant subset of CAVs monitor other vehicles and categorise their behaviour, building on the tradition of distributed systems research by establishing consensus (Abegaz et al., 2020). By using voting mechanisms, these monitoring CAVs can determine if a vehicle is deceitful and conduct an election to make a determination of how to respond to the deceitful vehicle (Abegaz et al., 2020). Invalid or malicious behaviour can then be reported and subsequently treated so that does not disrupt the travel eco-system.

6. Conclusions

In an era where connected and automated mobility will most likely be the mainstream standard for providing transport services, the deceitful CAV could be a force of disruption and imbalance with adverse repercussions in the individual and societal level for transport environments. This article is an interdisciplinary position contribution, combining narratives from transport planning, traffic engineering, computing and applied social science with lessons learnt from a rapidly emerging and diversifying CAV literature; a literature that is still relatively underdeveloped when it comes to understanding and contextualising the risks affiliated with the CAVs' potential ability to deceive and disrupt in unprecedent ways future travel eco-systems.

Our work aims to advance the understanding in this emerging thematic area, by introducing and thoroughly describing the deceitful CAV, which is an umbrella term that describes inappropriate CAV behaviour (and its consequences) that could be classified as: *(a) suppression/camouflage, (b) overloading, (c) mistake, (d) substitution, (e) target conditioning, (f) repackaging capability signatures, (g) amplification* and *(h) reinforcing impression*. The paper identifies the diverse dimensions and expressions of these disruptive phenomena considering their possible unwelcome impacts, and eventually generates a map-based taxonomy that contextualises critical attack examples reflecting deceitful behaviours according to their: *time frame*, *engagement*, *UTC infrastructure component, scale* (e.g., number of vehicles involved) and *impact* (how impactful a 'disruption' is for a user, few CAVs or a whole travel eco-system). We then propose mitigation strategies to safeguard CAV technology that focus not only on techno-centric measures looking at the machine-based triple of vehicles, communication, and control system but also consider and address the human factor (i.e., reflecting the user and the human controller externalities) that could generate, help facilitating or prevent deception or rogue behaviour. Our work testifies that even in an AI-centric travel eco-system it is humans who pose the greatest threat to CAVs in line with Sheehan et al. (2019) but at the same time they are the door-openers to solutions.

It should be acknowledged that the paper refers to a technology and a set of problems that are still in their infancy form. Forecasting long-term CAV technologies' adoption is thus not easy (Bansal and Kockelman, 2017); the challenges of the future could be somewhat different from what we describe in this study. CAVs and their eco-systems are a dynamic line of work with a high degree of uncertainty and unpredictability. However, setting up pro-actively the research scene for thinking as early as possible these new challenges and side-effects adds to the scientific state of art and prepares policy-makers for a number of cyber-threats and system-impacting behaviours. This could add to the overall effort that will allow us to prepare the ground for a more seamless transition to CAVs from a system of conventional human-driven and unconnected vehicles (Manivasakan et al., 2021). Our future studies will develop further this position contribution through work involving simulations, experiments, surveys, and elite interviews.

Author statements

Alexandros Nikitas: Conceptualisation, Data collection, Methodology, Analysis, Visualisation, Writing – original draft, Writing – review & editing. Simon Parkinson: Conceptualisation, Data collection, Methodology, Analysis, Visualisation, Writing – original draft, Writing – review & editing. Mauro Vallati: Conceptualisation, Data collection, Methodology, Analysis, Visualisation, Writing – original draft, Writing – review & editing.

Acknowledgements

We want to thank Mr Fernando Solis for helping us design Fig. 2. Dr Mauro Vallati was supported by a UKRI Future Leaders Fellowship [grant number MR/T041196/1].

References

Abegaz, B., Chan-Tin, E., Klingensmith, N., Thiruvathukal, G.K., 2020. Addressing rogue vehicles by integrating computer vision, activity monitoring, and contextual information. In: Proceedings Of 12th International Conference On Automotive User Interfaces And Interactive Vehicular Applications, pp. 62–64.

Ahmed, S., Juliato, M., Gutierrez, C., Sastry, M., 2021. Two-Point Voltage Fingerprinting: Increasing Detectability of ECU Masquerading Attacks. *ArXiv Preprint ArXiv*: 2102.10128.

Al-Sabaawi, A., Al-Dulaimi, K., Foo, E., Alazab, M., 2021. Addressing malware attacks on connected and autonomous vehicles: recent techniques and challenges. In: Stamp, M., Alazab, M., Shalaginov, A. (Eds.), Malware Analysis Using Artificial Intelligence and Deep Learning. Springer, Cham.

Amoozadeh, M., Raghuramu, A., Chuah, C.N., Ghosal, D., Zhang, H.M., Rowe, J., Levitt, K., 2015. Security vulnerabilities of connected vehicle streams and their impact on cooperative driving. IEEE Commun. Mag. 53 (6), 126–132.

Antoniou, G., Batsakis, S., Davies, J., Duke, A., McCluskey, T.L., Peytchev, E., et al., 2019. Enabling the use of a planning agent for urban traffic management via enriched and integrated urban data. Transport. Res. C Emerg. Technol. 98, 284–297.

Bansal, P., Kockelman, K.M., 2017. Forecasting Americans' long-term adoption of connected and autonomous vehicle technologies. Transport. Res. Pol. Pract. 95, 49–63.

Beg, A., Qureshi, A.R., Sheltami, T., Yasar, A., 2021. UAV-enabled intelligent traffic policing and emergency response handling system for the smart city. Personal Ubiquitous Comput. 25 (1), 33–50.

Beuran, R., Tang, D., Pham, C., Chinen, K.I., Tan, Y., Shinoda, Y., 2018. Integrated framework for hands-on cybersecurity training: CyTrONE. Comput. Secur. 78, 43–59.

Bishop, L.M., Morgan, P.L., Asquith, P.M., Raywood-Burke, G., Wedgbury, A., Jones, K., 2020. Examining human individual differences in cyber security and possible implications for human-machine interface design. In: International Conference On Human-Computer Interaction. Springer, Cham, pp. 51–66.

Blum, J.J., Neiswender, A., Eskandarian, A., 2008. Denial of service attacks on inter-vehicle communication networks. In: 11th International IEEE Conference On Intelligent Transportation Systems. IEEE, pp. 797–802.

Campisi, T., Severino, A., Al-Rashid, M.A., Pau, G., 2021. The development of the smart cities in the connected and autonomous vehicles (CAVs) era: from mobility patterns to scaling in cities. Infrastructure 6 (7), 100.

Chan, A.C.F., Zhou, J., 2014. Cyber–physical device authentication for the smart grid electric vehicle ecosystem. IEEE J. Sel. Area. Commun. 32 (7), 1509–1517.

Chavarro, J.M.M., Colman, E., Vásquez, L.F.N., 2020. The effect of deceiving vehicles in an autonomous intersection. In: 2020 IEEE 6th World Forum on Internet of Things (WF-IoT). IEEE, pp. 1–5.
Chen, J., Mao, G., Li, C., Zhang, D., 2019. A topological approach to secure message dissemination in vehicular networks. IEEE Trans. Intell. Transport. Syst. 21 (1), 135–148.
Chikaraishi, M., Khan, D., Yasuda, B., Fujiwara, A., 2020. Risk perception and social acceptability of autonomous vehicles: a case study in Hiroshima, Japan. Transport Pol. 98, 105–115.
Diakaki, C., Papageorgiou, M., Papamichail, I., Nikolos, I., 2015. Overview and analysis of vehicle automation and communication systems from a motorway traffic management perspective. Transport. Res. Pol. Pract. 75, 147–165.
Ebert, C., Weyrich, M., 2019. Validation of automated and autonomous vehicles. ATZelectronics Worldwide 14 (9), 26–31.
Edgar, T., Manz, D., 2017. Research Methods for Cyber Security. Syngress.
Ekman, F., Johansson, M., Karlsson, M., Strömberg, H., Bligård, L.O., 2021. Trust in what? Exploring the interdependency between an automated vehicle's driving style and traffic situations. Transport. Res. F Traffic Psychol. Behav. 76, 59–71.
Fagnant, D.J., Kockelman, K., 2015. Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. Transport. Res. Pol. Pract. 77, 167–181.
Furman, S., Theofanos, M.F., Choong, Y.Y., Stanton, B., 2011. Basing cybersecurity training on user perceptions. IEEE Secuirty. Privacy. 10 (2), 40–49.
Gu, Z., Yin, T., Ding, Z., 2021. Path tracking control of autonomous vehicles subject to deception attacks via a learning-based event-triggered mechanism. IEEE Transact. Neural Networks Learn. Syst. 32 (12), 5644–5653.
Guo, Q., Li, L., Ban, X.J., 2019. Urban traffic signal control with connected and automated vehicles: a survey. Transport. Res. C Emerg. Technol. 101, 313–334.
Gupta, M., Benson, J., Patwa, F., Sandhu, R., 2020. Secure V2V and V2I communication in intelligent transportation using cloudlets. IEEE Trans. Service. Comput. https://doi.org/10.1109/TSC.2020.3025993.
Ha, P., Chen, S., Du, R., Dong, J., Li, Y., Labi, S., 2020. Vehicle connectivity and automation: a sibling relationship. Front. Built. Environ. 199.
Han, K., Weimerskirch, A., Shin, K.G., 2014. Automotive cybersecurity for in-vehicle communication. IQT Quarterly 6 (1), 22–25.
Hu, S., Chen, Q.A., Sun, J., Feng, Y., Mao, Z.M., Liu, H.X., 2021. Automated discovery of denial-of-service vulnerabilities in connected vehicle protocols. In: 30th {USENIX} Security Symposium. USENIX Security 21.
Huang, Z., Ruj, S., Cavenaghi, M.A., Stojmenovic, M., Nayak, A., 2014. A social network approach to trust management in VANETs. Peer . Netwotk. Appl. 7 (3), 229–242.
Jeevitha, R., Bhuvaneswari, N.S., 2019. Malicious node detection in VANET session hijacking attack. In: 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT). IEEE, pp. 1–6.
Ji, H., Wang, Y., Qin, H., Wang, Y., Li, H., 2018. Comparative performance evaluation of intrusion detection methods for in-vehicle networks. IEEE Access 6, 37523–37532.
Jing, P., Huang, H., Chen, L., 2017. An adaptive traffic signal control in a connected vehicle environment: a systematic review. Information 8 (3), 101.
Jo, H.J., Kim, J.H., Choi, H.Y., Choi, W., Lee, D.H., Lee, I., 2019. MAuth-CAN: masquerade-Attack-Proof authentication for in-vehicle networks. IEEE Trans. Veh. Technol. 69 (2), 2204–2218.
Ju, Z., Zhang, H., Tan, Y., 2020. Deception attack detection and estimation for a local vehicle in vehicle platooning based on a modified UFIR estimator. IEEE Internet Things J. 7 (5), 3693–3705.
Kassens-Noor, E., Dake, D., Decaminada, T., Kotval-K, Z., Qu, T., Wilson, M., Pentland, B., 2020. Sociomobility of the 21st century: autonomous vehicles, planning, and the future city. Transport Pol. 99, 329–335.
Katrakazas, C., Theofilatos, A., Papastefanatos, G., Härri, J., Antoniou, C., 2020. Cyber Security and its Impact on CAV Safety: Overview, Policy Needs and Challenges. Policy Implications of Autonomous Vehicles. Elsevier, The Netherlands, pp. 73–94.
Khan, S.K., Shiwakoti, N., Stasinopoulos, P., Chen, Y., 2020. Cyber-attacks in the next-generation cars, mitigation techniques, anticipated readiness and future directions. Accid. Anal. Prev. 148, 105837.
Khayatian, M., Mehrabian, M., Andert, E., Dedinsky, R., Choudhary, S., Lou, Y., Shirvastava, A., 2020. A survey on intersection management of connected autonomous vehicles. ACM Trans. Cyber Phys. Syst. 4 (4), 1–27.
Kim, M.K., Park, J.H., Oh, J., Lee, W.S., Chung, D., 2019. Identifying and prioritizing the benefits and concerns of connected and autonomous vehicles: a comparison of individual and expert perceptions. Res. Trans. Business Manage. 32, 100438.
Kopelias, P., Demiridi, E., Vogiatzis, K., Skabardonis, A., Zafiropoulou, V., 2020. Connected & autonomous vehicles–Environmental impacts–A review. Sci. Total Environ. 712, 135237.
Lavrov, E.A., Volosiuk, A.A., Pasko, N.B., Gonchar, V.P., Kozhevnikov, G.K., 2018. Computer simulation of discrete human-machine interaction for providing reliability and cybersecurity of critical systems. In: 2018 Third International Conference on Human Factors in Complex Technical Systems and Environments (ERGO) S and Environments (ERGO). IEEE, pp. 67–70.
Lee, D., Hess, D.J., 2020. Regulations for on-road testing of connected and automated vehicles: assessing the potential for global safety harmonization. Transport. Res. Pol. Pract. 136, 85–98.
Li, Y., Velipasalar, S., 2020. Weighted Average Precision: Adversarial Example Detection in the Visual Perception of Autonomous Vehicles, 03751. *arXiv preprint arXiv*, 2002.
Liu, J., Xiao, Y., Wu, J., 2019. From ai to ci: a definition of cooperative intelligence in autonomous driving. In: International Conference on Internet of Vehicles. Springer, Cham, pp. 64–75.
Liu, N., Nikitas, A., Parkinson, S., 2020. Exploring expert perceptions about the cyber security and privacy of Connected and Autonomous Vehicles: a thematic analysis approach. Transport. Res. F Traffic Psychol. Behav. 75, 66–86.
Loukas, G., Karapistoli, E., Panaousis, E., Sarigiannidis, P., Bezemskij, A., Vuong, T., 2019. A taxonomy and survey of cyber-physical intrusion detection approaches for vehicles. Ad Hoc Netw. 84, 124–147.
Luo, Q., Cao, Y., Liu, J., Benslimane, A., 2019. Localization and navigation in autonomous driving: threats and countermeasures. IEEE Wireless Commun. 26 (4), 38–45.
Ma, Y., Wang, Z., Yang, H., Yang, L., 2020. Artificial intelligence applications in the development of autonomous vehicles: a survey. IEEE/CAA J. Automatica Sinica 7 (2), 315–329.
Manivasakan, H., Kalra, R., O'Hern, S., Fang, Y., Xi, Y., Zheng, N., 2021. Infrastructure requirement for autonomous vehicle integration for future urban and suburban roads–Current practice and a case study of Melbourne, Australia. Transport. Res. Pol. Pract. 152, 36–53.
May, A.D., Shepherd, S., Pfaffenbichler, P., Emberger, G., 2020. The potential impacts of automated cars on urban transport: an exploratory analysis. Transport Pol. 98, 127–138.
Milakis, D., Van Arem, B., Van Wee, B., 2017. Policy and society related implications of automated driving: a review of literature and directions for future research. J. Intelligent. Trans. Syst. 21 (4), 324–348.
Modas, A., Sanchez-Matilla, R., Frossard, P., Cavallaro, A., 2020. Toward robust sensing for autonomous vehicles: an adversarial perspective. IEEE Signal Process. Mag. 37 (4), 14–23.
Mostafizi, A., Koll, C., Wang, H., 2021. A decentralized and coordinated routing algorithm for connected and autonomous vehicles. IEEE Trans. Intell. Transport. Syst. 1–13.
Motallebi, S., Xie, H., Tanin, E., Qi, J., Ramamohanarao, K., 2019. Streaming route assignment for connected autonomous vehicles (systems paper). In: Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 408–411.
Mushtaq, A., Haq, I.U., Imtiaz, M.U., Khan, A., Shafiq, O., 2021. Traffic flow management of autonomous vehicles using deep reinforcement learning and smart rerouting. IEEE Access 9, 51005–51019.
Müter, M., Asaj, N., 2011. Entropy-based anomaly detection for in-vehicle networks. In: 2011 IEEE Intelligent Vehicles Symposium, vol. IV. IEEE, pp. 1110–1115.
Nagarajan, A., Allbeck, J.M., Sood, A., Janssen, T.L., 2012. Exploring game design for cybersecurity training. In: 2012 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER). IEEE, pp. 256–262.
Narayanan, S., Chaniotakis, E., Antoniou, C., 2020. Shared autonomous vehicle services: a comprehensive review. Transport. Res. C Emerg. Technol. 111, 255–293.
Nikitas, A., Njoya, E.T., Dani, S., 2019. Examining the myths of connected and autonomous vehicles: analysing the pathway to a driverless mobility paradigm. Int. J. Automot. Technol. Manag. 19 (1–2), 10–30.
Nikitas, A., Michalakopoulou, K., Njoya, E.T., Karampatzakis, D., 2020. Artificial intelligence, transport and the smart city: definitions and dimensions of a new mobility era. Sustainability 12 (7), 2789.
Nikitas, A., Thomopoulos, N., Milakis, D., 2021a. The environmental and resource dimensions of automated transport: a nexus for enabling vehicle automation to support sustainable urban mobility. Annu. Rev. Environ. Resour. 46, 167–192.
Nikitas, A., Vitel, A.E., Cotet, C., 2021b. Autonomous vehicles and employment: an urban futures revolution or catastrophe? Cities 114, 103203.
Paddeu, D., Shergold, I., Parkhurst, G., 2020. The social perspective on policy towards local shared autonomous vehicle services (LSAVS). Transport Pol. 98, 116–126.
Parkinson, S., Ward, P., Wilson, K., Miller, J., 2017. Cyber threats facing autonomous and connected vehicles: future challenges. IEEE Trans. Intell. Transport. Syst. 18 (11), 2898–2915.
Petrillo, A., Pescape, A., Santini, S., 2020. A secure adaptive control for cooperative driving of autonomous connected vehicles in the presence of heterogeneous communication delays and cyberattacks. IEEE Trans. Cybern. 51 (3), 1134–1149.
Pham, M., Xiong, K., 2021. A Survey on Security Attacks and Defense Techniques for Connected and Autonomous Vehicles. Computers & Security, p. 102269.
Pirri, P., Pahl, C., El Ioini, N., Barzegar, H.R., 2021. Towards cooperative maneuvering simulation: tools and architecture. In: 2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC). IEEE, pp. 1–6.
Porte, M., Dey, S., Joshi, A., Hespanhol, P., Aswani, A., Johnson-Roberson, M., Vasudevan, R., 2020. Detecting deception attacks on autonomous vehicles via linear time-varying dynamic watermarking. In: 2020 IEEE Conference on Control Technology and Applications (CCTA). IEEE, pp. 1–8.
Qayyum, A., Usama, M., Qadir, J., Al-Fuqaha, A., 2020. Securing connected & autonomous vehicles: challenges posed by adversarial machine learning and the way forward. IEEE Commun. Survey. Tutorial. 22 (2), 998–1026.
Rezgui, J., Gagné, É., Blain, G., St-Pierre, O., Harvey, M., 2020. Platooning of autonomous vehicles with artificial intelligence V2I communications and navigation algorithm. In: 2020 Global Information Infrastructure and Networking Symposium (GIIS). IEEE, pp. 1–6.
Sanders, C., Wang, Y., 2020. Localizing spoofing attacks on vehicular GPS using vehicle-to-vehicle communications. IEEE Trans. Veh. Technol. 69 (12), 15656–15667.
Shafahi, A., Najibi, M., Ghiasi, A., Xu, Z., Dickerson, J., Studer, C., et al., 2019. Adversarial Training for Free!. *arXiv preprint arXiv:1904.12843*.
Sheehan, B., Murphy, F., Mullins, M., Ryan, C., 2019. Connected and autonomous vehicles: a cyber-risk classification framework. Transport. Res. Pol. Pract. 124, 523–536.
Shrestha, R., Nam, S.Y., Bajracharya, R., Kim, S., 2020. Evolution of V2X communication and integration of blockchain for security enhancements. Electronics 9 (9), 1338.

Sitawarin, C., Bhagoji, A.N., Mosenia, A., Chiang, M., Mittal, P., 2018. Darts: Deceiving Autonomous Cars with Toxic Signs. *arXiv preprint arXiv*, 1802.06430.

Sun, X., Yu, F.R., Zhang, P., 2021. A survey on cyber-security of connected and autonomous vehicles (CAVs). IEEE Trans. Intell. Transport. Syst. https://doi.org/10.1109/TITS.2021.3085297.

Talebian, A., Mishra, S., 2018. Predicting the adoption of connected autonomous vehicles: a new approach based on the theory of diffusion of innovations. Transport. Res. C Emerg. Technol. 95, 363–380.

Taylor, S.J., Ahmad, F., Nguyen, H.N., Shaikh, S.A., Evans, D., Price, D., 2021. Vehicular platoon communication: cybersecurity threats and open challenges. In: 2021 51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W). IEEE, pp. 19–26.

Thomopoulos, N., Givoni, M., 2015. The autonomous car—a blessing or a curse for the future of low carbon mobility? An exploration of likely vs. desirable outcomes. Eur. J. For. Res. 3 (1), 1–14.

Thomopoulos, N., Nikitas, A., 2019. Smart urban mobility futures: editorial for special issue. Int. J. Automot. Technol. Manag. 19 (1–2), 1–9.

Török, Á., Szalay, Z., Uti, G., Verebélyi, B., 2020. Modelling the effects of certain cyber-attack methods on urban autonomous transport systems, case study of Budapest. J. Ambient Intell. Hum. Comput. 11 (4), 1629–1643.

Vallati, M., Chrpa, L., 2018. A principled analysis of the interrelation between vehicular communication and reasoning capabilities of autonomous vehicles. In: 2018 21st International Conference on Intelligent Transportation Systems (ITSC). IEEE, pp. 3761–3766.

Vallati, M., Chrpa, L., 2020. Reducing traffic congestion in urban areas via real-time Re-routing: a simulation study. In: Australasian Joint Conference on Artificial Intelligence. Springer, Cham, pp. 69–81.

Vallati, M., Scala, E., Chrpa, L., 2021. A hybrid automated planning approach for urban real-time routing of connected vehicles. In: 24th IEEE International Conference on Intelligent Transportation. IEEE.

van der Heijden, R., Lukaseder, T., Kargl, F., 2017. Analyzing attacks on cooperative adaptive cruise control (CACC). In: 2017 IEEE Vehicular Networking Conference (VNC). IEEE, pp. 45–52.

Vivek, S., Yanni, D., Yunker, P.J., Silverberg, J.L., 2019. Cyberphysical risks of hacked internet-connected vehicles. Phys. Rev. 100 (1), 012316.

Wang, P., Wu, X., He, X., 2020. Modeling and analyzing cyberattack effects on connected automated vehicular platoons. Transport. Res. C Emerg. Technol. 115, 102625.

Wang, W., Fida, M.H., Lian, Z., Yin, Z., Pham, Q.V., Gadekallu, T.R., Su, C., 2021. Secure-enhanced Federated Learning for Ai-Empowered Electric Vehicle Energy Prediction. IEEE Consumer Electronics Magazine.

Wu, W., Li, R., Xie, G., An, J., Bai, Y., Zhou, J., Li, K., 2019. A survey of intrusion detection for in-vehicle networks. IEEE Trans. Intell. Transport. Syst. 21 (3), 919–933.

Xie, Y., Gartner, N.H., Chowdhury, M., 2017. Editors' notes: special issue on connected and autonomous vehicles. Transport. Res. C Emerg. Technol. 100 (78), 34–36.

Xie, Y., Zhou, Y., Xu, J., Zhou, J., Chen, X., Xiao, F., 2021. Cybersecurity Protection on In-vehicle Networks for Distributed Automotive Cyber-physical Systems: State-of-the-art and Future Challenges. Practice and Experience, Software.

Xing, Y., Lv, C., Cao, D., Hang, P., 2021. Toward human-vehicle collaboration: review and perspectives on human-centered collaborative automated driving. Transport. Res. C Emerg. Technol. 128, 103199.

Zhang, T., Antunes, H., Aggarwal, S., 2014. Defending connected vehicles against malware: challenges and a solution framework. IEEE Internet Things J. 1 (1), 10–21.