IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received June 29, 2021, accepted July 8, 2021, date of publication July 14, 2021, date of current version July 26, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3097146

Context-Aware Security for Vehicles and Fleets: A Survey

DANIEL GRIMM, MARCO STANG, AND ERIC SAX
Institute for Information Processing Technologies (ITIV), Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany
Corresponding author: Daniel Grimm (daniel.grimm@kit.edu)

ABSTRACT Vehicles are becoming increasingly intelligent and connected. Interfaces for communication with the vehicle, such as WiFi and 5G, enable seamless integration into the user's life, but also cyber attacks on the vehicle. Therefore, research is working on in-vehicle countermeasures such as authentication, access controls, or intrusion detection. Recently, legal regulations have also become effective that require automobile manufacturers to set up a monitoring system for fleet-wide security analysis. The growing amount of software, networking, and the automation of driving create new challenges for security. Context-awareness, situational understanding, adaptive security, and threat intelligence are necessary to cope with these ever-increasing risks. In-vehicle security should be adaptive to secure the car in an infinite number of (driving) situations. For fleet-wide analysis and alert triage, knowledge and understanding of the circumstances are required. Context-awareness, nonetheless, has been sparsely considered in the field of vehicle security. This work aims to be a precursor to context-aware, adaptive and intelligent security for vehicles and fleets. To this end, we provide a comprehensive literature review that analyzes the vehicular as well as related domains. Our survey is mainly characterized by the detailed analysis of the context information that is relevant for vehicle security in the future.

INDEX TERMS Automotive engineering, context awareness, context modeling, fleet security, intelligent vehicles, security, situation awareness, threat intelligence, vehicles.

I. INTRODUCTION

In the course of the ever-increasing networking of electronic devices with each other and with the Internet, as in the areas of Internet of Things, Industry 4.0, Smart Grids, Smart Cities, and Avionics, this development is particularly noticeable in the automotive industry. Cyber-physical systems (CPS) that control and monitor their physical environment are becoming equal components of the World Wide Web, alongside the traditional personal IT devices and business networks. The growing interdependence of vehicles is fueled by the introduction of networking standards such as 5G and Wi-Fi. Thus, vehicles connected to the Internet present a prominent example of CPS. However, with every little opening of electronic devices for communication with the outside world, the system is wide open to cyberattacks. Especially with CPS such as vehicles, hacks and vulnerabilities can have life-threatening consequences for the users or massive financial implications for the manufacturers [1]–[3].

The automotive industry's most important problem is to ensure that fleet-wide attacks cannot take place, or at least only with minimal consequences [1], [5]. Nobody wants to imagine a scenario in which a whole fleet of vehicles is taken over remotely and used as a terrorist weapon, for example. Less obvious than the physical threat of a fleet-wide attack, mass theft of vehicles and even the sale of stolen user data is a problem, since there is a substantial financial benefit for attackers [1], [6]. Additionally, the in-vehicle network and the number of vehicle functions are growing, which manifests, for example, in 150 electronic control units (ECUs) today and an estimated 300 million lines of code in 2030 [6], which are required for current trends such as autonomous driving. Together with the various external interfaces, e.g. cellular networks, smartphones, Wi-Fi, Vehicle-to-Vehicle(V2V), Vehicle-to-Infrastructure(V2I), this constitutes an increasing threat surface and a growing potential for vulnerabilities. An overview of this ecosystem of the internal and external networks is shown in Figure 1. For simplicity, we focus on connectivity in this illustration and neglect, for example, the sensor technology that is required

The associate editor coordinating the review of this manuscript and approving it for publication was Jing Yan.

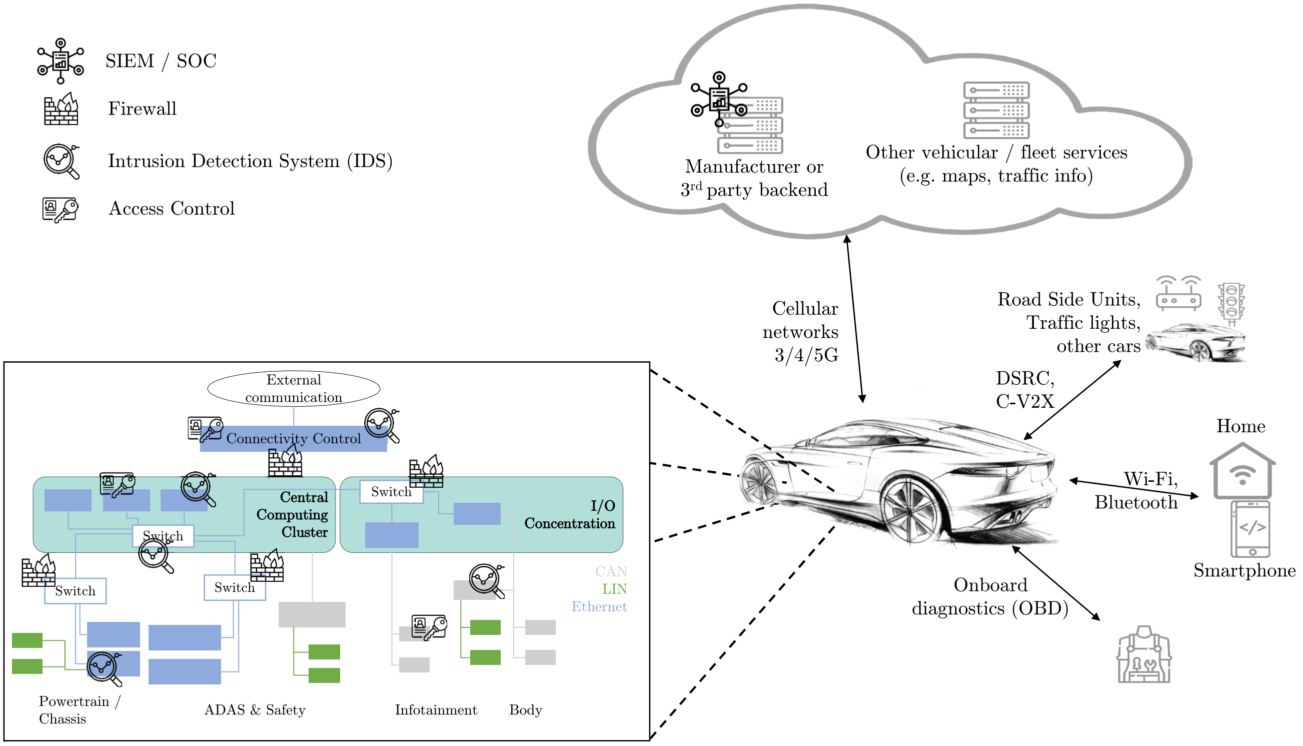

FIGURE 1. The connected vehicle, its internal network, and common security measures. Schematic illustration of the vehicle internal network adapted from [4]. Car image from Jaguar MENA[1], icons from flaticon[2].

for autonomous driving (i.e. cameras, radar, GPS, etc.). The section on the bottom left represents the internal network, consisting of ECUs of differing performance and several networking technologies (see Section III-A for more details). Each rectangular block symbolizes an ECU, and the color indicates their main networking technology. On the right side of the picture the external communication of the vehicle via the aforementioned interfaces is shown. In our example, the *Connectivity Control* ECU is responsible to exchange data with the manifold external communication partners (e.g. diagnostic, web servers, infrastructure).

Fleets of connected cars shared with multiple users and operating increasingly autonomously represent a viable target for hackers. In a recent report [1], KPMG defined a fleet from a cybersecurity perspective as ''A group of individual vehicles that connects to a common technology platform through shared operating systems, software, or hardware'' [1] With this definition, it becomes apparent that the term ''fleet'' encompasses even more risks than securing the vehicles of one vehicle manufacturer, for instance ever since standardized operating systems, e.g. based on the platforms of the AUTOSAR (AUTomotive Open Systems ARchitecture) foundation [7], have been introduced.

Having in mind physical and financial impacts, just closing vulnerabilities and fixing bugs with updates after malicious incidents is like shutting the stable door after the horse has bolted and is therefore not a viable solution. Instead, the cybersecurity solutions of vehicles and fleets should be (pro)active and react flexibly to the ever-changing threat landscape. This includes measures within the vehicle as well as holistic monitoring from a central backend. Inside the vehicle, in addition to well-known measures such as encryption and authentication, in literature and industry there exist active monitoring and reaction measures such as firewalls, intrusion detection, and intrusion prevention systems (IDS/IPS) as well as access controls. IDS for the automotive sector are currently on the way to standardization, an approach is published in version R20-11 of the AUTOSAR standard. On the other hand, access controls have been part of AUTOSAR since release R19-11 in the *Identity and Access Management* component [8]. In terms of the entire fleet, the elementary tool is the Vehicle Security Operations Center (V-SOC) [6], containing measures to correlate and analyze the data with Machine Learning [9] in e.g. a Security Information and Event Management (SIEM) System, which serves as the cornerstone of the V-SOC. As an example, some instances of these measures in the vehicle and backend are shown in Figure 1 (see Section III-C for more details on the security measures). In the vehicle, we may find multiple instances of IDS, firewalls and access controls. The SIEM and SOC systems are located in a backend server. For a security concept

[1] Jaguar C-X16 Design (2011) https://www.flickr.com/photos/56721991@N06/6213603198, accessed on 2021-03-01

[2] https://www.flaticon.com/

for a real vehicle, additional components such as encryption and hardware security would have to be added, which are not presented here for simplicity. Security measures in the vehicle act as *security sensors* in a network monitoring strategy [10] that provide the SIEM with log files and events. Security sensors shall document the actions of and on the system where they are deployed, e.g. if something suspicious was recognized. This data is transmitted to the SIEM via the vehicle's Internet connection, e.g. via the mobile network.

The recent legal regulation of UNECE makes monitoring and responding to cyberattacks on vehicles and their ecosystem a mandatory requirement to the homologation of new vehicle types [6], [11]. This will foster the widespread adoption of V-SOCs to ensure procedural and technical capabilities for monitoring and response. According to the UNECE regulation, a cyber risk management strategy shall be implemented by the manufacturers to be able to react to cyber threats and vulnerabilities along the entire lifecycle. To adhere to the regulations, new standards such as ISO/SAE 21434 [12] will define the common best practices and technical requirements and serve as the basis for assessing cybersecurity of new vehicles. The technical and organizational implementation along the value-added chain is of course still left to the manufacturers and their suppliers.

Both the in-vehicle security measures as well as the SIEM are responsible for analyzing data, behavior, or software actions and distinguish between benign and malicious events. Yet, both the decisions of the onboard measures and the correlation procedures of the SIEM face a severe problem: what is unusual but harmless, and what is suspicious or an actual attack depends on the context in which the vehicle or fleet are situated. To give just one example: a Parking Pilot System (PPS) as well as a Lane Keep Assistance System (LKAS) in general need to have access to the steering control of the vehicle. However, the PPS should only be allowed to send any control to the steering computer on the vehicle-internal network if we are about to park, whereas the LKAS should only be allowed to do the same on highways. Holle *et al.* [13] mention the increased dynamics of in-vehicle communication through the introduction of service-oriented architectures, updates over the air, or function on demand. As a result, trust boundaries and policies are dependent on the vehicle context. Static firewalls, for example, will therefore no longer be sufficient in the future. For service-oriented architectures, Gehrmann and Duplys [14] argue that an Intrusion Detection System should support dynamic reconfiguration and rule set adjustment. Context-sensitive IDS that take into account the data semantics of onboard networks are seen as a promising direction, although they have not been given much consideration in research so far [15].

Context-awareness and adaptiveness to situational constraints are thus highly compelling capabilities for security measures. It is not straightforward, though, to specify the relevant security context of vehicles and fleets, and it is a challenge to infer the current context and turn it into an instrument of automotive cybersecurity. To the best of our knowledge, literature does not provide an overview of context-aware vehicle security. Furthermore, contextual information is not yet generally considered as a tool for vehicle and fleet security. Only a few publications deal with the integration of context into tailored solutions, mentioning only exemplary context information that might be relevant for the specific application. Given the novelty of context-awareness in this field, we formulate a definition of the automotive security context. Our main contribution is to provide a comprehensive survey of context-awareness for vehicle and fleet security, taking into account related fields of work such as corporate IT and various industries. We differ significantly from other surveys, as this study is intended to be particularly beneficial for the application area of automotive security. Given the interdependence of corporate IT and the vehicle network, it should be emphasized that we also consider relevant work on context and security beyond the vehicle ecosystem. In addition, we define a new taxonomy, that describes the main categories of information that are context for automotive security. With this, we aim to pave the way for future research aimed at leveraging contextual information in security measures, processes, and tools.

The remainder of this work is structured as follows (see Figure 2 for a graphical overview of the paper's organization): to clarify our contribution's value, we continue to differentiate our survey from others in Section II and summarize the necessary background on vehicle and fleet security in Section III. Afterward, we define the automotive security context and emphasize the challenges the automotive industry faces in developing effective and efficient security measures and why contextual information is crucial for their solution (see Section IV). This work's central contributions are Section V and VI, surveying, summarizing, and comparing the relevant publications on context-awareness for automotive security purposes. An in-detail analysis identifies open challenges and research gaps. Finally, Section VII concludes this paper with a summary and outlook.

II. RELATED SURVEYS

Perera *et al.* [16] provide an extensive survey for context-awareness in Internet of Things applications. However, they only consider context-aware systems' security and privacy, e.g. secure data collection and storage and policies to access context information. Since various personally identifiable information is contained in the raw data and the abstracted contextual information, this is an essential component of context-aware systems. Yet, the usage of context information in security measures is out of their scope. The work of Li *et al.* [17] also covers security for context-aware systems. They demand encryption and monitoring of the information to identify anomalies and unauthorized access. Anyhow, this field is not the focus of our work and additionally, the aforementioned publications are not aiming at the automotive industry.

Various publications that provide an overview of automotive security cover different security measures, from

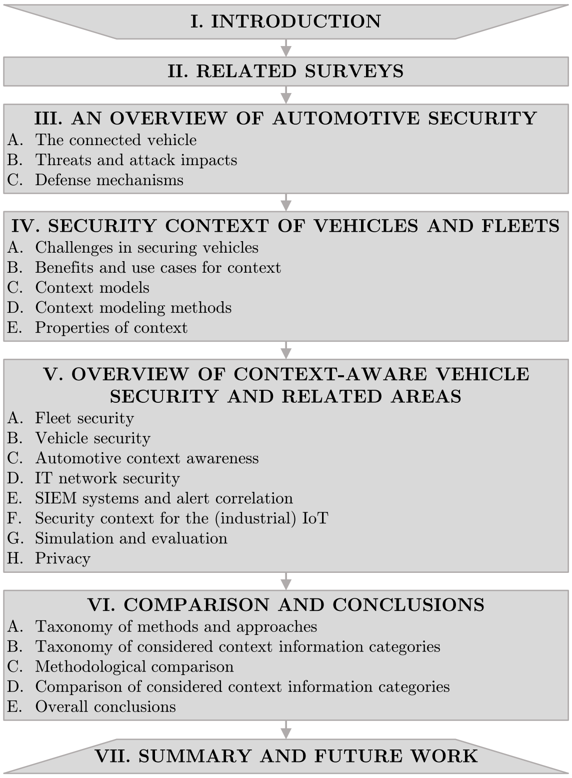
I. INTRODUCTION
II. RELATED SURVEYS
III. AN OVERVIEW OF AUTOMOTIVE SECURITY
A. The connected vehicle
B. Threats and attack impacts
C. Defense mechanisms
IV. SECURITY CONTEXT OF VEHICLES AND FLEETS
A. Challenges in securing vehicles
B. Benefits and use cases for context
C. Context models
D. Context modeling methods
E. Properties of context
V. OVERVIEW OF CONTEXT-AWARE VEHICLE SECURITY AND RELATED AREAS
A. Fleet security
B. Vehicle security
C. Automotive context awareness
D. IT network security
E. SIEM systems and alert correlation
F. Security context for the (industrial) IoT
G. Simulation and evaluation
H. Privacy
VI. COMPARISON AND CONCLUSIONS
A. Taxonomy of methods and approaches
B. Taxonomy of considered context information categories
C. Methodological comparison
D. Comparison of considered context information categories
E. Overall conclusions
VII. SUMMARY AND FUTURE WORK

FIGURE 2. Content and structure of this work.

specific hardware as a root of trust to encryption, intrusion detection and prevention systems, firewalls, and resilient architectures [15], [18]–[22]. But none of these mentions context-awareness of the measures, except for Al-Jarrah *et al.* [15] (see Section V-B). [15] and [18] focus on surveying Intrusion Detection Systems, while [19]–[22] provide a broader overview of security measures. Publications [21], [22] discuss several security measures and also methods that affect back-end vehicle connectivity, but are not a survey intended to provide a comprehensive review of related literature. In comparison to [19], publication [20] remains limited to the topic of updates in the area of fleet security. In contrast, our work surveys context information for both the fleet and the in-vehicle defense mechanisms and includes publications from the IT and IoT domains.

The survey of Vahdat-Nejad *et al.* [23] provides an overview of context-awareness in vehicular networks. Although they discuss VANETs (Vehicular Ad-hoc NETworks), which require vehicles with external V2V interfaces, the word *security* is not mentioned. Fernandez-Rojas *et al.* [24] present the most recent study on vehicle context-awareness, and it includes a section on security and data management. As such, their work is valuable for comparing it to ours. The section on security, though, does not contain works that are specifically targeted at the automotive industry. We defer to their work for a broader overview of context-awareness for vehicles, while we look specifically at the area of security for automotive systems.

TABLE 1. Comparison of related surveys with our work. ●: Work covers this topic, ○: Work covers this topic only partially.

Publication	Context	Security measures	Vehicle	Fleet	IT, IoT
[16]	●	○			●
[17]	●	○			
[15]	○	○	●		
[18]		○	●		
[21]		●	○	○	
[22]		●	○	○	
[19]		●	●	●	
[20]		●	●	○	
[23]	●			○	
[24]	●	○	○	○	
Our work	●	●	●	●	●

In summary, our work differs in essential aspects from previous work (see Table 1). Work that covers the breadth of automotive security measures has so far not sufficiently addressed the topic of context-awareness, where we make a decisive contribution. Work on context-awareness has so far only partially dealt with vehicles or their security. We take up the essential preliminary work and use it to supplement the selection of the work considered here.

III. AN OVERVIEW OF AUTOMOTIVE SECURITY

New demands on the vehicle, such as automated driving functions, are driving growth in software, processing capacity, and networking. We present the internal and external networks main components in Section III-A. However, we focus on the network technologies, interfaces and ECU platforms and omit details on e.g. vehicle sensors and automated driving functions, which are already captured in a vast amount of literature, e.g. [25], [26]. With networking and increasing functionality, the risk for attacks increases, which manifests in the steadily growing numbers (see Section III-B). For this reason, various defense measures are the subject of research in the automotive sector (see Section III-C).

A. THE CONNECTED VEHICLE

In Figure 1, an exemplary Electrical/Electronical (E/E) architecture of a future vehicle and its connections to external systems are shown. The E/E architecture consists of the various ECUs, the operating systems and software running on them, and the different bus and network systems that interconnect the ECUs.

1) INTERNAL NETWORK

The most relevant bus system technologies are Controller Area Network (CAN), Local Interconnect Network (LIN), FlexRay, Media Oriented Systems Transport

(MOST), and Automotive Ethernet. Each technology has its specific application areas, which differ in terms of required data rates and the need for real-time guarantees. To name some examples, CAN supports data rates of up to $1 Mbit/s$ (High-Speed CAN), $5 - 8 Mbit/s$(CAN-FD) and $10 Mbit/s$ (CAN XL). With Ethernet, data rates of $10 Mbit/s$ (10BASE-T1S [27]), $100 Mbit/s$ (100BASE-T1 [28]), $1 Gbit/s$ (1000BASE-T1 [29]) and even more are possible. Along with the higher bandwidth offered by Ethernet networks, the physical transceivers' cost is higher, and more processing power is required on behalf of the ECUs. In current vehicles, multiple instances of each of these technologies can be used to achieve an optimum of cost and required data rates. For example, the E/E architecture of a 2019 Audi A8 consists of seven CAN buses, one MOST bus, and one FlexRay [30], besides numerous LIN buses.

Future architectures, such as the one in Figure 1, will make greater use of Ethernet that replaces former bus systems [31], [32], leading to convergence with traditional hierarchical IT networks. To cope with the growing number of functions, service-oriented communication is increasingly used, which is mainly implemented via Ethernet using the Scalable Service-oriented Middleware over IP (SOME/IP) [33] and Data Distribution Service (DDS) [34] protocols. In contrast to the previous so-called signal-oriented communication, which is mainly used via CAN, LIN, and FlexRay, the communication patterns are no longer statically defined at development time. Instead, a network is built up dynamically via service discovery, in which new communication paths can emerge at any time while others are discarded. Concerning the ECUs, the trend towards autonomous driving and connectivity leads to centralization and more performant computing platforms [4], [35]. For example, a central computing cluster and an I/O cluster are introduced in Figure 1. Whereas many ECUs for individual tasks were connected via gateways in the past, the focus here is on general-purpose computing platforms that can provide several functions.

In our example, the ECUs in the infotainment and body domain are connected via CAN or LIN to the I/O cluster because infotainment and body functions are mostly user I/O driven with low data rates. Hence, the cheaper CAN / LIN is preferred to Ethernet. Traditionally, functional domains (chassis, driver assistance (ADAS), infotainment, body) were physically separated via different buses for the domains. While in our example this separation still exists, in Ethernet networks alternatively a virtual separation via a different Virtual Local Area Network (VLAN) tag for each domain is possible. Operating systems such as the POSIX-based (i.e. Linux) AUTOSAR Adaptive, AUTOSAR Classic, Android Automotive[3] or GENIVI[4] running on powerful ECUs support the general-purpose platform approach. As with bus systems, operating systems have different application areas that vary in required real-time capability, safety, the scale of software, and user interaction.

2) EXTERNAL NETWORK

For the outside world, the external interfaces connect the vehicle to other cars, the user's smartphone, or home. Different protocols are used for wireless communication within the external ecosystem. In addition to the cellular 3/4/5G network, Wi-Fi is a vital interface to connect the car to backend servers on the Internet ultimately. The market for backend platform providers is still very new and not yet consolidated. Large IT enterprises, for example, offer their cloud services and join forces with corresponding technology partners, such as Microsoft and Volkswagen, for integration into the vehicle ecosystem [36]. For communication with other cars and infrastructure Dedicated Short Range Communication (DSRC) or Cellular V2X (C-V2X) are used. The Onboard Diagnostic (OBD) interface is used to connect mechanics in the workshop to the car to conduct fault diagnosis of the vehicle or perform software updates.

B. THREATS AND ATTACK IMPACTS

The risk of cyberattacks is increasing with the evolution toward autonomous, connected, and shared vehicles controlled by software. A variety of attacks and potential threats to the entire ecosystem have been carried out by both white-hat and black-hat hackers (i.e., criminals with malicious intent). In general, the number of attacks has increased in recent years [37]. The main risk for cyber incidents originates from the external interfaces through which attackers can gain access to the vehicle. From there, further compromises can be made via the internal network. The CAN bus as an in-vehicle network vulnerable to e.g. spoofing, replay, fuzzing, and denial-of-service attacks [38] as CAN does not bring inherent security features. Newer technologies such as Ethernet and service-oriented software architecture pose new risks to vehicle security [20], [39], [40]. Furthermore, with standardized hardware, software, or operating systems (e.g., AUTOSAR, Linux-based systems), the potential for attacks to affect diverse vehicles is exceptionally high. Sensors can be fooled by manipulated data or jamming as another example, e.g. the global navigation satellite system (GNSS) or blinding a camera [41]. With the advent of Machine Learning for automated driving, adversarial attacks against the automated driving systems are on the rise. Using some physical stickers, some systems can be fooled to misclassify street signs [42]. Other researchers demonstrated an attack against the object tracking system [43]. In addition, manufacturers' servers are hackable from the vehicle side if the attacker gains access to a car's telematics control unit [44]. Finally, every single line of code carries the risk of creating new vulnerabilities [6].

The disclosed vulnerabilities and attacks on BMW (14 remote and local vulnerabilities) [45], Tesla (remote control of e.g. brakes) [46], or Honda [47] show that almost no vehicle manufacturer has managed to leave unscathed. Highly valuable are theft of consumer data, theft from multiple

[3] https://source.android.com/devices/automotive/

[4] https://www.genivi.org/

vehicles at once, and attacks that threaten physical security and thus could be exploited for terrorism or extortion [1]. By attacking to manipulate automated driving functions, for example, hackers can gain control over a vehicle's physical behavior [3]. Ultimately, this can have safety-critical consequences for occupants and the surrounding environment.

C. DEFENSE MECHANISMS

In the automotive ecosystem, different assets need to be secured, e.g. the applications running in the vehicle and on backend servers, the communication on the network, and the hardware (i.e. the ECUs) themselves. Depending on the asset type, typically different security measures are used. For more detail on the specific solutions please refer to the literature [15], [20], [40], [48]–[50].

1) FLEET/BACKEND

The broadest countermeasures are actions that are executed in a backend server of the manufacturer or a third-party provider. The data from the entire fleet can be collected and analyzed here. Typical actions include correlations of events by the SIEM, the detection of outliers on the level of fleet data (misbehaving and anomalous vehicles), and the overarching V-SOC where all activities are coordinated. After identifying an incident, Software updates over the air (SOTA) are deployed fleet-wide to fix vulnerabilities or to improve the vehicle's configuration. Alongside monitoring and secure updates, vulnerability and risk management processes support fleet security throughout the entire lifecycle.

2) EXTERNAL INTERFACES

The most important aspect is to reduce the number of interfaces as much as possible without neglecting customer needs. In addition, applications that are allowed to communicate vehicle-externally must be restricted in their access to onboard networks and applications. Diagnostic access should be secured to a particular degree, as diagnostics bring permission to perform hazardous operations (e.g. flashing new firmware).

3) APPLICATIONS

The secured operation of applications is ensured with measures such as access control mechanisms to grant or deny actions on resources (e.g. execute other applications, read or write files), and isolate applications in different execution environments with separated OS. Applications can also be secured with Host-based Intrusion Detection Systems that observe and monitor applications and the OS of an ECU. Detected deviations from normal behavior could indicate the execution of uncommon system functions.

4) NETWORK

The vehicle's network can be secured with different measures. First of all, communication with different safety criticality and disparate customer functions (e.g. body and comfort vs. powertrain) should be separated into different physical or virtual networks. In addition, firewalls enable blocking any communication on a network based on a set of predefined rules, both vehicle-internally as well as via external interfaces. Lastly, Network-based Intrusion Detection Systems identify deviations of the communication from the expected behavior of a network and are thus crucial for detecting CAN or Ethernet security incidents.

5) DATA

Encryption and authentication mechanisms are the major elements for securing the data. Here, common mechanisms to authenticate CAN communication (i.e. Secure Onboard Communication (SecOC) [51]) and to encrypt and authenticate Ethernet communication (e.g. IPSec, TLS, DTLS) are relevant. CAN communication is typically not encrypted because the computational overhead is too significant for low-cost microcontrollers.

6) ECU & HARDWARE

The innermost defenses are hardware security measures and procedures to manage the keys that are required for authentication or encryption. The roots of trust are tamper-proof modules of microcontrollers that provide the hardware for e.g. encryption and storing of keys on that ECU [19]–[21]. Based on this hardware anchor firmware code such as the bootloader and the update mechanism are authenticated and maintain a chain of trust with secure boot and secure update [52]. Another hardware measure is sensor redundancy. Although this is more considered as safety measure, redundant sources of information are also helpful when one sensor is attacked (e.g., jammed or blinded).

7) EXAMPLE SETUP

In our example (see Figure 1) we highlight some positions in the E/E architecture and the backend systems where security mechanisms could be deployed. For example, it makes sense to control network traffic to the outside of the car with a Firewall. Intrusion Detection Systems are integrated where important applications are running (e.g. safety-critical functions such as the powertrain) or where network data can be easily captured (e.g. a switch). Access control is essential in the connectivity controller, where e.g. diagnostic access to the car is handled. Hardware roots of trust, authentication, and encryption mechanisms are not shown in the example for simplicity.

In the backend, a SIEM and V-SOC system could be run on a server of the manufacturer or external servers. Regardless of where the systems run, the data of the fleet is collected and analyzed in the SIEM / V-SOC. A schematic overview of the analysis process in the SOC using the SIEM system is shown in Figure 3. While the assets, concrete rules, and actions are different between IT security and automotive (fleet) security, the workflow and structure of SIEM and SOC are likely to be reused, see e.g. [53]. First, the events that are sent or pulled from the fleet are transformed to a common data format, if necessary. The collection

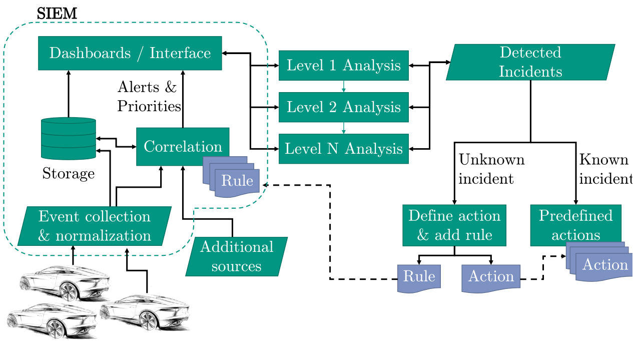

FIGURE 3. Schematic overview of a SIEM system as part of a SOC.

process also includes persistent storage of the events to make them searchable for the analysis processes, for example in a database. Correlation in a SIEM ranges from aggregation of frequently occurring events, counting common attributes, e.g. IP addresses, to hand-written if-then rules, to machine learning-based alerting. The correlation process can involve additional information, such as known vulnerabilities or common attack patterns (see also following Section IV). Correlation is used to prioritize events and alert analysts in the SOC, who access the data through a user-friendly interface, such as graphical dashboards or tables. An alert triggered by the SIEM can be analyzed by multiple layers of (human) analysts in the SOC [54], [55]. If the analysis requires more time or knowledge, the alert is pushed to a higher level. Hence, at the lowest level, all alerts are continuously monitored, while the higher-level analysis should be done by automotive experts who investigate the critical cases. Then, for known incidents, a series of predefined actions are taken to address the incident. In the coming years, the more common case will be that the incident is unknown and an individual response to the incident needs to be developed. The actions that are taken in such cases are not clear today but could span from informing the vehicle in minor cases to e.g. blocking the vehicle's internet connection [55], [56]. Often, an over-the-air update of the vehicle's software to eradicate the problem or vulnerability will be required. However, it must be ensured that an update does not jeopardize the safety of the vehicle [57]. Finally, new SIEM rules that capture the detected incident can be added to the correlation process to facilitate analysis of similar incidents in the future.

IV. SECURITY CONTEXT OF VEHICLES AND FLEETS

Context and context-awareness are initially abstract concepts that have to be tailored to our automotive applications. Since context-awareness has not yet well established for security in the automobile, we initially formulate a terminology. The term *context* has been defined various times, and the most popular definition of Dey and Abowd [58]

> "Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves."

is still valuable. To elaborate on this, the entities of automotive security context span - by means exhaustive - from the ECUs, data, and network of the car to its functions, the environment it is situated in up to the vehicles themselves, threats, and vulnerabilities. Likewise, we will follow a definition given by Jovanovikj *et al.*, who refined Dey and Abowd's definition to introduce the term *security context*: "Security context is a set of contextual information considered relevant for the process of security, regarding a particular task or activity." ([59]) Further on, Gartner defined *context-aware security*: "Context-aware security is the use of supplemental information to improve security decisions at the time they are made, resulting in more accurate security decisions capable of supporting dynamic business and IT environments." ([60]) Closely related to context-aware security is the term *cyber threat intelligence* (CTI), defined e.g. by Rob McMillan [61]. Context information about the (enterprise or vehicle) network assets forms a part of threat intelligence. However, the scope of CTI is more focused on knowledge that is not originating from enterprise-internal sources, but external sources like social media or publications. The goal of CTI is to link indicators of compromise with specifically known threat actors (i.e. "hackers") and their typically used methods. In addition, it is not exclusively technical information, but also gives e.g the SOC manager or security architect a bird's eye view of the security situation. First CTI providers and information sources for the automotive industry are arising, e.g. the Automotive Information Sharing and Analysis Center (AUTO-ISAC) [62] or Upstream Security's AutoThreat feed [63]. These sources provide valuable context on ongoing or successful attacks for the entire fleet. The necessity of context-awareness concerning this *external* context will increase in the future with the rise of new threat actors, but we also elaborate here on *internal* context-awareness, aiming at a situational understanding of the vehicle and fleet security. Based on the aforementioned definitions, we define *context-aware automotive security* as

> *the use of contextual information to improve the security decisions at the time they are made, resulting in more accurate, efficient or timely security decisions capable of supporting dynamic automotive vehicle and fleet environments.*

A. CHALLENGES FOR THE DEVELOPMENT OF SECURITY MEASURES OF VEHICLES AND FLEETS

As already included in our definition, context-awareness is required due to the dynamic environment and because accurate, efficient, and timely decisions are crucial for automotive security. Some of the inherent challenges of vehicle and fleet defense were stated by Mercedes-Benz [64]. In his talk, Guy Harpak outlines the fleet's scale and dynamics, the strict requirements on remediation time, and the diversity of car types as the key challenges. To give more emphasis, we outline some of the major challenges the industry faces in developing security measures for vehicles and fleets.

A) First of all, a vehicle can be exposed to an unlimited variety of situations and environments during its life cycle. Though the manufacturer does not know what users will do with their vehicles in advance, security must be guaranteed in all these situations.
B) Secondly, new vulnerabilities and incidents can manifest themselves throughout the entire life cycle. Given the potentially devastating consequences of a successful attack, the time required to resolve known issues must be as short as possible, up to a real-time response in the vehicle if a successful attack is highly probable. The sale and exchange of successful attack methods and tools amongst criminals increase the risk posed by vulnerabilities to fleets.
C) Thirdly, the volume of functions and software applications executed in a vehicle causes a massive expansion of in-vehicle networking. Powerful computers with standardized operating systems and a wide variety of bus systems and protocols are required for such a development. Security measures must cope with this tangle of different data and interpretations.
D) At the same time, users expect security to be provided ‘‘in-kind’’, making it difficult to justify additional customer costs. Accordingly, development efforts for in-vehicle security measures are aimed at tying up as few hardware and software resources as possible, since this does not bring in a significant profit for the manufacturer.
E) About SIEM and V-SOC, the amount of potential data to be analyzed poses a particular challenge. With millions of vehicles on the road, it is expected that there will be many alarms and incidents. Satisfactory prioritization of the analysis of security incidents is already difficult in business IT. This is all the more true for the automotive industry. It is impossible to collect and store all available raw data from the vehicle's sensors, network, and computing units in a central instance. Moreover, data protection and privacy reasons make it difficult to analyze or store raw data from vehicle fleets.
F) Finally, a fleet consists of different vehicle types and variants. No vehicle from one manufacturer is the same as another. In addition, software updates and on-demand functions increase the number of variants and configurations over time. However, security solutions must be able to cope with all these factors and consider the risks posed by third parties' external services (maps, car sharing).

The effectiveness of the security measures applied must be guaranteed for all situations and be prepared to deal with the challenges mentioned above. A central aspect for this purpose is the knowledge about the security situation of the individual vehicle and the fleet. Security measures both within the vehicle and in the backend require an understanding of the vehicle situation, the vehicle network, the functions, weak points, environment, vehicle configuration, etc. to be able to react adequately according to the situation, since what is normal or permitted behavior depends on the context.

B. BENEFITS AND USE CASES FOR CONTEXT

Summing up the aforementioned challenges, it depends on the (security) context, what our security measures should do. To secure vehicles and fleets effectively and efficiently, the respective security measures, tools, and processes must be provided with relevant contextual information. Thus, providing context information means making the required knowledge about the environment, situation, status, and landscape generally accessible to all security components. By turning context into a tool of automotive cybersecurity, significant benefits can be achieved. We enable security measures to:

1) Understand the situations and react accordingly
2) Bring semantic meaning to the in-vehicle network and fleet data
3) Uniformly access contextual information across the entire brigade of security measures
4) Match the relationships between alerts and events, vulnerabilities and attacks
5) Reduce the scale of fleet data for analysis and the data transferred for security purposes from vehicles to the backend

As a result, context-aware and intelligent security measures can be developed. In the vehicle, adaptive security is possible, which flexibly adjust its decisions to their environment based on the contextual knowledge provided. In the backend, the context-awareness manifests as a more intelligent analysis of security events and incidents, correlating security events with fleet behavior, in-vehicle network data, or vulnerabilities.

To elaborate on this, let us have a look at some primary use cases of security context, as outlined by Sinha *et al.* [65]. They mention the *configuration* of security measures, the *placement* of security measures in the network, and the *modification* of the measures themselves as the key concepts for the use of context information. Benefits *1)* and *2)* are linked to the *configuration* aspect. With an understanding of situations, network and vehicle type differences, the thresholds of an IDS system, the rules of a firewall, or the attributes of an access control system can be configured and adapted to fit the context. Benefit *3)* is associated with *modification*, since the security measures do not need to infer context on their own but can rely on provided context. However, benefits *4)* and *5)* go beyond the key use cases mentioned by Sinha *et al.* [65]. Thus, we opt to add *analytics* as another main use case. In *4)* and *5)*, context information serves as relevant threat intelligence for the processes and tools involved in monitoring the entire fleet and reduces the analysis efforts. As security is a relatively new challenge to the automotive domain but will become a new dimension of quality in vehicles [6], also *modification* and *placement* are use cases for context information. The optimization of the monitoring strategies of the vehicle network and closing the loop with the development of mature software and appropriate and efficient security measures are potential applications.

C. CONTEXT MODELS

According to Dey and Abowd, the above definition of context, similar to other definitions, leaves a lot of room to define what is context information and what is not. Usually, for the individual use case, a selection of information must always be made, which is considered context information. The set of selected information is called the context model, according to Henricksen's definition [66]:

> "A context model identifies a concrete subset of the context that is realistically attainable from sensors, applications and users and able to be exploited in the execution of the task. The context model that is employed by a given context-aware application is usually explicitly specified by the application developer, but may evolve over time."

This selection can be specified more or less formally. We distinguish here between implicit and explicit context models (see also [16], [67]). Implicit models are only defined within a specific application and the identification of the current context takes place within the application. An example for vehicle security would be if the firewall subscribes to the required raw data like the vehicle speed and draws its own conclusions about the context like *fast*, *slow* or *standing*. In contrast, an explicit context model is defined outside the application and the derivation of the current context takes place outside. For the firewall in our example, it is not apparent where the context information *fast* originates from, it is only made available to the firewall.

This approach's advantage is that an explicit context model provides a clear separation of the utilization of context information and the definition of the relevant information and its derivation. Furthermore, the context can be made available to several applications simultaneously and uniformly. This is particularly advantageous concerning the configuration of several security measures. If, for example, firewall and IDS both use the speed as raw information to derive the context *fast* and *slow* therefrom, conflicting configurations of the overall vehicle security can occur if each application uses its implicit context model. On the other hand, it can be disadvantageous for the system's performance to use an explicit context model, because the access to the context information is carried out via a communication medium or middleware instead of within the application [16].

D. CONTEXT MODELING METHODS

For the specification of explicit context models, various methodologies exist in the literature and multiple surveys provide an extensive overview of these context modeling methods [16], [17], [68]–[71]. In general, any method for describing available information, data, and its relationships can serve as a context modeling method. The context can be modeled from a general perspective or application-specific. Although generic models help to get an idea of the universal categories that are understood as contextual information, an application-specific model is required to use contextual information. To give an example, Hoda *et al.* present in their paper [72] the ECOPPA ontology for modeling the context of pervasive computing applications to improve the system users' physical activity. Here the authors show how they derive an application-specific model from a generic one (upper layer vs. domain level in Figure 4). While *location* is generally important for different applications, only specific locations are actually of interest for each application.

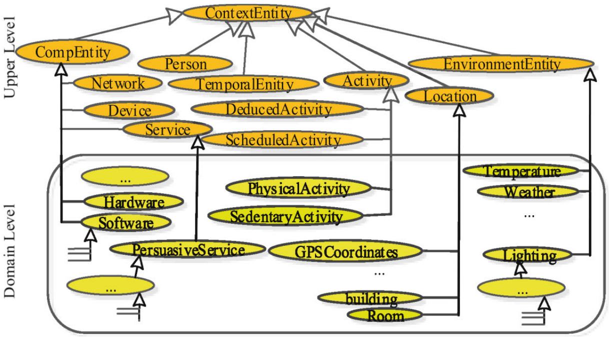

FIGURE 4. ECOPPA ontology of Hoda *et al.* [72].

We summarize the most important modeling methods *key-value based, markup language, relational, object-oriented* and *ontological*. In contrast to other studies, we refrain from including graphical models here, because although they are well suited for developing a context model, they still require some implementation to be used in applications. The implementation is always a specific data structure that can be filled with the current context information at runtime, which can be accessed and queried to retrieve and use the information. Accordingly, a graphical model is a helpful way to achieve one of the realizations described below. For example, the Unified Modeling Language (UML) or specific tools can be used to create graphical models. For ontologies, Protégé[5] is very often used as a modeling tool. However, there are several other modeling tools, e.g. Enterprise Architect, which supports both UML, XML schemas, and ontologies based on the UML extension Ontology Definition Metamodel.[6] Additionally, we briefly describe how the context models can be persisted, depending on the technology.

1) KEY-VALUE MODEL

Key-value models are the simplest form for context description. For this purpose, we define a list of relevant information, which are the keys. During runtime, the values are accessed like a dictionary. The enumerated attributes (keys) are easy to design and retrieve. However, no structure, hierarchy, or relationships can be modeled. This means that no meta-information about any context attribute (e.g.

[5] https://protege.stanford.edu/

[6] https://www.omg.org/spec/ODM/

uncertainty, time) can be captured. In addition to storing key-value data in proprietary file formats, key-value databases could be used to persist context information. In the automotive sector, the persistent storage of both key-value pairs and files similar to a desktop PC file system is covered by the AUTOSAR adaptive standard and can therefore be used by applications without significant development effort [73].

2) MARKUP LANGUAGE

A context model based on the markup language is a machine-readable document that delivers structured content. The content here is the context information, which is hierarchically structured. The markup part of the document makes it machine-readable. The most important form for our use case is the XML (extended markup language) format. Here, a schema defined at development time ensures the validation of the document format and, in part, the value ranges of the data. Compared to key-value storage, schema validation and hierarchization are advantageous, while relationships and reasoning are not supported either. The XML format is primarily useful for persistent storage as files and for exchanging and transferring data between applications (e.g. [74]).

3) RELATIONAL MODEL

Relational models originate from the development of relational databases. Data is thereby described as tables (relations) of several attributes. Each tuple of a relation has a unique identifier. Several relations can be connected by defining an additional relation for their dependence. Hierarchical relations can be represented implicitly as well, so a relational database has a higher expressiveness than an XML notation. For querying the data, standardized languages e.g. SQL exist, which increases the usability. Although relational models were not typically considered a tool for contextual modeling, they are considered here, since several methods treated as graphical modeling in other publications can be mapped to a realization as a relational database. These include, for example, object-role models and entity-relationship models. The implemented databases follow a previously defined schema, which on the one hand allows validation, but on the other hand, makes it difficult to extend. Relational databases are industrially established, the development procedure is well-known and efficient implementations are available for embedded systems.

4) OBJECT-ORIENTED MODEL

Object-oriented modeling techniques are well known in the automotive industry, especially since there are established tools and workflows for graphical development, e.g. the structural class and object diagrams of UML. Class diagrams and object orientation use hierarchies and relationships, which are very important for context modeling. Since object orientation is closely linked to programming languages, the implementation and integration of context objects are straightforward. However, there is no standard high-level query approach, so contextual security mechanisms require more effort to access the context information. Nevertheless, no additional time delays are expected when accessing the information in a context object, as is the case with a database query. For persistence, there are object-oriented databases or methods for the serialization of files, e.g. JSON.

5) ONTOLOGICAL MODEL

Ontologies are synonymously called ’vocabulary’ and are the cornerstone of the so-called Semantic Web.[7] They are a semantic model of the terms of a particular application domain and contain the terminology and their relations among each other. The elementary construct on which ontologies are based are triples of subject, predicate, and object, e.g., *the car* (subject) *is* (predicate) *fast* (object), thus representing an annotated connection between two terms. These triples follow the RDF standard and can be stored in various machine-readable formats. To differentiate the terms and resolve ambiguities or draw automatic conclusions from terms, logical languages exist to extend the RDF standard, such as OWL. All standards are managed by the W3C. For context modeling, ontologies are seen as quasi-standard in different publications. In addition to the automatic inferences already mentioned, the uniform and consistent definition of terms, independent of data sources, and the reuse of already publicly available vocabularies for various areas are of particular interest to us. For the automotive sector, for example, the VSSo ontology exists (see Section V-C). A disadvantage for the automotive industry is that the methodology comes from the academic and scientific world and therefore engineers have little experience in its use. In addition, queries using the available query languages (e.g. SPARQL) can be resource inefficient. An interesting aspect is the interrelation of ontologies with knowledge graphs, which use ontologies as schema (formal semantics) for a graph-based database representation. The most well-known knowledge graphs are the Google Knowledge Graph,[8] DBPedia [75], and Geonames.[9] In addition to graph databases as long-term persistent storage, XML is suitable for storing the RDF triples.

E. PROPERTIES OF CONTEXT

For a technical use of context, it is important to consider that context information has or may have different properties. A particularly important concept is that context is an abstraction [69]. Compared to raw data that represent low-level information, such as sensor readings, context information is generic and on a higher level (see Figure 5). Since context is

[7] Introduction: https://www.w3.org/2009/Talks/0615-SanJose-tutorial-IH/Slides.pdf

[8] https://blog.google/products/search/introducing-knowledge-graph-things-not/

[9] Founded by Marc Wick https://www.geonames.org/

TABLE 2. Properties of context information.

Context property	Description	Scale
	Defined at design time	
Persistent	context element whose changes should be monitored and archived	yes / no
Dynamicity	frequency of change of a context element	static / dynamic
Sensitivity	level of PII	defined levels, e.g. *high*, *medium*, *low*
	Varying at runtime	
Trusted	context element whose sources are well-known and not subject to manipulation	percentage
Certain	context element whose sources are physically reliable and accurate	percentage
Valid	context element that is not outdated	yes / no

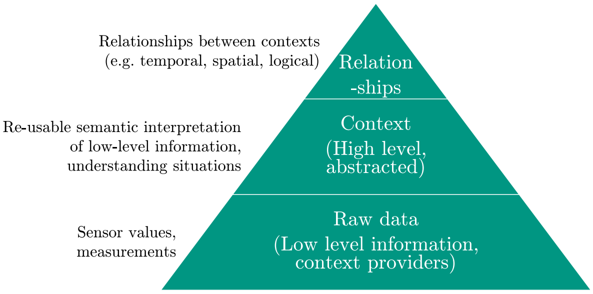
Relationships between contexts
(e.g. temporal, spatial, logical)
Relation -ships
Re-usable semantic interpretation of low-level information, understanding situations
Context (High level, abstracted)
Sensor values, measurements
Raw data (Low level information, context providers)

FIGURE 5. Abstraction levels of context, adapted from [69].

a description of a situation (see definition of Dey and Abowd in section IV), relationships between contexts form another level of abstraction. This includes both temporal and logical relationships, e.g. a sequence of contexts in which our vehicle is situated. Sanchez et. al [76] emphasized the distinction between raw (sensor) data and context, created by processing the raw data. In addition, the authors considered consistency an essential property.

Wan and Alagar [77] outline the properties of persistency, dynamics, and trust that some of the context elements and classes may have. Feld and Müller [78] elaborate on the aspects that should be taken into account to model knowledge in the automotive domain. Their publication particularly addresses the high proportion of dynamic context information in automobiles and recommends a careful consideration of the factor time using a validity period of the context information. The authors also refer to the uncertainty of context information, caused by uncertain raw data on the one hand and heuristics and inaccuracies in the derivation of the actual context on the other hand. Finally, Feld and Müller address privacy issues, since context information in the vehicle can contain a high degree of personally identifiable information (PII), such as location, usage habits, or driving style. Accordingly, the sensitivity of the context information must be considered as a characteristic. These properties of context information can be divided into two categories (see Table 2). Persistence, dynamics, and sensitivity are determined by the context class and are therefore defined system-specifically once at development time. However, trust, validity, and certainty may differ from context element to element and may also vary at runtime.

V. OVERVIEW OF CONTEXT-AWARE VEHICLE SECURITY AND RELATED AREAS

To consider the state of the art of context and context-awareness in the field of vehicle and fleet security, it is necessary to clarify the scope of the consideration. This work aims to shed light on the following questions:

- What is the security context of vehicles and fleets?
- What methodology can we use to model and store context information?
- How can we extract context from data and what are relevant data sources?

There are still some questions left unanswered beyond our work scope, primarily how the context is determined in detail and how security measures can make use of this knowledge. Nevertheless, we see the biggest challenge in the definition of the relevant information, which is why our work focuses on this issue.

The choice of publications considered here provides essential information on one or more of the aforementioned questions. However, none of the works found addresses all questions to our satisfaction. Since context-awareness for vehicle and fleet security is a very new topic, only a few papers directly refer to this subject. For this reason, we have decided to include other areas in our overview in addition to the publications that directly relate to the automotive application of context information for security (see Figure 6). This includes publications that deal with the modeling methodology of context and the corresponding tools already outlined in Section IV-D. Moreover, papers are relevant that deal with the application of context in the automotive sector, though not for security. Additionally, we are interested in methods that deal with the aggregation, abstraction, and discovery of relations in vehicle and fleet data, since abstraction is one of the elementary concepts of context (see Figure 5). Outside the automotive domain, application areas of interest impose similar constraints on their security measures, such as real-time capability and limited resources. Accordingly, for example, works from the domains Internet of Things (IoT) and Supervisory Control and Data Acquisition (SCADA) are considered, which use context information. As mentioned

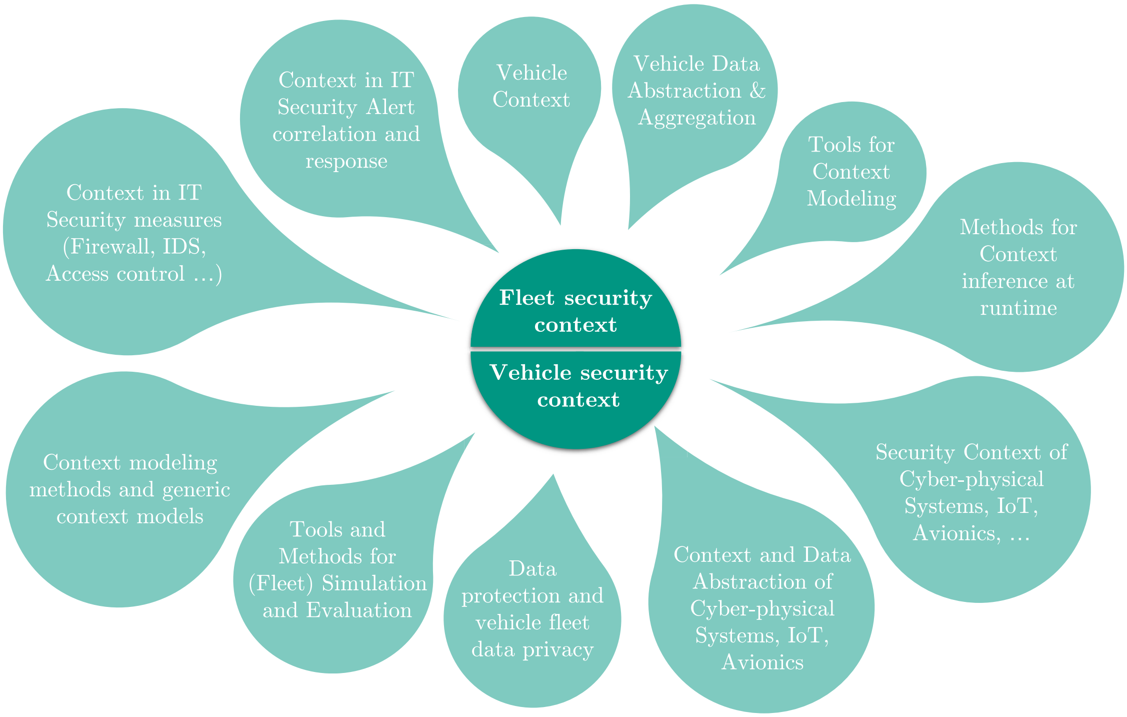
Fleet security context
Vehicle security context
Context in IT Security Alert correlation and response
Vehicle Context
Vehicle Data Abstraction & Aggregation
Tools for Context Modeling
Methods for Context inference at runtime
Security Context of Cyber-physical Systems, IoT, Avionics, …
Context and Data Abstraction of Cyber-physical Systems, IoT, Avionics
Data protection and vehicle fleet data privacy
Tools and Methods for (Fleet) Simulation and Evaluation
Context modeling methods and generic context models
Context in IT Security measures (Firewall, IDS, Access control …)

FIGURE 6. Overview of relevant areas for this survey.

above, the architecture of the vehicle is converging more and more with conventional networks. POSIX-based operating systems, high bandwidths over Ethernet networks, and higher performance hardware lead to the fact that context-aware security measures from traditional IT needs to be considered. In addition to the consideration of context for individual measures, we place particular emphasis on the area of SIEMs, alert correlation, and threat intelligence.

Several issues are also relevant to the development of context-aware security measures for vehicles and fleets. One major problem in developing security measures for the automotive domain is the absence of suitable datasets with both benign and malicious behavior [15]. Especially for vehicle fleets, no publicly accessible data set contains the required information such as real-world attacks. Besides, the available data sets from the in-vehicle network usually lack the coverage of all network types and the decoding of the user data (i.e. semantics of the data). For example [79] and [80] contain only CAN data, but no Ethernet or service-oriented communication. Thus, simulation of the vehicle, its environment, its internal and external communication as well as attacks is an important category of publications to consider here. Finally, as outlined in Section IV-E, privacy is a crucial aspect, especially for fleet security. With the recent regulations, e.g. in the EU [81], unless a person has agreed to data processing for a specific purpose, personal data may not be processed. Excepted from this is the fulfillment of other legal obligations, for example, the protection of vital interests of a person or the fulfillment of a task in the public interest. Therefore, some PII may be accessible in backend systems, if such legal basis exists. However, this will not be the case for most in-vehicle data. Appropriate mechanisms, such as anonymization, are therefore required to make the tremendous amount of in-vehicle data accessible in the backend for fleet security purposes.

A. FLEET SECURITY

Context-aware security solutions for monitoring and analyzing an entire fleet are not frequent in scientific literature. Regarding the analysis of multiple vehicles, the focus of scientific research lies in the field of misbehavior detection in VANETs based on V2V communication and less on centralized fleet monitoring. As indicated by recent surveys [82], [83], the VANET solutions, in general, are mostly designed for use in Road-Side Units or a Central Authority and not in a backend system. To give an example, Ghaleb *et al.* [84] introduce context-awareness in their solution. The authors consider the mobility information of surrounding vehicles in the VANET given by the internationally standardized *basic safety messages* (BSM) and *cooperative awareness messages* (CAM). They locally identify misbehaving vehicles in the VANET via multiple reference models of the surrounding vehicles' behavior. However, local detection of misbehaving vehicles in VANETs is not our focus, although using V2V messages such as BSM and CAM as context information for security decisions in a backend may be an exciting idea.

In contrast, industrial approaches are already being heavily promoted. There are various providers of SIEM, SOC, or monitoring solutions on the market, e.g. Argus Fleet Protection [54], [85], Arilou SIEM/SOC Backend Solution [86], HARMAN SHIELD [87], and Upstream Security's C4 Platform [88], [89]. Especially Upstream and Argus advertise to foster context-awareness. In publicly available material, Upstream gives as an example of context whether the vehicle is driving or parking and which actions are permitted in this state. Context can be used in their industrial product to define specific rules for identifying incidents. Kaneti [54] emphasizes that context aggregation within the vehicle instead of deriving context from raw data in the cloud results in data transfers of only 1-15 Mbyte per year instead of 0.1-5 GByte per hour. Thus, a massive reduction of the required bandwidth and consequently costs for data transfer to the SIEM / SOC can be achieved. As examples of context, the geolocation of the vehicle, build details, component versions, service history, warranty data, and the IDS configuration are mentioned. Additionally, some scientific publications emphasizing SIEMs and fleet-wide analysis are mentioned in the following.

1) "AUTOMOTIVE SIEM AND ANOMALY DETECTION USING SAND-SPRINKLED ISOLATION FOREST" [90]

Haga *et al.* propose to use a SIEM system in the backend, that analyzes in-vehicle communication and logged data. It is based on a specific adaption of the machine-learning technique Isolation Forest to reduce false alarms, which are familiar to anomaly-based IDS [90]. Additionally, they suggest integrating vulnerability information as context / CTI from the Auto-ISAC [62].

2) "POSTER: ANOMALY-BASED MISBEHAVIOUR DETECTION IN CONNECTED CAR BACKENDS" [91]

Berlin *et al.* introduce SeMaCoCa (Security Management of Services in Connected Cars), a SIEM system that aims to detect misbehavior based on anomaly detection using machine learning [91]. However, in addition to vehicle data, the authors mention various external context data sources that could improve the detection process, e.g. from external service providers, data from service workshops, weather, or road conditions. Although it is not explicitly referred to as context-awareness in the respective publication, it can be understood as a step in this direction.

3) "ADVANCED ANALYTICS FOR CONNECTED CARS CYBER SECURITY" [92]

Levi *et al.* [92] propose a security system that is located in a backend and uses a Hidden Markov Model (HMM), which is a machine learning approach to predict event sequences (see Figure 7). Instead of transmitting raw vehicle data (e.g. CAN data, operating system data) to the backend, they use a rule-based approach to convert the raw data into *events* before transmission (e.g. *Logging*, *Engine Stop*, *Download App*). The HMM analyzes the events in conjunction with attributes defined per event (e.g. location and speed). Although the authors aim to develop an anomaly-based IDS that monitors each vehicle of a fleet individually, the concept of *events* is interesting from the perspective of context-awareness. Events represent a form of abstraction and can thus describe the context. On the one hand, this enables a universal analysis independent of the vehicle type in SIEM/SOC, while at the same time reducing the amount of data to be transmitted and simplifying the understanding of the underlying raw data.

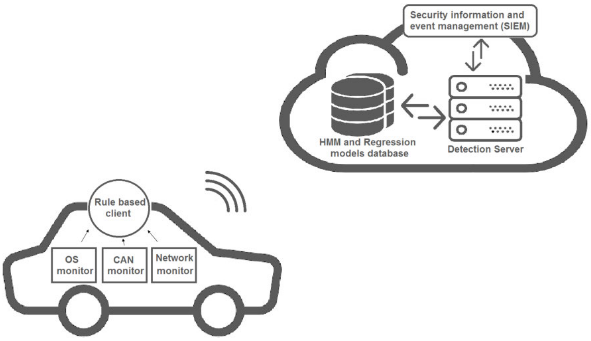
Security information and event management (SIEM)
HMM and Regression models database
Detection Server
Rule based client
OS monitor
CAN monitor
Network monitor

FIGURE 7. System overview of Levi *et al.* [92].

B. VEHICLE SECURITY

Concerning security measures within a vehicle, context has already been addressed, used, or at least mentioned in other words by several authors. For example, Al-Jarrah *et al.* [15] have conducted an extensive survey that shows the current state of the art in the field of IDS systems for in-vehicle networks. Context-aware IDS systems are identified as a research gap, as only a few publications in this field could be found. According to the authors, the challenge is to develop a vehicle-global profile of normal behavior for which different data sources have to be combined. The publications listed here include primarily intrusion detection systems and access control measures and some more generic or overarching approaches.

1) "TOWARDS A SECURITY ARCHITECTURE FOR PROTECTING CONNECTED VEHICLES FROM MALWARE" [93]

In their recent publication, Iqbal et. al [93] argue for a complete security architecture for connected and autonomous vehicles. In addition to well-known security mechanisms for protection and detection such as firewalls and host-based IDS, they emphasize the need for specific management components for vehicle security. All components are seen as parts of an advanced operating system. The task of one of these components is to administer the vehicle context. According to Iqbal *et al.*, the vehicle context should be used to switch between different security modes. The context-dependent security modes are coupled with other preventive measures, such as a dynamic access control mechanism that restricts applications and connectivity based on the vehicle context. However, the authors do not make a clear statement about the relevant context information or how this information can be derived from vehicle data and sensor information. Nevertheless, they give as examples of context information the type

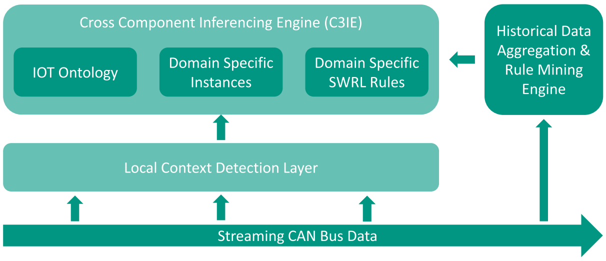
Cross Component Inferencing Engine (C3IE)
IOT Ontology
Domain Specific Instances
Domain Specific SWRL Rules
Historical Data Aggregation & Rule Mining Engine
Local Context Detection Layer
Streaming CAN Bus Data

FIGURE 8. **System architecture of Narayanan [95].**

of road, if someone other than the owner is driving, and the number of connected untrusted peripherals.

2) "DATA-DRIVEN EXTRACTION OF VEHICLE STATES FROM CAN BUS TRAFFIC FOR CYBERPROTECTION AND SAFETY" [94]

Although Moore *et al.* [94] do not explicitly mention the concept of context, their publication is of interest here. The authors present a method to derive physical vehicle states from CAN messages, which can be considered an aspect of context. They use the raw payload of CAN frames without decoding it into specific vehicle signals. However, they only infer the states *idle, acceleration, maintain speed, deceleration*.

3) "USING SEMANTIC TECHNOLOGIES TO MINE VEHICULAR CONTEXT FOR SECURITY" [95]

Narayanan suggests a two-step context abstraction to detect abnormal data flows and scenarios (see Figure 8). Data of the CAN bus is decoded by the "Local Context Detection" layer to the sensor values. In addition, the first layer transforms the raw sensor values to an abstracted measurement (e.g. "high speed" instead of 137.34 km/h), that is associated with an ontology entry. The ontology is located at the level of the "Cross Component Context Inference Engine" (C3IE), which represents the global system state. The knowledge base in C3IE is divided into four components: an extended version of the IoT-Lite Ontology [96] (which describes IoT networks in terms of sensors and actuators), the Domain Instances, some Domain Specific SWRL rules, and a Reasoner. The domain instances are the vehicle-specific sensors, such as a speed sensor, that are mapped to the generic classes in the ontology. The SWRL rules are used to describe high-level context about the system state, which can be "normal" or "anomalous" here. The Reasoner is dynamically checking the SWRL rules to infer the security state of the system. Additionally, they point out that the rules should be based on automatic rule-mining processes instead of specifying them manually.

4) "A FRAMEWORK FOR DETECTING ANOMALOUS BEHAVIORS IN SMART CYBER-PHYSICAL SYSTEMS" [97]

In his Ph.D. thesis [97], Narayanan outlines a method to derive an abstracted vehicle behavior, besides the approach already presented in [95]. The so-called *Automatic Behavioral Abstraction Technique* (ABATe) uses data of regular vehicular operation to learn a context model. Initially, the system generates state vectors from raw sensor data by aggregating vectors whose Euclidean distance is below a certain threshold into a single state. This defines a fixed number of possible system states. Subsequently, a neural network is used to generate an embedding of the state vectors into a lower dimension, taking into account the previously seen state vectors. The network learns to map state vectors that generally occur together in a sequence to context vectors that are close to each other. However, an interpretation of the resulting context vector is difficult and may not necessarily correspond to a human recognizable relationship. Used as a security measure, the context model is used to identify unknown system states or system states that do not conform to normal operations.

5) "CONTEXT-AWARE INTRUSION DETECTION IN AUTOMOTIVE CONTROL SYSTEMS" [98]

Wasicek *et al.* propose the Context-aware Intrusion Detection System (CAIDS) [98]. CAIDS designates itself as context-aware because it incorporates the vehicle control systems' physical context into an IDS. Data from sensors and actuators exchanged between ECUs form the basis for a reference model of the control system. The reference model based on an autoencoder (special architecture of a neural network) provides the first context-level by turning raw data into events when the raw data deviates too much from the reference model. In a subsequent step, it becomes context on a second abstraction level by supplementing the information about deviations from the physical reference model with global information about the vehicle operating status. According to the authors, state machines or decision trees can be used for this purpose, which, depending on the state (e.g. idle, on the highway), ensure a reasonable interpretation of the first-level context. Reference model plus information about the state thus serve to improve the understanding of the semantics of the data by IDS. However, an explicit context model is not propagated and the system is focused on the in-vehicle context.

6) "ROAD CONTEXT-AWARE INTRUSION DETECTION SYSTEM FOR AUTONOMOUS CARS" [99]

Jiang *et al.* introduce the Road Context-aware IDS (RAIDS) [99] (see Figure 9). They infer the road context from camera images and sensor data to detect intrusions in CAN frames. The basic idea here is that the road context influences and determines the correct and benign vehicle signals. The authors emphasize that with the introduction of autonomous driving this connection becomes more and more stringent. While today the road context induces different reactions by differing drivers and thus varying data in the vehicle network, autonomous cars are controlled by algorithms and computers. Therefore the same road context should generate the same network traffic. Jiang *et al.* define the road context as "the

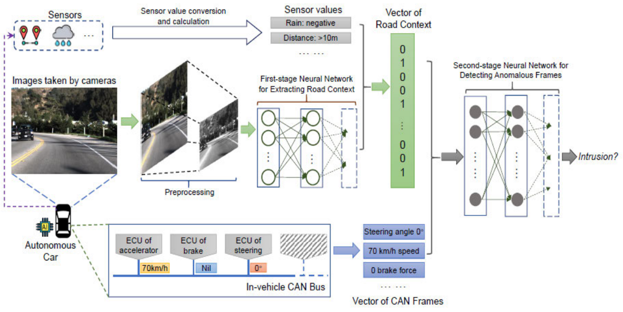

FIGURE 9. **Concept of the road context-aware IDS by [99].**

information an autonomous car is encountering when it is cruising'' [99]. Despite the lack of a formal model, the authors outline some of the most important classes of their context model:

- Road conditions (lane marking, crossroad, junction)
- Traffic lights, pedestrians, vehicles, obstacles, bumps, and pits around the autonomous car
- Weather conditions (rain, fog, cloud, snow)
- Sunrise, sunset, night, and tunnel lights

A neural network, which operates on camera data, in combination with other, not more exactly specified sensor information shall provide this context information. A second neural network is responsible for the detection of abnormal CAN frames. RAIDS is trained in an end-to-end approach, which implies that no explicit context model is given. The inference of the context information is not validated in detail. Accordingly, although the authors name some contextual information relevant from their point of view, it is not evident from the concept presented whether this information can actually be obtained by the first processing step.

7) ''CONTEXT-AWARE ANOMALY DETECTOR FOR MONITORING CYBER ATTACKS ON AUTOMOTIVE CAN BUS'' [100]

Another context-aware IDS approach was presented by Kalutarage *et al.* [100]. They present an anomaly-based IDS, aiming at contextualizing CAN messages. Yet, for the authors, context is nothing more than a frequency distribution of the sequences of CAN messages. The authors use a statistical distribution estimated from normal behavior as context to identify abnormal CAN messages. The main argument for this is that the vehicle manufacturers do not disclose the exact functions for the specific CAN IDs. From the development perspective of the manufacturer, this argument is irrelevant.

8) ''INTEGRATION OF ATTRIBUTE-BASED ACCESS CONTROL INTO AUTOMOTIVE ARCHITECTURES'' [101]

Rumez *et al.* [101] outline a distributed attribute-based access control (ABAC) mechanism for automotive architectures. The presented approach allows access control of ECU requests both within a domain and between different domains. The ABAC policies are based on dynamic attribute values that are inferred by a Policy Information Point (PIP) typically located in a gateway. In a gateway, the PIP has access to several buses inside the vehicle and thus the necessary information to determine the current attribute values. From the context-awareness perspective, the attributes represent the context. With the proposed grouping of attributes into several categories (subject, action, environment, resource) in the standard XML-based access control policy data format XACML, the methodology already goes beyond a simple key-value representation. However, the authors do not go into further detail regarding the definition of relevant attributes. Examples mentioned include vehicle status, seat occupancy, and location.

9) ''oMAC: OPEN MODEL FOR AUTOMOTIVE CYBERSECURITY'' [102]

Another approach for an access control mechanism was outlined by Hugot et. al [102]. Their main goal is the integration of an access control mechanism in the SOME/IP protocol. The authors emphasize that future vehicles need dynamic cybersecurity mechanisms, where access permissions to vehicle resources such as actuators, sensors, and applications are adjusted depending on the state of the vehicle. The access control is based on functional flows, which are defined as automatons. According to the state of the automaton, different policies are active, i.e. allow for specific access. The states of the functions are determined via so-called reference monitors, that are distributed across the in-vehicle network. In this publication, the vehicle status is mainly defined by the status of the functions. The status can be seen as part of the vehicle security context.

10) ''CONTEXT-AWARE ACCESS CONTROL IN NOVEL AUTOMOTIVE HMI SYSTEMS'' [103]

Finally, Gansel *et al.* [103] present a context-aware access control mechanism for automotive purposes. The system presented serves to share access to a specific resource, the screen in the cockpit, among different applications. The applications have both safety-critical and non-critical tasks, which is why the authors focus primarily on real-time capable implementation. Context is, on the one hand, derived from vehicle sensors and communication events such as speed, time, location, or phone calls. On the other hand, it is the specific state of the managed applications. However, the system context based on vehicle data is always directly linked to the application context and is not considered generically or provided to other security measures.

C. AUTOMOTIVE CONTEXT-AWARENESS AND DATA SEMANTICS

The application areas of context-awareness in the automotive domain are very broad, thus we select some exemplary contributions in the field of intelligent, personalized, and adaptive systems. For more comprehensive information, please refer to further publications, e.g. [23], [24]. Aside from that, different ontologies have been published as domain knowledge models

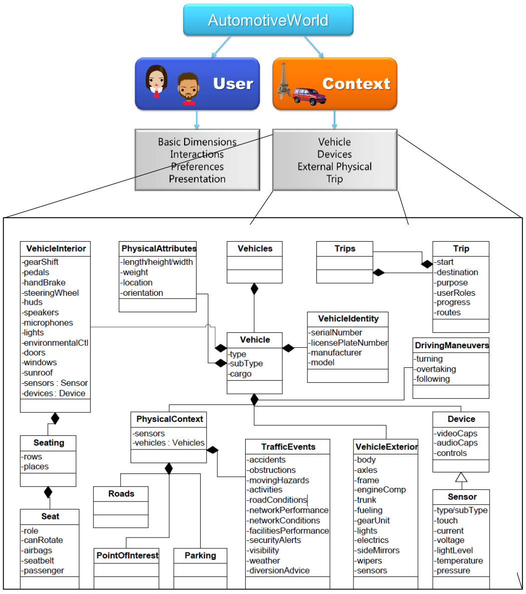

FIGURE 10. Automotive ontology and its context model [78].

for specific automotive applications, e.g. semantic middleware, autonomous driving, and testing (see [104]–[107] for example).

1) "OSGi BASED SERVICE INFRASTRUCTURE FOR CONTEXT AWARE AUTOMOTIVE TELEMATICS" [108]

An early design of context-aware architecture can be found in the work of Zhang *et al.* [108]. Intending to enable modular context processing, they build on the Java-based service architecture OSGi[10] and develop an ontology model for information representation. The ontology is divided into the four core concepts location, user, vehicle, and computing unit, building on the work of Wang *et al.* [109]. The processing uses hardware or software signals as input. Context providers perform an initial abstraction of the raw data, whereupon context interpreters convert low-level context to high-level context allowing context managers to distribute it to services. An example of a use case is the detection of a stranger in a car, at night and the inference of a burglary. While the Java-based implementation may not be practical for realistic embedded real-time automotive environments, the ontology model is attractive for our purposes.

2) "THE AUTOMOTIVE ONTOLOGY: MANAGING KNOWLEDGE INSIDE THE VEHICLE AND SHARING IT BETWEEN CARS" [78]

A description of a possible automotive ontology is provided by Feld and Müller, with a focus on personalized intelligent systems. It is intended to provide semantic knowledge about the user, the car, and the driving situation. For modeling this knowledge, an ontology is used that is divided into the two sub-domains *user* (e.g. preferences) and *context* (see Figure 10). The latter consists of an abstracted view of the interior's sensor data and a description of other data such as POIs, trips, and an external physical context. In addition, suggestions for capturing time series, dealing with uncertain values, and privacy are given. For our purposes, while the extensive modeling of physical attributes of the car is too fine-granular (e.g. weight of the car) and therefore not relevant, other aspects such as traffic events and trips may be of interest for security purposes.

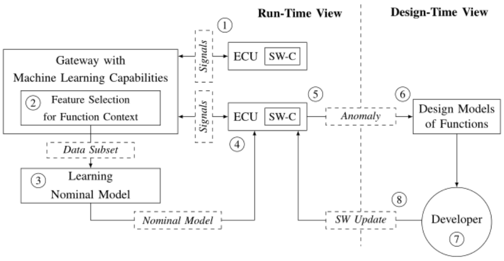

FIGURE 11. System overview of [110] for extracting the functional context at runtime.

3) "SYNCHRONIZATION BETWEEN RUN-TIME AND DESIGN-TIME VIEW OF CONTEXT-AWARE AUTOMOTIVE SYSTEM ARCHITECTURES" [110]

In [110], a framework is designed to detect abnormal use of a particular vehicle function. The framework is divided into eight steps (see Figure 11). The data (''signals'') of all vehicle functions (software components, ''SW-C'') is collected in the vehicle gateway (1). The gateway is a central ECU in the vehicle that is connected to several buses and can access all vehicle data. The amount of data is reduced there by a feature selection algorithm (2) that selects the signals that are correlated with the given vehicle function. The selected vehicle signals define the context of the function, which is used to train a nominal model (3). It is learned by acquiring statistical parameters of the function context. An anomaly is defined by a deviation from the learned statistical parameters (4). When a deviation is detected, it is reported to the backend (5). A state machine with two states (''normal'', ''anomalous'') is generated to represent the situation of the function (6). If anomalies are detected, the responsible engineers are notified (7) and countermeasures (e.g. an update) can be initiated (8).

Beyond, the authors extend their work in [111]–[113]. Identifying the context of a function (i.e. correlated signals of the in-vehicle network) is used to derive a personalization model that reflects when a driver deactivates certain functions. This is the personal nominal model of functional behavior. Deviations from this behavior (i.e. anomalies) can result either from software or hardware faults as well as cyberattacks. Compared to the previous work [110] the identified subset of vehicle signals was subdivided into six groups in [113]. The group *driving context* contained signals such as

[10] https://www.osgi.org/

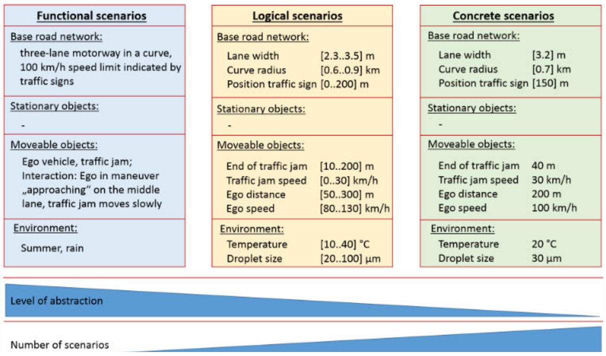

FIGURE 12. Scenario abstraction levels in PEGASUS [114].

road type, speed limit, and temperatures. To identify the vehicle's security context, the feature selection approach seems promising to derive the relevant context information of any user-level function (e.g. cruise control). However, the authors note that the relevant subset of signals varies largely from driver to driver.

4) "ONTOLOGY BASED SCENE CREATION FOR THE DEVELOPMENT OF AUTOMATED VEHICLES" [115]

In the PEGASUS project,[11] an ontology-based method for creating virtual scenes for testing autonomous vehicles was developed [115]. Driving scenes are modeled with different aspects such as road layout and topology, the traffic infrastructure and temporal manipulations of the aforementioned (e.g. construction sites), static and dynamic objects on the road (e.g. cars, maneuvers), and finally environmental conditions (e.g. weather). Of course, the maneuvers and precise descriptions are relevant for developing vehicles that act safely in every possible situation. However, as security attacks can influence the physical behavior of a vehicle, a very abstracted view of the driving scenario may be of interest for security purposes, e.g. in the form of functional scenarios (see Figure 12).

5) "VEHICLE SIGNAL SPECIFICATION - STANDARDIZED WAY TO DESCRIBE AUTOMOTIVE DATA" [116]

The Vehicle signal Specification (VSS) is a taxonomy for the internal vehicle signals, which is standardized by the World Wide Web Consortium (W3C) and developed by the GENIVI Alliance. The uniform semantic definition of the signals existing in the vehicle enables a common and consistent pool of data for applications in the vehicle. The taxonomy focuses on conventional sensors, actuators, and data of infotainment systems.

6) "GENERATING SEMANTIC TRAJECTORIES USING A CAR SIGNAL ONTOLOGY" [117], [118]

Based on VSS (see Section V-C5) and further ontologies,[12] Klotz *et al.* create an extensive ontology for vehicle signals and sensors [118]. The VSS ontology (VSSo) is used in combination with the STEP ontology [119] to map trajectories using the ontology's vocabulary and enrich it with semantic information in the referenced publication. For this purpose, they develop modules for adding and annotating the sensor observations, as well as for graphical representation. The static part of the knowledge base is generated by parsing vehicle configuration files. Dynamic aspects, i.e. observed sensor measurements are added as new RDF instances. To relate the sensor values to points in time and space, the stored triples are tagged with markers for the coordinates and the time. The project is validated using a data stream from the vehicle simulator OpenXC. While the aforementioned VSS taxonomy (Section V-C5) is seen as the data model for applications in the vehicle, the W3C promotes the VSS ontology as the data model for applications in the cloud.[13]

7) "DESIGN OF AN EVENT-BASED ARCHITECTURE FOR THE INTRA-VEHICULAR CONTEXT PERCEPTION" [120]

Terroso-Saenz *et al.* aim to determine the context in the vehicle to provide personalized and intelligent services in the vehicle and infrastructure. The focus here is on identifying travel routes concerning the places visited, classifying them, and deriving information about the vehicle occupancy [120], [121]. A Complex Event Processing (CEP) approach handles data from different sources and identifies elements from a defined context model (see Figure 13). The model consists of dynamic (location, activity, time, and identity) and static (vehicle owner/features) context. In [120] an architecture with five context modules is designed to extract the different aspects. The vehicle state module uses a fuzzy logic approach to determine whether the vehicle is in motion or has come to a stop. In addition, the Occupancy module determines passenger groups based on the behavior when boarding the vehicle. Finally, the vehicle location coordinates are clustered, assigned to landmarks, and connected to routes between landmarks with itineraries or itinerary segments.

8) "AUTOMOTIVE CONTEXT-AWARE POLICY SYSTEM FOR CAR CONNECTIVITY REQUESTS" [122]

In her work, Copeland describes the Context-Aware Policy System (CAPS) for vehicle connectivity based on enterprise policies. The system aims at the fact that vehicle manufacturers operate a fleet of devices that are connected to the mobile network and as such incur costs for mobile network connectivity. Therefore, awareness of the car context shall encourage the selection of the best available access network and roaming network, e.g. to reduce these costs and mitigate risks for commercially sensitive data. The author's important statement is that the vehicle context must be distinguished from the driver context, as different drivers must be expected in the case of shared mobility. Moreover, the different data sources for context are emphasized in detail, listed separately for each stakeholder of the system's goal (carrier, vehicle

[11] https://www.pegasusprojekt.de/de/

[12] e.g. Semantic Sensor Network Ontology (SOSA)https://www.w3.org/TR/vocab-ssn/

[13] https://www.w3.org/auto/

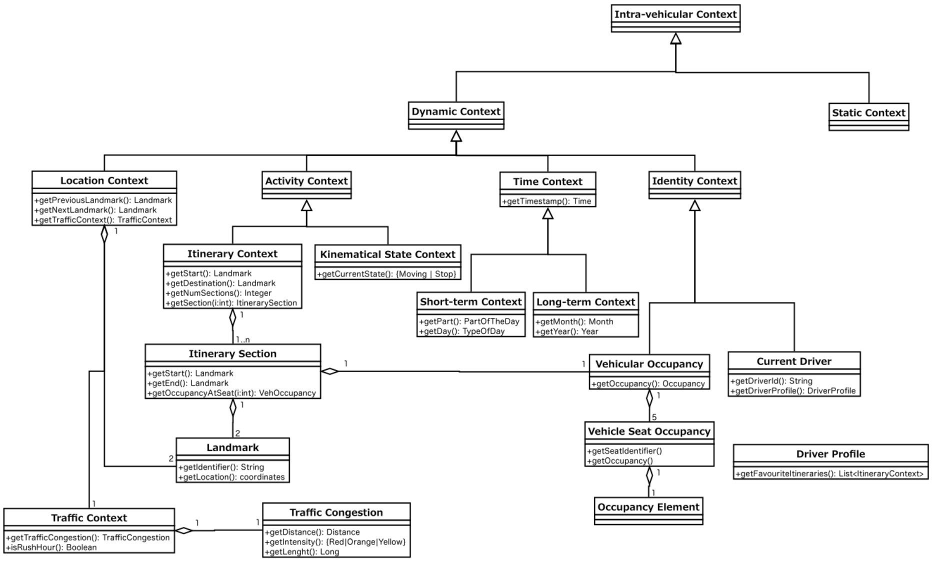

FIGURE 13. Intra-vehicular context of Terroso-Saenz *et al.* [120].

TABLE 3. Key factors of a vehicle's context [122].

KF	Factor	Description
S	Space	Tagged location, Abroad,
T	Timing	Work / leisure, Scheduled
A	Activity	Direction / Speed / Stationary, business trip, garaged
U	Urgency	Accident / breakdown , away-from-home/office,
C	Charging	Bands for roaming, bands per car type…
I	Integrity	Apps Risks, Confidentiality, Discrepancies
D	Destination	Apps / website / P2P , approved / untrusted, Corp. apps
N	Network type	WiFi hotspots, Enterprise car-park, Mobile 3G/4G
M	Media type	Video / voice / text, Duration / Volume

service provider, car/fleet). The publication mentions nine key factors (KF) (see Table 3).

9) "SEMANTIC COMPARISON OF DRIVING SEQUENCES BY ADAPTATION OF WORD EMBEDDINGS" [123]

Ries *et al.* adapt ideas from Natural Language Processing to enable the semantic interpretation of driving data in their respective contexts. The Word2Vec approach, which was developed initially for embedding words in language processing tasks, is transferred to automotive data sets. To enable this, they convert multivariate time series from e.g. test drives for the development of autonomous driving algorithms into discrete driving states using a binning approach. Subsequently, a dimension reduction is performed using a semantic embedding with neural networks (see figure 14) to obtain a low-dimensional space instead of the high-dimensional input space. In this space, distances between two journeys can be calculated and semantically similar journeys automatically arrange themselves with a small distance between them. However, there is no ad-hoc grouping of the embedded space into meaningful or coherent regions or classes. The analysis and interpretation of why driving sequences are similar and *close* to each other must be made manually. In addition, the binning of the signals must be specified. Therefore, this approach is only feasible for a small number of signals. Nevertheless, this is a helpful approach to abstract the physical behavior of the vehicle. Whether the neural networks are suitable for use in the vehicle, however, remains open.

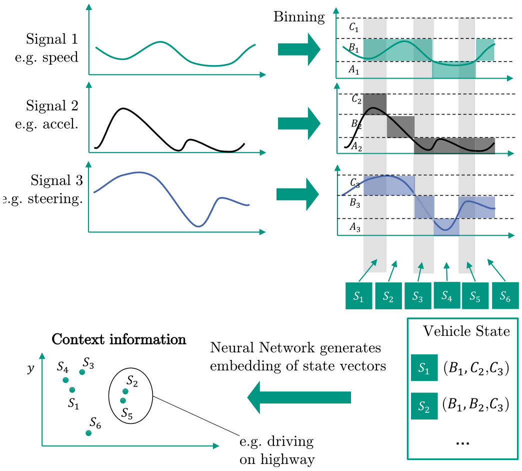

FIGURE 14. Semantic embedding in word2vec manner [123].

10) "UNSUPERVISED DRIVE TOPIC FINDING FROM DRIVING BEHAVIORAL DATA" [125]

Bando *et al.* [125] and their preliminary work in [124] emphasize an Unsupervised Machine Learning method for

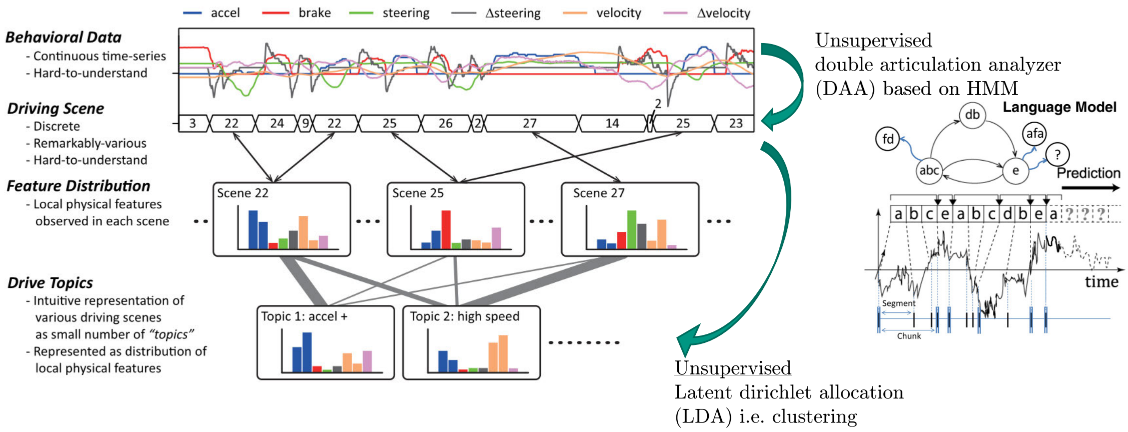

FIGURE 15. Unsupervised data aggregation [124], [125].

segmenting driving behavioral data into driving scenes and likewise in drive topics (see Figure 15). As the first step, they use a method based on Hidden Markov Models (HMM) called double articulation analyzer (DAA) to convert the behavioral data to driving scenes [124]. However, the authors elaborate that for one hour of driving, the algorithm finds more than 400 different driving scenes, which are discrete but not intuitive to understand. Thus, as the second step, they propose to use Latent Dirichlet Allocation (LDA), to group the scenes to a manageable number of drive topics. The clustering is performed according to the distribution of the physical input features. Additionally, each topic's name (a label) is found automatically as the most frequent feature of the topic's statistical distribution. A significant challenge with machine learning methods is that a large data set must be available to represent the totality of the driving scenes.

D. IT NETWORK SECURITY

The range of publications that mention *context*, *context-awareness* or *situation-awareness* for known security mechanisms in enterprise IT and mobile devices is very wide. Both areas are generally not safety-critical and are therefore not subject to the strict real-time requirements that we have in the automotive and other industries. Only a small percentage of publications deal with systematic context-awareness, many publications consider only one item of contextual information (e.g. a variable and hence adaptive blacklist of IP addresses [126]), few items (e.g. location [127]), or do not give more precise information about which context information is used. The most frequent applications of context-awareness include Intrusion Detection Systems and Access Controls.

1) "ONE SIZE DOES NOT FIT ALL: 10 YEARS OF APPLYING CONTEXT-AWARE SECURITY" [65]

In their overview work, Sinha *et al.* [65] emphasize three elementary types of context information. The authors describe the vulnerability profile, which comprises the set of known vulnerabilities relevant to the system (machine, OS, applications). Secondly, the attack surface is mentioned. On the one hand, this comprises the occurring attacks and, in addition, all kinds of information about the attackers, such as their goals and methods. The last aspect is the usage model, which aims at prioritizing the defense according to the importance of the services and applications. In addition to the frequency of use of the services, information such as the importance of the data and the opportunity costs in case of failure are referred to. For the application of context information to the use cases already outlined in Section IV-B, there are two key challenges according to the authors: diversity among networks and the dynamic nature of context. This is also true for the application in vehicles.

2) "ONTOSECURE: A SEMANTIC WEB BASED TOOL FOR NETWORK SECURITY STATUS PREDICTION" [128]

Bhandari and Singh introduce the tool *OntoSecure*, which is based on ontologies to improve the network security situational awareness. The system context in the form of ontology is intended to support the system administrator in making the right cyber-defense decisions. Based on the created OWL models for network, services, and attacks, as well as dynamic context updates based on SWRL rules, an abstract network security status is determined and predicted (see Figure 16). The SWRL rules are used to process the inputs in the form of configuration changes, new vulnerabilities, and attacks and to update the ontology. Further details of the underlying ontologies are not presented.

3) "DROID MOOD SWING (DMS): AUTOMATIC SECURITY MODES BASED ON CONTEXTS" [130]

Iqbal and Zulkernine introduce a component for the Android mobile phone operating system called Droid Mood Swing (DMS) that detects the phone's context. The context defines security modes, which enable automatic permission restrictions, an inter-process firewall, restricted network access, and file system access control. DMS is based on the author's work *Flamingo* [129] (see Figure 17). The context parameters used are described carefully (see table 4), but

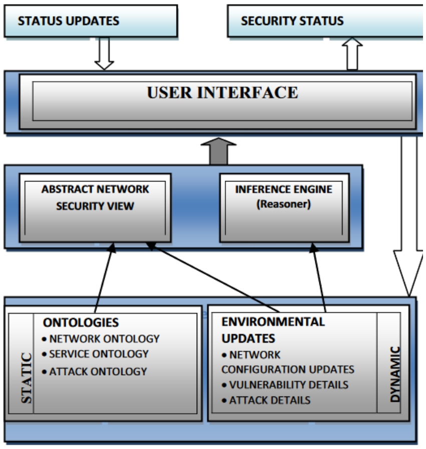

FIGURE 16. Overview of the OntoSecure system [128].

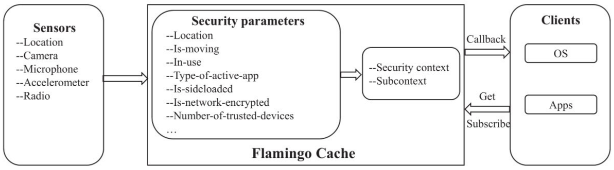

FIGURE 17. The Flamingo framework for security context management in Android smartphones [129].

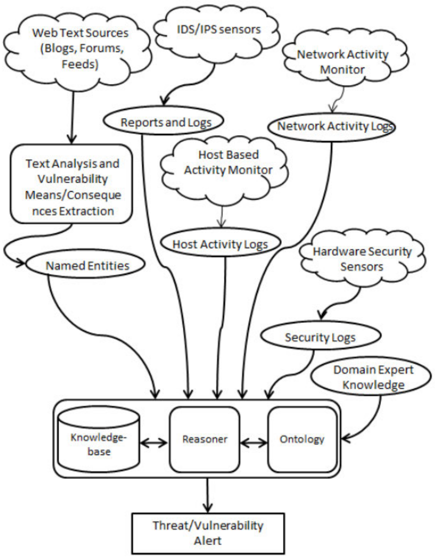
Web Text Sources (Blogs, Forums, Feeds)
IDS/IPS sensors
Network Activity Monitor
Reports and Logs
Network Activity Logs
Text Analysis and Vulnerability Means/Consequences Extraction
Host Based Activity Monitor
Host Activity Logs
Named Entities
Hardware Security Sensors
Security Logs
Domain Expert Knowledge
Knowledge-base
Reasoner
Ontology
Threat/Vulnerability Alert

FIGURE 18. Situation aware intrusion detection [133].

they are not formally structured (e.g., with an ontology or other model). Different sub-contexts can be active for the superordinate security context that is determined based on the location. The mechanism for temporarily holding the current context in DMS is not described in greater detail.

TABLE 4. Context parameters of Flamingo [129].

Context parameter	Possible values	Sensors and sources
Location	Home, Office, Outdoor	GSM, WiFi, GPS
Is moving	Yes / no	activity = still
Type of activity	in vehicle, on bicycle, on foot, running, still, tilting, walking, unknown	Google Activity recognition API
Location place type	restaurant, gym, park, café, hospital, etc.	Google Places API
In use	on, off on / off status of screen, status of microphone and speaker	
locked	Yes / no	
type of active app	type from Google play store	
active app is sideloaded (not from Google Store)	Yes / no	
Network type	2G, 3G, LTE, WiFi	
Network encrypted	Yes / no	
Camera on	Yes / no	
Microphone on	Yes / no	
Storage encrypted	Yes / no	
Number of trusted devices		
Financial app	Yes / no	Type of active app = financial
Security Context	Sub-context	Sources
Home	Casual, Private, Financial, Sideloaded	Camera, Microphone
Office	Casual, In-meeting, Financial, Sideloaded	meeting schedule from calendar
Outdoor	Place-type, Financial, Sideloaded	

4) "ENHANCING THE ACCURACY OF NETWORK-BASED INTRUSION DETECTION WITH HOST-BASED CONTEXT" [131]

In an early work, the open-source Bro Network Intrusion Detection System, which is now known as *Zeek*[14] is extended with context information from the hosts that are part of the monitored network. The authors deal specifically with the monitoring of HTTP traffic, as this makes up the majority of Internet traffic. The advantage of context extraction on the host is that, unlike IDS, the host can decrypt the traffic,

[14] https://zeek.org/

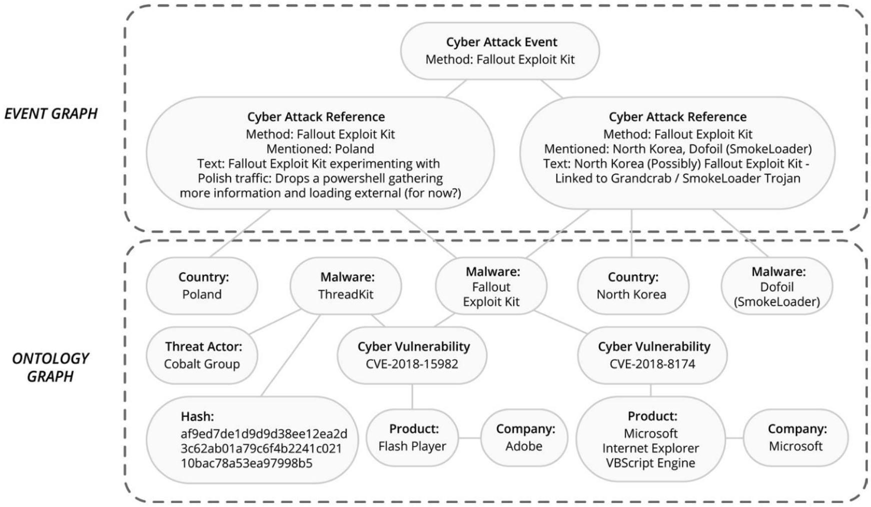
EVENT GRAPH
Cyber Attack Event
Method: Fallout Exploit Kit
Cyber Attack Reference
Method: Fallout Exploit Kit
Mentioned: Poland
Text: Fallout Exploit Kit experimenting with Polish traffic: Drops a powershell gathering more information and loading external (for now?)
Cyber Attack Reference
Method: Fallout Exploit Kit
Mentioned: North Korea, Dofoil (SmokeLoader)
Text: North Korea (Possibly) Fallout Exploit Kit - Linked to Grandcrab / SmokeLoader Trojan
ONTOLOGY GRAPH
Country: Poland
Malware: ThreadKit
Malware: Fallout Exploit Kit
Country: North Korea
Malware: Dofoil (SmokeLoader)
Threat Actor: Cobalt Group
Cyber Vulnerability CVE-2018-15982
Cyber Vulnerability CVE-2018-8174
Hash: af9ed7de1d9d9d38ee12ea2d 3c62ab01a79c6f4b2241c021 10bac78a53ea97998b5
Product: Flash Player
Company: Adobe
Product: Microsoft Internet Explorer VBScript Engine
Company: Microsoft

FIGURE 19. Security intelligence graph [134].

process the HTTP request thoroughly, and knows all redirections to addresses other than the one originally requested. On this basis, the IDS is provided with the following context information: the hosts and TCP ports, the original request string, the final URL after redirection, the name of the file being served, and the HTTP reply code.

5) "ConXsense - AUTOMATED CONTEXT CLASSIFICATION FOR CONTEXT-AWARE ACCESS CONTROL" [132]

Miettinen *et al.* introduce a framework for context-aware access control on mobile phones which is called *ConXsense*. The framework's contextual information is the GPS-based location context of visited areas, the set of visited WiFi areas, and the social context of the device. The social context is defined by the Bluetooth devices that are typically around the device of interest. In addition to very detailed explanations on the calculation of the individual context information, the authors determine statistically derived information from the mentioned context, e.g. maximum total visit time of any GPS-based context, average encounter time of familiar Bluetooth devices. These features are used as input for a machine learning approach, which classifies the context into secure and insecure or public and confidential respectively.

6) "A KNOWLEDGE-BASED APPROACH TO INTRUSION DETECTION MODELING" [133]

An approach for a situational aware Intrusion Detection System is presented by More *et al.* [133]. The system uses an ontology with base classes *means, targets* and *consequences* for creating a knowledge base and detecting intrusions that are related to known vulnerabilities or attack methods. With different data sources, either properties of ontology classes are added (e.g. IP address) or object triples of a class are generated (e.g. a specific buffer overflow attack), and the knowledge base is filled. Figure 18 shows the sources and some of the information integrated into the ontology and knowledge base.

E. SIEM SYSTEMS AND ALERT CORRELATION

Besides integrating context information in security measures such as access controls, context is especially required in the process of intrusion response. Here, alerts are analyzed and correlated with other alerts. However, to gain insight into what an alarm could mean for an enterprise or network, for example, to answer the questions *was this an attack*, *what impact does the attack have*, *what is affected by it*, additional information is required. The selected publications are all based on the common idea of a security knowledge graph, which is in some way supported by one or more ontologies. The ontologies define the vocabularies and relationships for performing the respective actions, e.g. attack reconstruction, alarm aggregation, identification of malicious entities. Nevertheless, the information contained and the applications differ.

1) "THE SECURITY INTELLIGENCE GRAPH: WHITE PAPER" [134]

Recorded Future provides context about threat actors and captured information of the internet with its *Security Intelligence Graph*. The graph-based data structure shall enable analysts as well as algorithms to pivot through the stored relationships, e.g. vulnerabilities, threat actors, or organizations associated with the actors. The security intelligence context is subdivided into an ontology graph and an event graph which are connected (see Figure 19). While the ontology graph captures the static information about entities in the cybersecurity landscape (companies, countries, IP addresses), the event graph contains the real-world cyber attacks.

2) "DEVELOPING AN ONTOLOGY FOR CYBER SECURITY KNOWLEDGE GRAPHS" [135]

Another approach for a knowledge graph is presented by Iannacone *et al.* [135]. The STUCCO ontology is developed

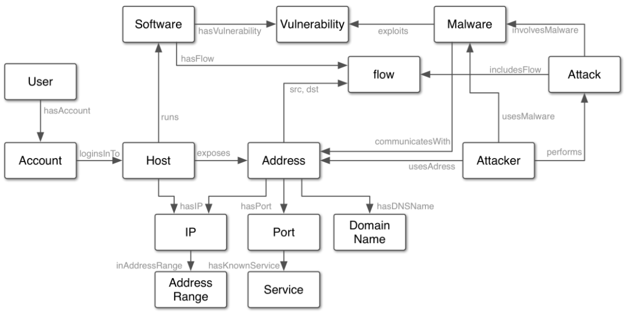

FIGURE 20. STUCCO ontology [135].

as the schema for the integration of different structured and unstructured information sources in the knowledge graph and is part of an open-source cyber intelligence platform.[15] The ontology consists of 15 entity types and 115 properties (see Figure 20). In comparison to other approaches, the authors use the GraphSON format for describing the triples of their knowledge base instead of RDF triples. The graph shall serve for incident response, e.g. searching for impacted hosts and flows, and automated response such as adjusting firewall rules.

3) "ONTIDS: A HIGHLY FLEXIBLE CONTEXT-AWARE AND ONTOLOGY-BASED ALERT CORRELATION FRAMEWORK" [136]

The ONTIDS framework for alert correlation is presented by Sadighian *et al.* [136], [137]. The authors aim to reduce the false alarm rates of Intrusion Detection Systems by integrating available context information. ONTIDS is based on an ontology for the representation and storage of alert information, alert context, vulnerability information, and attack scenarios (see Figure 21). For integrating the alert data of different IDS systems, the IDMEF[16] is used. The correlation is based on SQWRL and SWRL rules that implement two steps: context- and vulnerability-based filtering of given alerts to determine possible attack scenarios and afterward, try to match a sequence of previous alerts to reconstruct the entire attack scenario.

4) "SEMANTIC AWARE ATTACK SCENARIOS RECONSTRUCTION" [138]

Saad and Traore outline an attack scenario reconstruction approach based on alerts of multiple IDS systems. With their concept, the alerts that are related and thus form one specific attack scenario are identified. The authors build on a self-defined intrusion ontology in whose format the alerts are initially transformed. More details on the ontology can be found in an earlier work of the authors, see [139]. Interestingly, the authors calculate a semantic relevance between two alerts a and b in the ontology (e.g. *has the same goal*) based on the number of relations between these alerts and construct a graph from it. Attack scenarios are extracted from the graph as maximum cliques. Another transformation step converts a group of related alerts found into a new graph, which takes causal relationships into account. A causal relationship between two alerts a and b exists if the effects of a have a large intersection with the required preconditions of b. Besides, the authors describe a step for filtering false alerts based on environmental data such as network topology or security policies, however, they do not emphasize how to do that in detail.

5) "A PRACTICAL APPROACH TO CONSTRUCTING A KNOWLEDGE GRAPH FOR CYBERSECURITY" [140]

Jia *et al.* introduce an approach to construct a cybersecurity knowledge graph based on an ontology that is automatically populated from structured and unstructured information. A machine learning approach called conditional random field is implemented to extract elements (i.e. classes, instances for the ontology) from the data sources, e.g. vulnerability and attack databases. Additionally, the authors employ rules to find new relationships and attribute values.

6) "UCO: A UNIFIED CYBERSECURITY ONTOLOGY" [141]

The increasing number of existing cybersecurity ontologies and open knowledge bases was unified by Syed *et al.* in their Unified Cybersecurity Ontology (UCO) [141]. The authors reuse their preliminary work (e.g. the STUCCO ontology of [135]) as well as the Structured Threat Information eXpression (STIX) format [142] and prominent standards and ontologies, e.g. CVE, CCE, CVSS, CAPEC, CYBOX, KillChain, DBpedia [75], and Yago [143]. The authors describe their unified ontology's basic classes and how they map to classes in the reused ontologies. Furthermore, exemplary SPARQL queries are discussed, that search the available triples of information in selected use cases, e.g. *Suggest similar software to given software* and *Vulnerabilities associated with PDF Readers*.

7) "RelExt: RELATION EXTRACTION USING DEEP LEARNING APPROACHES FOR CYBERSECURITY KNOWLEDGE GRAPH IMPROVEMENT" [144]

Pingle *et al.* base on the work of Syed *et al.* (see Section V-E6) and define UCO 2.0. This ontology is using classes and relationships defined in STIX 2.0. Accordingly, their work describes attacks and the techniques, actors, and indications for such. The focus of the work, though, is that the relations described in the ontology are obtained from publicly available security texts (e.g. social media) using machine learning. The proposed system called *RelExt* consists of three steps. First, a Named Entity Recognizer extracts security-relevant terms from the text. Then, a Word2Vec model trained on cybersecurity terms transforms the textual terms in vectors of fixed length. Finally, each two of the extracted vectors are classified using a feedforward neural network concerning the relation existing between them.

[15] http://stucco.github.io/

[16] Intrusion Detection Message Exchange Format

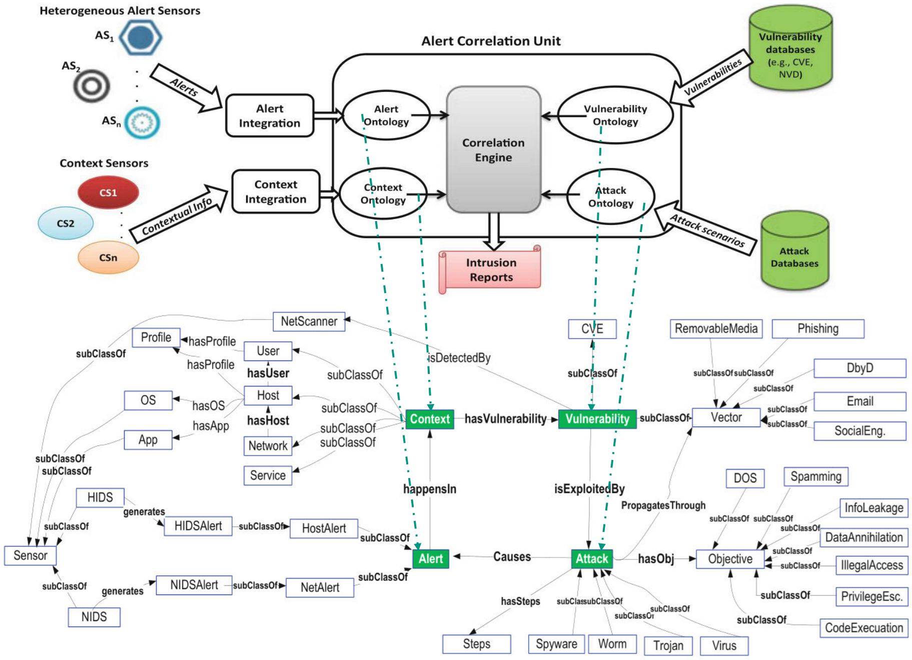

FIGURE 21. ONTIDS alert correlation and ontology [136].

8) "The SEPSES KNOWLEDGE GRAPH: AN INTEGRATED RESOURCE FOR CYBERSECURITY" [145]

While either most of the existing cybersecurity ontologies published in the literature are not available for download and the respective security knowledge graphs are not populated with actual knowledge extracted from the web and databases, Kiesling *et al.* aim to close this gap. Their SEPSES knowledge graph [145], as well as the underlying ontology, are freely available for download and query. In addition, a process is implemented that frequently accesses public databases such as CVE, NVD, CVSS, CPE, CWE, and CAPEC to automatically update and broaden the knowledge base. In contrast to other publications, the knowledge graph is not focused on solving specific problems, e.g. vulnerability analysis. Therefore, the knowledge base must be extended to include the appropriate information for each use case. However, the authors present examples from intrusion detection and vulnerability assessment using their knowledge graph.

9) "MalRank: A MEASURE OF MALICIOUSNESS IN SIEM-BASED KNOWLEDGE GRAPHS" [146]

Finally, likewise, the recent work of Najafi *et al.* introduces a knowledge graph for SIEM systems [146]. The schema of the knowledge graph is not very comprehensive but well emphasized and spans information from system logs and cyber threat intelligence. On the side of the logs, proxy, DNS, and DHCP logs are integrated, while the open databases integrate IP Ranges, ASN (Autonomous System Numbers), X.509 certificates, DNS Resource Records, and known malicious domains and IPs. However, the publication focuses on a graph-based algorithm called MalRank, which identifies malicious entities in the knowledge graph. Based on a set of known malicious entities, a precise scoring metric is used to infer a rating of the other entities.

F. SECURITY CONTEXT FOR THE (INDUSTRIAL) INTERNET OF THINGS

Since the Internet of Things and industrial control system devices are subject to similar resource constraints as in the automotive field, it is worth taking a brief look at adaptive and context-aware security for this field. In addition to conceptual work, practical approaches to e.g. authentication are being worked on.

1) "CONTEXT-AWARE SECURITY SOLUTIONS FOR CYBER-PHYSICAL SYSTEMS" [77]

Wan and Alagar outline their idea of context-aware security for cyber-physical systems (CPS). As the vehicle is a CPS as well, such concepts may be of interest to us. Besides defining properties of context information for CPS (see Section IV-E), the authors mention system context, user context, environmental context, temporal context, and physical context as the relevant categories of context information. It is expected that specific ontologies are used to model this information. However, they do not specify one. Instead, the authors focus on a formal notation of context, which is used to describe access control policies.

2) "HANDLING A TRILLION (UNFIXABLE) FLAWS ON A BILLION DEVICES" [147]

An overview of open challenges in securing the IoT world was given by Yu *et al.* They emphasize the need for context-aware security policies because of the strong dependencies between the different IoT devices. The publication follows a brute force approach to include context-awareness in policies by specifying the set of possible states the entire system may be in as $|S| = \prod_{i,j} |C_i| \times |E_j|$. C_i is the security context of the devices, e.g. *normal, suspicious* and E_j are the environmental context information with discrete values, e.g. temperature = high/low. While being expressive, this approach yields a possible explosion of system states with a rising number of devices.

3) "AN ONTOLOGY-BASED SECURITY FRAMEWORK FOR DECISION-MAKING IN INDUSTRIAL SYSTEMS" [148]

Mozzaquatro *et al.* performed a literature review of IoT technology, information security, ontologies, and other knowledge representation schemes and used the MENTOR method to derive an ontology for IoT Security called IoTSec [149]. Their reference ontology includes assets, threats, vulnerabilities, and security mechanisms. In [148], this ontology is used in a security framework for IoT devices to enable adaptive security mechanisms. The mechanisms are based on rules, e.g. access control and authorization. However, the authors don't emphasize the adaptation mechanisms. Besides, the approach to data collection to populate the knowledge base at runtime is only superficially presented.

4) "NETWORK SECURITY SITUATION AWARENESS BASED ON SEMANTIC ONTOLOGY AND USER-DEFINED RULES FOR INTERNET OF THINGS" [150]

The idea of network security situation awareness (NSSA) for IoT is introduced by Xu *et al.* [150]. The maturity of the NSSA can be divided into three levels: perception, evaluation, and prediction. The authors propose a situation reasoning method based on an ontology and various manually defined SWRL rules to integrate the heterogeneous information from different sources, addressing the first two levels of NSSA. In addition to categories frequently used in other publications such as attacks, alerts, vulnerabilities, and configuration (here named with context), the specially defined ontology includes Netflows (see Figure 22). The publication does not deal with the implementation or experimental evaluation in depth.

5) "PUTTING THINGS IN CONTEXT: SECURING INDUSTRIAL AUTHENTICATION WITH CONTEXT INFORMATION" [151]

The group of Anton *et al.* addresses with several publications the security of industrial IoT and SCADA environments [151]–[153]. In [151], the authors focus on including context information in the authentication mechanism RADIUS. They use a simple scheme of location, access time, and the criticality of the action that shall be authorized (i.e. default access or privileged access). The location context is provided by the IP address, allowing for differing between on-site and off-site access. In [152] they present an aggregation model of industrial IoT data, which the authors call *context-based* data. They introduce three levels of abstraction, with raw network packets on the first level, network flow on the second, and a network flow graph on the third level. Furthermore, the authors aim to collect the *effect* of network traffic in the form of e.g. machine log files as well as data for the *causes* of the traffic, e.g. configuration logs, or application data. However, the model is straightforward and has to be concretized to fulfill a specific use case.

6) "MANAGING CONTEXT INFORMATION FOR ADAPTIVE SECURITY IN IOT ENVIRONMENTS" [154]

Hernandez-Ramos *et al.* are developing a security framework that can be used, for example, to adapt the identity management and authorization of IoT devices to a situation. Without going into detail about the intended contextual information they want to use, the authors plan to use the modeling language SensorML[17] in its JSON form to describe the IoT data. The authors argue that this approach is less resource-intensive than modeling the data with ontologies and therefore better suited for IoT. Furthermore, Complex Event Processing is envisaged as a technology for concluding high-level contexts. However, the work remains on a more theoretical level.

G. SIMULATION AND EVALUATION

As the literature analysis has shown, some works remain on a theoretical level and do not deal with evaluation. Context-awareness is a challenge in this respect. To check the validity and performance of methods to derive context from data, it has to be defined which raw information is needed. Above all, a data set with this raw information and ideally known context has to be available. Such a data set should therefore contain on the one hand the issue of security via attacks and alarms, on the other hand, the vehicle and its network communication as well as its environment. In addition, data should not only be available for one vehicle but a fleet. Otherwise, solutions for analysis in SIEM systems can only be validated to a limited extent. However, there is a lack of public data sets with attacks [15] in the area of automotive security, which makes it very difficult to compare techniques. Data sets of vehicle-internal communication are generally hardly available, not in realistic complexity (e.g. no Ethernet and traffic from service-oriented communication), or cannot be used because the decoding of the payload data is manufacturer-specific. Special attention in the development of context-adaptive security measures will therefore have to be paid to the simulation of such data sets. For example, simulations are usually used to investigate security measures for VANETs [82], since almost the same problem arises here.

We do not want to go into detail about individual publications here, but rather name some popular simulators that are maintained and refer the reader to them. Ahmed *et al.* give a comprehensive overview of simulators for VANETs and emphasize three categories of simulators: mobility generators, network simulators, and VANET simulators, where the latter is either a standalone product or a framework combining simulators from the other two categories. In addition, we look at the publication of Lekidis and Barosan, which consider simulators for vehicle internal networks [155]. Besides

[17] https://www.ogc.org/standards/sensorml

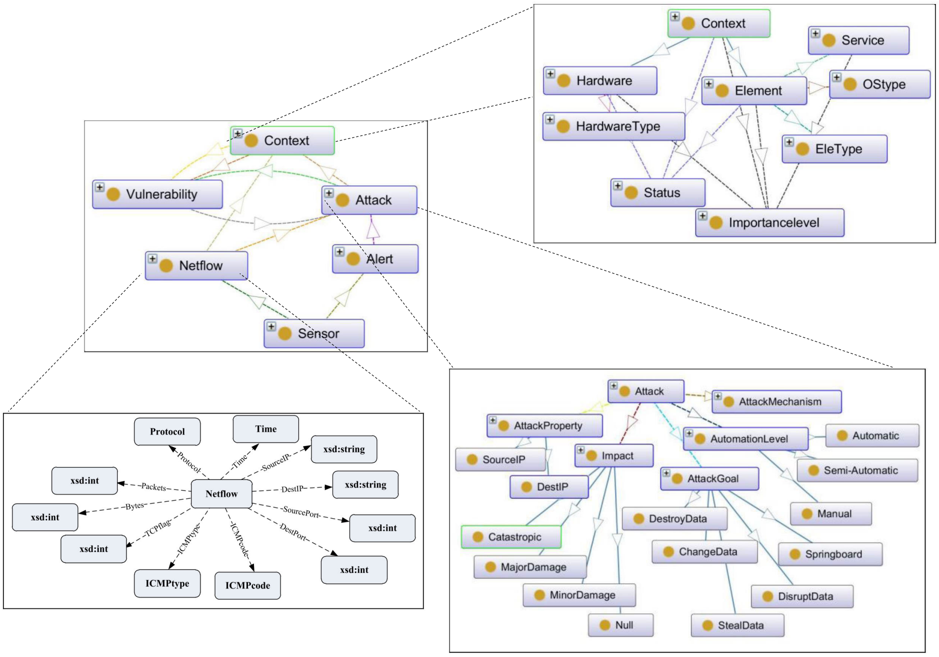

FIGURE 22. **Ontology for network situation awareness including the sub-concepts [150].**

in-vehicle networks, external communication, and mobility of fleets, the vehicles' physical environment has to be simulated to some extent, to yield the driving behavior. For this purpose, we refer to the publication of Tong *et al.* [156], where simulators and tools for this purpose are compared. We include CARLA [157], LGSVL,[18] and IPG CarMaker[19] in our listing (see Table 5). Moreover, attacks need to be simulated, e.g. generated in a reproducible way. Evaluation platforms may provide a way to generate such attacks, e.g. [158], [159]. Nevertheless, the solutions known from enterprise IT such as Kali [160] and Metasploit [161] are increasingly important due to the transformation of the in-vehicle architectures towards IT networks.

H. PRIVACY

Finally, also privacy is a field we have to look into when developing context-aware security for vehicles and fleets. Vehicle data initially belongs to the user or owner of the vehicle and includes a varying degree of sensitivity [162]. Particular protection is required for direct communication of the user via the vehicle, e.g. telephone, user's calendar as well as personal data and preferences, such as the identity and usage patterns of vehicle applications. Less sensitive are data about the environmental conditions or the technical condition of the vehicle [162]. Since even a speed signal is sufficient to trace PII such as location [163], the transmission of raw data from the vehicle to a backend to establish behavior monitoring is very limited. For processing and storage, especially outside the vehicle, either explicit authorization must be granted or legal regulations must require it. Otherwise, suitable anonymization or pseudonymization measures must be used to remove personal data and ensure privacy. In pseudonymization, identification features are replaced by pseudonyms (e.g. arbitrary code). With the help of this mapping principle, pseudonymized data can be traced back to the original data set, if required. In contrast, anonymization procedures do not allow the identity to be traced back. In the literature different possibilities for this are presented [164]:

- Remove the sensitive attribute from the dataset, e.g. the *vehicle identification number* (VIN). However, it may not be sufficient to remove only characteristics, if identification can be made from other available side-data sets.
- k-anonymity [165]: In k-anonymity, multiple instances are grouped in such a way that at least k instances form each group and share some quasi-identifiers. The quasi-identifiers have the same value for all k instances in each group. To turn attributes into such quasi-identifiers, values must be generalized, i.e., instead of speed $10.42\frac{m}{s}$, for example, the speed is specified in a range, e.g., speed between $10\frac{m}{s}$ and $20\frac{m}{s}$.
- l-diversity and t-closeness [166], [167]: l-diversity complements k-anonymity to ensure that the sensitive attributes within a group with the same quasi-identifiers are different. This is important because otherwise the values of the quasi-identifiers can be used to immediately infer the value of the sensitive attributes. t-closeness furthermore attempts to approximate the distribution of sensitive values within a group to the distribution in the entire data set.

[18] https://www.lgsvlsimulator.com/

[19] https://ipg-automotive.com/products-services/simulation-software/carmaker-release-90/

TABLE 5. **Tools for fleet simulation.**

Category	Tool	Scope
Environment & physics	CARLA	open source, based on Unreal Engine
	LGSVL	open source, integrated with open source autonomous driving projects
	IPG CarMaker	commercial, sophisticated physics model
	MATLAB / Simulink	commercial, toolboxes for automated driving and vehicle dynamics, model-based systems engineering
Traffic	SUMO	open source, microscopic simulation, good compatibility and popularity
	PTV Vissim	commercial, microscopic, mesoscopic and hybrid simulation
	VEINS	open source, framework combining SUMO and OMNeT++
Network & Communication	Vector CANoe	commercial, interface to MATLAB / Simulink, specialized to all automotive in-vehicle protocols
	OMNET++	open source, various extensions e.g. for vehicle networks[20]
	NS-3	open source, Python and C++ API

Besides these approaches for anonymizing an entire dataset, *differential privacy* has been established as an alternative scheme that can be applied to streaming data, e.g. single instances of data sent from a vehicle to a backend. The basic approach here is to add a certain amount of noise to the data set. This allows to specify and ensure a mathematically precise level of privacy. Data that has been anonymized using differential privacy allows making statistical inferences about the entire data set. However, statements about a single dataset are only of limited use, since the added noise makes an exact answer difficult. Nelson and Olovsson provide a good overview of how the method can be used for vehicle data [168]. Recently, Hassan *et al.* published a general survey on differential privacy, which deals with cyber-physical systems [164].

In a certain way, the abstraction or generalization of raw data into contextual information helps to ensure privacy. Since this generally no longer involves fine-grained information, but only abstract values that specify ranges or categories, the de-anonymization of data is made more difficult.

VI. COMPARISON AND CONCLUSION

For the central step of this work, the comparison of the analyzed publications and the deduction of open research areas based on their results, it is necessary to define the comparison criteria. On the one hand, a comparison concerning the methods, objectives, and technologies used is necessary to identify typical approaches independent of the contextual information considered (see Section VI-A). In addition, for precise and fine-grained modeling of the relevant automotive security context, it is important to compare which context information was used in the considered work. We present the required taxonomy (see Section VI-B) before the comparison is discussed in detail.

A. TAXONOMY OF METHODS AND APPROACHES

To compare the different approaches and highlight interesting features, a comparative taxonomy is required (see Table 6). For this purpose, we take some parts of the notation of Perera *et al.* [16] with the concepts *modeling* and *reasoning*. We extend their taxonomy by the points *used technologies*, if the *resources* are *available* and the *scope* of the publication. The publications analyzed in this survey originate from application areas of context-awareness rather than being generic approaches. We therefore interpret the term reasoning more broadly than, for example, Perera *et al.*. In our work, in addition to the primarily ontology-based automatic reasoning methods (e.g., included in Apache Jena[21]) that are well known in the context-awareness domain, we consider reasoning to encompass in general the methods used to extract context from given data. To highlight publications that pursue a generic approach to context-awareness, e.g., through a context management component, or extract context that can be used in a way that is not only application-specific, we extend the taxonomy by adding the item *generic / management*. Items in brackets indicate partial fulfillment of the respective category.

B. TAXONOMY OF THE CONSIDERED CONTEXT INFORMATION CATEGORIES

For an analysis of the considered works, it is interesting from our point of view, in addition to the conceptual comparison such as the modeling methods used, which information was considered as context information for the respective work. To this end, it is necessary to form the superset or, depending on the application area, the subset of the utilized context information. The set of selected information helps us the answer our question 'what is context'.

After our initial analyses, we identified the following categories of context information that emerged in the publications we examined and are of interest to us:

[21] https://jena.apache.org/documentation/inference/

TABLE 6. Taxonomy for comparison of relevant publications.

Used Technologies and methods	List of mainly used techniques (e.g. SPARQL, MATLAB, ..), evaluation datasets and methods, if applicable
Generic / Management	✓: System contains a component for context management or context can be used generically and independent of application
Modelling	Key-value modelling (K), Markup schemes (M), Object oriented modelling (Ob), Ontology-based modelling (On), no explicit model (-)
Reasoning	supervised learning (S), unsupervised learning (U), semi-supervised learning (S/U), rules / hand-crafted (R), fuzzy logic (F), not defined (-)
Resources Available	✓: Results or resources of the project are available (e.g. context model is published online for re-use)
Scope of application	

1) TIME

Time is an essential factor when it comes to expressing the course of an attack, for example, or to describe that two vehicles are attacked simultaneously. For contextual information, simultaneously does not necessarily mean *exactly* at the same time, but for example on the same day. Thus, time context is more than just containing the exact current time.

2) ENVIRONMENT

Information describing the physical environment of the vehicle is considered as category ‘‘environment context’’. Since the physical environment can change very frequently while driving, environmental contexts are the more stationary or slowly changing parts of these environmental conditions, e.g. the weather.

3) LOCATION

Location context information is, for example, the current country we are in or some visited distinct locations. The location could have been considered as part of the environment. However, as publications tend to focus either on the physical conditions or the positioning and points of interest (POI) aspect, we separate these categories. Additionally, while the weather might change more or less frequently, it is changing very seldom if a specific GPS position is part of one country or another.

4) USER

The user category describes the contextual information about the user(s) of the vehicle. The term user can encompass much more than just the driver or passengers. Information about other persons related to the vehicle, e.g. mechanic, backend operator, can also be seen as context.

5) NETWORK

Information describing the communication network is a relevant context for security mechanisms, since attacks, for example, typically manifest themselves via network communication. Under network context, we include all information that concerns data flows on the one hand and the applications in the network on the other. Network context in the vehicle includes the Ethernet networks familiar from IT as well as context relating to legacy automotive bus systems such as CAN.

6) BEHAVIOR

For the vehicle as a cyber-physical system, information about the physical state and behavior has been included as context by various publications. In terms of security, attacks with physical impact may be critical to safety, and accordingly, behavioral information is relevant contextual information.

7) SECURITY ALERTS

Depending on the publication’s use case, security events and alerts triggered by security measures were considered as context information. Their history, origin, or the reason for triggering are highly immediate pieces of information that provide context for other security measures. Information about attacks, whether actually taking place or descriptions of typical methods, falls under this category as well.

8) SECURITY STATUS

Under the category security status we summarize here all context information, which was indirectly derived from the input data of the considered systems. For example, we include a metric that derives an overall system security status such as ‘secure’, ‘attacked’ based on alarms and network data as such context information.

9) VULNERABILITIES

The vulnerabilities category contains all contextual information describing a security weakness of the system, i.e. some way to get into the system or manipulate it. This can be, for example, a reference to public databases or a criticality metric that assesses the severity of a vulnerability.

10) CONFIGURATION

Since vulnerabilities are usually associated with specific versions and combinations of software, operating systems, or hardware, it is helpful to consider such information as a security context. Many publications that include vulnerabilities have therefore included some form of description of the system configuration, i.e. the static state of the system. In vehicles today, the configuration is relatively static after the start of production, though it will change dynamically in the future due to the increased use of OTA updates or customer functions that can be purchased on-demand.

11) EXTERNAL

Additional contextual information, that is not included in the aforementioned categories can be interesting context information in specific cases. For example, this could be data that resides outside the company's ecosystem and is retrieved as needed. Among these, we have taken traffic information, weather forecasts, or news stories such as tweets into account. Some publications focus exclusively on gaining context from such external sources.

C. COMPARISON FINDINGS BASED ON THE METHODOLOGICAL TAXONOMY

Table 7 illustrates the comparison of the around 50 papers we covered in the survey according to the taxonomy of Table 6. Works from the categories *Simulation* (see Section V-G) and *Privacy* (see Section V-H) are not under comparison here. In the first column, the methods and technologies used, and in the last column, the areas of application, the publications diverge widely, as already revealed in the detailed descriptions in Section V. The only exception is automotive security, where several papers combine intrusion detection systems with context-awareness.

However, it is particularly striking that in the area of automotive and fleet security, only two of the papers analyzed take a generic approach and only two others could be applied at least to some extent to a broader field of application (second column). In other areas, such as IT security and SIEMs, generic applicability is much more pronounced. The third column (modeling) shows a clear picture: either ontologies (On) are used, or no model is defined at all (-), but only superficially a set of relevant contextual information is mentioned. A few papers that at least exhaustively define the space of contextual information used in their work can be roughly classified into key-value approaches (K/-), without following stringent notations. Two publications stand out, which use a markup-based approach [101] and an object-oriented model [120], [121].

In the fourth category of comparison, the reasoning methods, most of the publications work either with a semi-supervised or rule-based approach. Most rule-based approaches are employed in combination with ontologies and implemented in the form of Semantic Web technologies such as OWL and SWRL. This follows the traditional approach from the context-awareness literature. Semi-supervised approaches, in contrast, tend to be used without an actual model. The popularity of this category of machine learning methods is deeply rooted in the underlying problem. In the security field, it is characteristic that quite a high amount of data of normal, inconspicuous behavior is available, while data with real failures, attacks, and intrusions is rare and, above all, not made public. Training of pure supervised classification methods is difficult in this case. It is worth noting that only a few publications combine rule-based approaches and other methods. Supervised methods have been used where labeled data is available, such as in the extraction of context from a text (e.g. [140]).

Although publications like to argue with the reuse of knowledge when selecting ontologies as a modeling method, only a few of the publications analyzed here basically follow this approach themselves. Only five of the papers make data, ontologies, or implementation freely available online. Particularly positive here are the works [101], which provide both code and simulation environment, and [145], which offer a publicly accessible database including query interface.

D. COMPARISON FINDINGS BASED ON THE TAXONOMY OF CONTEXT INFORMATION CATEGORIES

In table Table 8, the individual works are classified regarding the included context information. A filled circle indicates a particularly extensive or detailed consideration of the category. Unfilled circles indicate that the publication is based on information in this area, but has either not elaborated on it in detail, does not describe how and from where the context information is obtained, or only covers a very selected subset of the category. For example, if only the category 'behavior' is mentioned, but no more detail is given on which behaviors to differentiate, this is marked with an unfilled circle.

Overall, it is evident that no publication covers all areas of context information. This is understandable concerning the use of the information considered. Since some context information is more relevant in the vehicle, it was therefore considered more in publications on dedicated security measures. In contrast, others are more relevant in SIEM are more relevant in SIEM systems for correlation. Papers in the area of automotive context and data aggregation do not aim to increase security and, as expected, have low coverage of directly security-related categories such as alerts, vulnerabilities, and security status. Publications in the field of IT-SIEM systems, on the other hand, do not consider physical aspects relevant to our use case or information related to the non-virtual world. Attack descriptions sometimes contain information about attacker groups that typically use the attack method. However, such information has not been categorized by us as 'user' because it is not legitimate system users, but merely information related to the attack.

Some expected correlation between the analysis results in this section and the previous one can be noted. Works that do not define an explicit context model or only minimal lists of considered information (see Table 7) rarely achieve broad and detailed covered categories in Table 8. However, individual context aspects are described in detail in such works and methodically extracted from data. Furthermore, due to space limitations in publications, it is challenging to describe all categories of context information with the methods used to derive them from data when several categories are considered.

In the application area fleet security, almost only superficial information of companies has been considered so far. Accordingly, in this category, few contextual information has been elaborated in detail, but only categories have been

TABLE 7. Summary of literature reviewed in this survey.

	Reference	Technologies & Methods	Generic / Management	Modelling	Reasoning	Resources	Scope
Fleet security	[54]	Argus SOC		-	-	-	SOC, SIEM, processes and challenges
	[90]	Sand-sprinkled Isolation Forest, real CAN data		-	S/U		CAN IDS, SIEM
	[91]	(concept)		-	-	-	SIEM anomaly detection
	[92]	Hidden Markov, rule-based abstraction, SUMO		K/-	R, S/U		SIEM / backend IDS
Automotive Security	[93]	(concept)	✓	-	-	-	in-vehicle security architecture, context-aware modes
	[94]	Hidden Markov, Convolutional Neural Network, real CAN data		-	U, S		physical vehicle state
	[95]	SWRL, IoT-Lite Ontology, CAN data	✓	On	R		security context, anomaly Detection
	[97]	Neural Network, CAN data (OpenXC)		-	S/U		behavior abstraction, anomaly detection
	[98]	Neural Network, CAN data (OBD)		-	S/U		intrusion detection
	[99]	Neural Networks, image data, CAN data		-	S		intrusion detection
	[100]	N-gram distribution, CAN data (from [79])		-	S/U		intrusion detection
	[101]	XACML, ABAC, CAN	(✓)	M	-	✓	access control, diagnostic access
	[102]	SOME/IP, automata, reference monitors		-	R		(function) access control
	[103]	formal methods, cockpit demonstrator, virtual machines	(✓)	-	R		HMI display access control
Automotive Context and data aggregation	[108]	OSGi, OWL	✓	On	R		generic context-aware SOA
	[78]	UML	✓	On	(F)		Intelligent HMI, adaptive functions
	[110], [111], [113]	Fisher correlation score, real CAN data (BMW)	(✓)	-	U		Anomaly Detection, Intelligent Functions
	[117]	OWL Lite, SPARQL, Flask, Triple Pattern Fragments, OpenXC		On	R, O	✓	Semantic trajectories, vehicle data
	[114], [115]	SWRL, OWL-API[23]		On	R		scenario-based testing of autonomous vehicles
	[120], [121]	CEP, UbikSim, Esper, SUMO, clustering	✓	Ob	R, U, F		intelligent functions, itineraries
	[122]	Mobile networks, conceptual	(✓)	K / -	-		network access policies
	[123]	Word Embedding, Neural Networks, own test drive dataset		-	U		driving sequence comparison
	[124], [125]	Hidden Markov Model, Latent Dirichlet Allocation		-	U		driving sequence description

Continues on next page

[23] http://owlcs.github.io/owlapi/

TABLE 7. ***(Continued.)*** **Summary of literature reviewed in this survey.**

	Reference	Technologies & Methods	Generic / Management	Modelling	Reasoning	Resources	Scope
IT Security	[65]	address space monitoring, honeynets, blacklists		-	-		multiple security measures
	[128]	OWL, SWRL, SQWRL	(✓)	On	R		system administration, security status
	[129], [130]	ZoneDroid, androzoo database[170], real Android phone	✓	K	R		smartphone security mode management, firewall, file system
	[131]	Zeek, server HTTP traffic instrumentation		-	-		intrusion detection
	[132]	FlaskDroid, real Android phones	(✓)	-	R, S		access control, smartphone malware and misuse
	[133]	OWL, N3, web data streams, OpenCalais	(✓)	On	R		semantic intrusion detection
SIEM and alert correlation	[134]	web text sources, language processing	✓	On	S/-		graph for threat intelligence provision
	[135]	JSON, GraphSON, RabbitMQ, Titan Graph Database	✓	On	-	✓	knowledge graph development
	[136], [137]	OWL-DL, IDMEF, SWRL, SQWRL, Prelude, DARPA 2000 & UNB ISCX dataset	✓	On	R		alert fusion & verification, attack path reconstruction
	[138]	SWRL, OWL, Snort, DARPA 2000 dataset, Treasure Hunt Dataset		On	R		attack pattern identification, alert verification
	[140]	Linear CRF[24]	(✓)	On	S, R		vulnerability & attack knowledge base
	[141]	OWL-DL, RDF, various public ontologies, Apache Jena Fuseki	(✓)	On	R	✓	ontology unification
	[144]	web text sources, Named Entity Recognition, Word2Vec, STIX		On	S, U		knowledge graph entity identification
	[145]	RDF, SPARQL, caRML, Apache Jena, various open databases	✓	On	-	✓	open cyber security knowledge
	[146]	Apache Spark GraphX, Apache Kafka, Kubernetes, real-world enterprise SIEM data		(On)	S/U		maliciousness ranking of network nodes
Industrial / IoT sec.	[150]	SWRL, OWL, SPARQL	✓	On	R		network situation reasoning
	[148]	OWL, SPARQL, RDF, more conceptual	✓	On	R		adaptive rule-based IoT security
	[151]	RADIUS		K / -			IoT authentication
	[77]	formal semantics, theoretical work	✓	(On)	(R)		CPS access controls
	[147]	*µmboxes*, Squid, Snort		K / -			IoT security gateways and policies

[24] Conditional Random Fields

TABLE 8. Categories of context information mentioned in this survey.

		Context categories										
	Reference	***Time***	***Environment***	***Location***	***Person / User***	***Network / Applications***	***(physical) Behavior / Activity***	***Alerts / Events***	***Security Status***	***Vulnerabilities***	***(Vehicle) configuration / features, maintenance***	***External***
Fleet security	[54]	○		○				○		○	○	
	[90]	○		○	○	○		●		●		
	[91]			○			○				○	○
	[92]			○	○	○	○	○	○			
Automotive Security	[93]		○	○	○	○	○		○			
	[94]						○					
	[95]						○					
	[97]						●					
	[98]						●					
	[99]		●				●					
	[100]					○						
	[101]			○	○		○					
	[102]						○					
	[103]		○		○		○					
Automotive Context and data aggregation	[108]			●	●	○	○					○
	[78]		●	●	●	○	●	○				○
	[110], [111], [113]				○		●					
	[117]	○		○			●				○	
	[114], [115]		●				○					
	[120], [121]	●	○	●	○		○					○
	[122]	●		●		○	○		○			
	[123]						○					
	[124], [125]						○					
IT Security	[65]					●				●		
	[128]					○		○	○	○	○	
	[129], [130]		●	●		○	○		○			
	[131]					○						
	[132]			●	○	○						
	[133]					●		●		●		○
SIEM and alert correlation	[134]							○		●		●
	[135]					●		○		●	●	●
	[136], [137]					○		●		●	●	
	[138]							●		○	○	
	[140]							○		○		●
	[141]							○		●	●	●
	[144]									●	●	●
	[145]							○		●	●	●
	[146]					●			●			○
Industrial / IoT security	[150]	○				●		●	●	●	●	
	[148]					○		○		○		
	[151]	●		●	○		○					
	[77]	○	○	○	○	○	○					
	[147]		○						○			

mentioned. The work [90] is the only example in this category that goes into more detail. In contrast, in the area of vehicle security, there is a strong focus on behavior and, in some cases, on the context of the legacy CAN network. Here, in particular, the work goes into methodological detail about how data becomes context. An exception to this is [93], which takes a more theoretical approach but describes a broader coverage of categories. In the overview, it is noticeable that publications from the automotive security area do not yet cover the breadth and/or depth of information as is the case in IT.

In the area of automotive context-awareness and data aggregation, the picture is very uneven, since the works pursue very different application areas. However, they all have vehicle behavior in common to varying degrees of depth and sophistication. As far as the other categories in the table to the left of this are concerned, context-awareness of location and environment, for example, is vital for (autonomous) driving functions and is, therefore, more common. Some works also include external context here, such as data from other vehicles.

The context category network is particularly strongly represented in the application area of IT Security. It is interesting to see that certain papers from IT security also include environmental contexts. These are works in the field of mobile IT security. This manifests a difference from conventional network security, which does not focus on such a context. Compared to the publications in automotive security, there are conventional IT security publications that already integrate vulnerabilities or alarms as context.

Work in the field of SIEM and alert correlation provides a very homogeneous picture. Here, the external information is particularly strongly represented, the context information typically referred to as threat intelligence. However, the external context used in the area of SIEM and correlation is primarily to be seen as a data source for the other categories, such as vulnerabilities and attacks. For example, [144] focus just on the extraction of context from external sources. Especially the field of vulnerabilities and configuration is heavily used as context by the work in the SIEM environment. The publications considered here already go into a depth that fuels the actual usability of contextual information in practice. Network information, on the other hand, plays only a subordinate role.

Finally, the last area examined, industrial and IoT security presents a scattered picture. Overall, there is an increased focus on these cyber-physical systems' physical aspects, similar to the automotive environment. However, unlike the publications in automotive security, for example, [150] already goes into a detailed consideration of alert, vulnerability, and configuration contexts.

E. OVERALL CONCLUSIONS

Several things can be concluded from the current state of science and research. Overall, the following points lead to the conclusion that context-awareness in the automotive security environment is still very much in its infancy.

1) IMMATURITY IN USING A CONTEXT MODEL

The most important and essential question is 'what is the context', or 'in which context'. Research in the field of context-awareness in automotive security has only rarely used a modeling methodology. Context is thus not or only insufficiently defined. Progress in all areas of context-awareness can be expected when complete models allow an actual comparison of the used contexts of two systems.

2) IMMATURITY IN USING THE RIGHT MODELING METHODS

In application areas outside the automotive industry, ontologies have become strongly established to model context. In the SIEM environment, this has given rise to knowledge graphs, which are very well suited to the big data challenges and the integration of external knowledge. However, it is still an open question which modeling methodology is suitable for the real-time critical applications in the vehicle or the analyses of the fleet.

3) IMMATURITY OF UTILIZED CATEGORIES AND ASPECTS OF CONTEXTUAL INFORMATION

As can be seen in the publications of the other industries and application areas, the contexts considered are differing greatly. In general, however, there is a much broader and in some places deeper coverage of context categories than in the automotive security domain. When context models for vehicle and fleet security are defined in the future, care should be taken to take this breadth into account.

4) IMMATURITY OF THE INFERENCE METHODOLOGY

In the future, special attention should be paid to the improved inference of the current context from the given data. So far, strong separation of rule-based methods tied to ontologies and machine learning-based methods that do not feed into explicit context models can be observed in the literature, independent of the application area. Combinations of different approaches have not been pursued widely, while machine learning is gaining importance in the literature due to increasing data volumes. But of course, we have to keep it with Occam's Razor[24]: Of several satisfactory solutions to the same problem, the simplest theory is preferable. Thus, use machine learning where required but stick to more straightforward, hand-crafted rules where sufficient.

5) IMMATURITY OF APPLICABILITY

Context-awareness and semantic description forms of data originate from science. The more widespread use of Knowledge Graphs in the industry (e.g. [134]) or at least industrial research [169] finally shows the applicability of these ideas. However, many projects on ontologies or other forms of context description have remained dry theory and have never

[24] https://www.britannica.com/topic/Occams-razor

been able to prove in practice their usefulness. Work on automotive security context-awareness should always seek practical relevance (how can the modeled knowledge be applied) not to die virtuously. Practical usability includes validation of context-awareness, which can only be done within defined use cases. Different contexts may be relevant for different use cases (different vehicle architectures, functions, context users).

6) IMMATURITY OF REUSE

The previous two sections have shown that methods have already been identified for various context aspects to derive them from the vehicle and network data. Likewise, extensive context models have already been developed for security applications and the vehicle context. Thus, it is unnecessary to reinvent the wheel for automotive security context-awareness but to combine the existing state of the art.

In summary, these challenges form the research gaps that need to be filled. The first step in this direction can be to combine the context models for security and the vehicle domain known in the literature in a meaningful way to obtain initial substantive definitions of the vehicle and the fleet security context. An open dialog with industry and academia is needed to identify relevant categories of information in this effort. Methods to refine the context information used based on validation in real-world applications are also open research areas. In this paper, we present some use cases in the security domain at the beginning. With the increasing availability of data from vehicle fleets, the enterprise itself, and the ecosystem, other sub-fields of security may benefit from a semantic description of data in the future that we cannot yet imagine today.

VII. SUMMARY AND FUTURE WORK

The development of security measures such as encryption, intrusion detection, firewalls, or monitoring and correlation systems such as SIEMs is becoming increasingly urgent for the automotive industry. The introduction of legal regulations and the high risks for the safety of passengers and the immense financial impact of attacks have led to increased investments in the industry and many publications in this field. However, the vehicle's current security measures are designed too statically to deal with the constantly changing threat situation. At the same time, in the backend of the vehicle manufacturer or a third-party vendor, where the security alarms accumulate in the Security Operations Center, human analysts are faced with a pile of data from which no decisions can be easily deduced. Thereby, both in the vehicle and in the backend, it requires context information and context-sensitivity to better deal with the myriad challenges.

Overall, we would like to point out that we are still at the very beginning of the rather vague topic of context-awareness for security. But if we look ahead to the future, opening up vehicles to (authorized) access from outside is indispensable for new vehicle functions and the creation of value in the automotive industry. We therefore already need to start thinking about the future vehicle and fleet security. In addition to established development processes, a structured test strategy, rapid fleet-wide updates, and comprehensive monitoring, context-awareness can become an important component of security concepts.

This work thus serves to develop adaptive, intelligent security measures that adapt to the context of the vehicle and fleet. Our work makes a valuable contribution by defining context for vehicle and fleet security as well as characterizing challenges for which context information is necessary. We identified

- unlimited situation diversity,
- new vulnerabilities continuously,
- architectural complexity (ECUs, lines of code) and dynamics,
- cost constraints,
- enormous amounts of data,
- and diversity of variants

as the six key challenges. Some examples were given on the use of contextual information for security. In addition, we have outlined the main characteristics of context information and presented the established methods for context modeling.

The main contribution is a comprehensive survey conducted by identifying and summarizing work on context-awareness from the areas of vehicle onboard and fleet security, vehicle data, IT security and SIEM systems, and industrial security. The subsequent comparison of the considered publications revealed some open research areas and is intended to be a guidance for the future development of intelligent security in the automotive domain. A taxonomy of relevant categories of contextual information was established based on the comparison of the analyzed papers. The quintessence of the comparison is that the solutions currently published under the goal of context-awareness in the vehicle security domain are not yet genuinely context-aware. Neither do the current works reach the breadth of categories that the taxonomy spans, nor is it sufficiently clear for all aspects of context how it can be extracted from the amount and variety of data in the vehicle. However, the analysis also shows that substantial preliminary work has already been done in application areas outside of security as well as in other industries. This is where we can start in the future to improve security measures for vehicles and fleets and make them more intelligent.

Necessary steps for this have already been presented in the previous section of this paper. We hope that our survey will make it easier for successive work to get started on the topic of context-awareness, adaptivity, and intelligence for security measures. Other topics that should be focused on for the development of such measures in the future are privacy approaches as well as datasets and their generation utilizing simulations, and the reproducible generation of attacks to enable meaningful validation. In the future, we plan to develop explicit context models for the vehicle security field to achieve a fine-grained definition of the relevant context. Apart from that, it is necessary to extend the interfaces to security measures such as intrusion detection systems, access controls, and SIEM systems to enable the use of context information. Last but not least, more use cases for context information should be explored to gain the maximum benefit from the effort of semantically describing data from vehicles and fleets in the future. Another aspect that was outside the scope of this work is securing context information. When the context is used to control other security measures, it itself becomes an asset worth protecting.

REFERENCES

[1] G. Silberg, J. Anderson, A. Baritchi, D. Le, R. Plesco, and M. Krajecki. (2017). Protecting the fleet… and the car business: Today's cyber-physical threats disrupt automotive operating models. KPMG. Accessed: Aug. 17, 2020. [Online]. Available: https://assets.kpmg/content/dam/kpmg/jm/pdf/protecting-the-fleet-webfile.pdf

[2] Ponemon Institute LLC, IBM Security. (2019). *2019 Cost of a Data Breach Report*. Accessed: Aug. 17, 2020. [Online]. Available: https://www.ibm.com/security/data-breach

[3] C. Miller and C. Valasek. (Aug. 10, 2015). *Remote Exploitation of an Unaltered Passenger Vehicle*. Accessed: Mar. 1, 2021. [Online]. Available: http://illmatics.com/Remote%20Car%20Hacking.pdf

[4] M. Tischer. (2018). *The Computing Center in the Vehicle: AUTOSAR Adaptive*. Accessed: May 10, 2020. [Online]. Available: https://assets.vector.com/cms/content/know-how/_technical-articles/AUTOSAR/AUTOSAR_Adaptive_ElektronikAutomotive_201809_PressArticle_EN.pdf

[5] M. Smith. (Jul. 17, 2017). Elon musk's top cybersecurity concern: A fleet-wide hack of teslas. CSO. Accessed: Nov. 4, 2020. [Online]. Available: https://www.csoonline.com/article/3208035/elon-musk-s-top-cybersecurity-concern-preventing-a-fleet-wide-hack-of-teslas.html

[6] O. Burkacky, J. Deichmann, B. Klein, K. Pototzky, and G. Scherf. (2020). Cybersecurity in automotive: Mastering the challenge. McKinsey & Company and GSA. Accessed: Aug. 12, 2020. [Online]. Available: https://www.gsaglobal.org/wp-content/uploads/2020/03/Cybersecurity-in-automotive-Mastering-the-challenge.pdf

[7] AUTOSAR Foundation. (2003). *AUTOSAR—Automotive Open System Architecture*. Accessed: Aug. 13, 2019. [Online]. Available: https://www.autosar.org/

[8] AUTOSAR Foundation. (Nov. 28, 2019). *Specification of Identity and Access Management*. [Online]. Available: https://www.autosar.org/fileadmin/user_upload/standards/adaptive/19-11/AUTOSAR_SWS_IdentityAndAccessManagement.pdf

[9] J. Deichmann, B. Klein, G. Scherf, and R. Stützle. (Oct. 10, 2019). *The Race for Cybersecurity: Protecting the Connected Car in the Era of New Regulation: The Car Industry's Digital Transformation Exposes New Cybersecurity Threats Learn What OEMs Can Do to Protect Their Cars and Customers From Hackers*. Accessed: Aug. 12, 2020. [Online]. Available: https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/the-race-for-cybersecurity-protecting-the-connected-car-in-the-era-of-new-regulation

[10] D. Grimm, F. Pistorius, and E. Sax, "Network security monitoring in automotive domain," in *Advances in Information and Communication*, vol. 1129, K. Arai, S. Kapoor, and R. Bhatia, Eds. Cham, Switzerland: Springer, 2020, pp. 782–799. [Online]. Available: http://link.springer.com/10.1007/978-3-030-39445-5_57, doi: 10.1007/978-3-030-39445-5_57.

[11] M. Dehm and M. Tschersich, "Road vehicles' life-cycle: Mapping of relevant standards and regulations for automotive cybersecurity," in *Proc. 17th Escar Eur.*, Stuttgart, Germany, Nov. 2019. [Online]. Available: https://www.escar.info/images/Datastore/2019_escar_EU/lectures/16_Dehm.pdf

[12] ISO/SAE. (Sep. 20, 2018). *Road Vehicles—Cybersecurity Engineering*. [Online]. Available: https://www.iso.org/standard/70918.html

[13] J. Holle, R. Jung, S. Shukla, and T. Lothspeich, "About the future of layered security: How to achieve security in next generation EE-architectures," in *Proc. Escar Eur.*, Brussels, Belgium, Nov. 2018.

[14] T. Gehrmann and P. Duplys, "Intrusion detection for SOME/IP: Challenges and opportunities," in *Proc. Euromicro Conf. Digit. Syst. Design*, A. Trost, A. Žemva, and A. Skavhaug, Eds., Piscataway, NJ, USA, 2020, pp. 583–587, doi: 10.1109/DSD51259.2020.00096.

[15] O. Y. Al-Jarrah, C. Maple, M. Dianati, D. Oxtoby, and A. Mouzakitis, "Intrusion detection systems for intra-vehicle networks: A review," *IEEE Access*, vol. 7, pp. 21266–21289, 2019, doi: 10.1109/ACCESS.2019.2894183.

[16] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context aware computing for the Internet of Things: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 414–454, 1st Quart., 2014, doi: 10.1109/SURV.2013.042313.00197.

[17] X. Li, M. Eckert, J.-F. Martinez, and G. Rubio, "Context aware middleware architectures: Survey and challenges," *Sensors*, vol. 15, no. 8, pp. 20570–20607, 2015, doi: 10.3390/s150820570.

[18] W. Wu, R. Li, G. Xie, J. An, Y. Bai, J. Zhou, and K. Li, "A survey of intrusion detection for in-vehicle networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 919–933, Mar. 2020, doi: 10.1109/TITS.2019.2908074.

[19] L. B. Othmane, H. Weffers, M. M. Mohamad, and M. Wolf, "A survey of security and privacy in connected vehicles," in *Wireless Sensor and Mobile Ad-Hoc Networks*, D. Benhaddou and A. Al-Fuqaha, Eds. New York, NY, USA: Springer, 2015, pp. 217–247, doi: 10.1007/978-1-4939-2468-4_10.

[20] F. Luo and S. Hou, "Cyberattacks and countermeasures for intelligent and connected vehicles," *SAE Int. J. Passenger Cars-Electron. Electr. Syst.*, vol. 12, no. 1, pp. 55–66, Oct. 2019, doi: 10.4271/07-12-01-0005.

[21] M. Ring, D. Frkat, and M. Schmiedecker, "Cybersecurity evaluation of automotive E/E architectures," in *Proc. ACM Comput. Sci. Cars Symp. (CSCS)*, 2018, pp. 1–7.

[22] T. Rosenstatter and K. Tuma. (Oct. 16, 2019). A state-of-the-art investigation: CyReV deliverable D1.1. Cyber Resilience for Vehicles—CyReV, CyReV Consortium. Accessed: Aug. 25, 2020. [Online]. Available: https://autosec.se/wp-content/uploads/2019/11/CyReV_D1.1_A-State-of-the-Art-Investigation.pdf

[23] H. Vahdat-Nejad, A. Ramazani, T. Mohammadi, and W. Mansoor, "A survey on context-aware vehicular network applications," *Veh. Commun.*, vol. 3, pp. 43–57, Jan. 2016, doi: 10.1016/j.vehcom.2016.01.002.

[24] R. Fernandez-Rojas, A. Perry, H. Singh, B. Campbell, S. Elsayed, R. Hunjet, and H. A. Abbass, "Contextual awareness in human-advanced-vehicle systems: A survey," *IEEE Access*, vol. 7, pp. 33304–33328, 2019, doi: 10.1109/ACCESS.2019.2902812.

[25] E. Yurtsever, J. Lambert, A. Carballo, and K. Takeda, "A survey of autonomous driving: Common practices and emerging technologies," *IEEE Access*, vol. 8, pp. 58443–58469, 2020, doi: 10.1109/ACCESS.2020.2983149.

[26] B. Paden, M. Čáp, S. Z. Yong, D. Yershov, and E. Frazzoli, "A survey of motion planning and control techniques for self-driving urban vehicles," *IEEE Trans. Intell. Veh.*, vol. 1, no. 1, pp. 33–55, Mar. 2016, doi: 10.1109/TIV.2016.2578706.

[27] *IEEE Standard for Ethernet—Amendment 5: Physical Layer Specifications and Management Parameters for 10 Mb/s Operation and Associated Power Delivery Over a Single Balanced Pair of Conductors: 10BASE-T1S*, Standard IEEE 802.3cg-2019, 2019.

[28] *IEEE Standard for Ethernet Amendment 1: Physical Layer Specifications and Management Parameters for 100 Mb/s Operation Over a Single Balanced Twisted Pair Cable: 100BASE-T1*, Standard IEEE 802.3bw-2015, 2015.

[29] *ISO/IEC/IEEE 8802-3:2017/Amd 4-2017 Information Technology—Telecommunications and Information Exchange Between Systems—Local and Metropolitan Area Networks—Specific Requirements—Part 3: Standard for Ethernet Amendment 4: Physical Layer Specifications and Management Parameters for 1 Gb/s Operation Over a Single Twisted-Pair Copper Cable: 1000BASE-T1*, Standard IEEE 802.3bp-2016, 2016.

[30] *Audi A8 (Type 4N) Electrics and Electronics: Self Study Programme 664*, Audi Service Training I/VK-35, Audi AG, Ingolstadt, Germany, 2017.

[31] V. M. Navale, K. Williams, A. Lagospiris, M. Schaffert, and M.-A. Schweiker, "(R)evolution of E/E architectures," *SAE Int. J. Passenger Cars Electron. Elect. Syst.*, vol. 8, no. 2, pp. 282–288, 2015, doi: 10.4271/2015-01-0196.

[32] M. Maul, G. Becker, and U. Bernhard, "Service-oriented EE zone architecture key elements for new market segments," *ATZelektronik Worldwide*, vol. 13, no. 1, pp. 36–41, Feb. 2018, doi: 10.1007/s38314-017-0092-4.

[33] *SOME/IP Protocol Specification: Release 1.1.0*, document 696, AUTOSAR Foundation, 2017.

[34] Object Management Group. (Apr. 2015). *Data Distribution Service (DDS): Version 1.4*. Accessed: Sep. 13, 2019. [Online]. Available: http://www.omg.org/spec/DDS/1.4

[35] O. Burkacky, J. Deichmann, G. Doll, and C. Knochenhauer. (2018). Rethinking car software and electronics architecture. McKinsey & Company. Accessed: May 10, 2020. [Online]. Available: https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/rethinking-car-software-and-electronics-architecture

[36] Volkswagen AG. (Mar. 2019). *Automotive Cloud: Volkswagen and Microsoft Develop Mobility Ecosystem*. Accessed: Aug. 28, 2020. [Online]. Available: https://www.volkswagenag.com/en/news/stories/2019/03/automotive-cloud-volkswagen-and-microsoft-develop-mobility-ecosy.html

[37] Upstream Security. (2020). *Upstream Security Global Automotive Cybersecurity Report 2020: Research Into Cyber-Attack Trends in the Smart Mobility Ecosystem*. Accessed: Jan. 23, 2020. [Online]. Available: https://www.upstream.auto/upstream-security-global-automotive-cybersecurity-report-2020/

[38] European Union Agency for Cybersecurity. (2019). *ENISA Good Practices for the Security of Smart Cars*. Accessed: Aug. 23, 2020. [Online]. Available: https://op.europa.eu/publication/manifestation_identifier/PUB_TP0219881ENN

[39] D. Grimm, M. Weber, and E. Sax, "An extended hybrid anomaly detection system for automotive electronic control units communicating via Ethernet—Efficient and effective analysis using a specification- and machine learning-based approach," in *Proc. VEHITS*, M. Helfert and O. Gusikhin, Eds. Setúbal, Portugal: SCITEPRESS-Science and Technology Publications Lda, 2018, pp. 462–473, doi: 10.5220/0006779204620473.

[40] M. Rumez, D. Grimm, R. Kriesten, and E. Sax, "An overview of automotive service-oriented architectures and implications for security countermeasures," *IEEE Access*, vol. 8, pp. 221852–221870, 2020, doi: 10.1109/ACCESS.2020.3043070.

[41] J. Petit, B. Stottelaar, M. Feiri, and F. Kargl. (2015). Remote attacks on automated vehicles sensors: Experiments on camera and LiDAR. Blackhat Europe. [Online]. Available: https://www.blackhat.com/docs/eu-15/materials/eu-15-Petit-Self-Driving-And-Connected-Cars-Fooling-Sensors-And-Tracking-Drivers-wp1.pdf

[42] K. Eykholt, I. Evtimov, E. Fernandes, B. Li, A. Rahmati, C. Xiao, A. Prakash, T. Kohno, and D. Song, "Robust physical-world attacks on deep learning models," 2017, *arXiv:1707.08945*. [Online]. Available: http://arxiv.org/abs/1707.08945

[43] Y. Jia, Y. Lu, J. Shen, Q. A. Chen, Z. Zhong, and T. Wei, "Fooling detection alone is not enough: First adversarial attack against multiple object tracking," 2019, *arXiv:1905.11026*. [Online]. Available: http://arxiv.org/abs/1905.11026

[44] A. Tierney. (Apr. 24, 2020). From a TCU to corporate domain admin. Pen Test Partners. Accessed: Aug. 15, 2020. [Online]. Available: https://www.pentestpartners.com/security-blog/from-a-tcu-to-corporate-domain-admin/

[45] Tencent Keen Security Lab. (May 22, 2018). *New Vehicle Security Research By KeenLab: Experimental Security Assessment of BMW Cars*. Accessed: Aug. 15, 2020. [Online]. Available: https://keenlab.tencent.com/en/2018/05/22/New-CarHacking-Research-by-KeenLab-Experimental-Security-Assessment-of-BMW-Cars/

[46] Tencent Keen Security Lab. (Sep. 19, 2016). *Car Hacking Research: Remote Attack Tesla Motors*. Accessed: Aug. 15, 2020. [Online]. Available: https://keenlab.tencent.com/en/2016/09/19/Keen-Security-Lab-of-Tencent-Car-Hacking-Research-Remote-Attack-to-Tesla-Cars/

[47] Z. Whittaker. (Jul. 31, 2019). Security lapse exposed weak points on Honda's internal network. TechCrunch. Accessed: Aug. 15, 2020. [Online]. Available: https://techcrunch.com/2019/07/31/security-lapse-exposed-weak-points-on-hondas-internal-network/

[48] M. Weber, "Untersuchungen zur anomalieerkennung in automotive Steuergeräten durch verteilte observer mit fokus auf die plausibilisierung von kommunikationssignalen," Ph.D. dissertation, KIT, Karlsruhe, Germany, Jan. 2019, doi: 10.5445/IR/1000092815.

[49] M. Dibaei, X. Zheng, K. Jiang, S. Maric, R. Abbas, S. Liu, Y. Zhang, Y. Deng, S. Wen, J. Zhang, Y. Xiang, and S. Yu, "An overview of attacks and defences on intelligent connected vehicles," Jul. 2019, *arXiv:1907.07455*. [Online]. Available: http://arxiv.org/abs/1907.07455

[50] J. Petit and S. E. Shladover, "Potential cyberattacks on automated vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 546–556, Apr. 2014, doi: 10.1109/TITS.2014.2342271.

[51] *Specification of Secure Onboard Communication*, document 654, AUTOSAR Foundation, 2017.

[52] A. Barisani, "Insecure boot," in *Proc. 17th Escar Eur.*, Stuttgart, Germany, Nov. 2019, pp. 128–131. [Online]. Available: https://www.escar.info/images/Datastore/2019_escar_EU/lectures/18_Barisani.pdf

[53] D. R. Miller, *Security Information and Event Management (SIEM) Implementation* (Network Pro Library). New York, NY, USA: McGraw-Hill, 2010.

[54] L. Kaneti and J. Sultanik, "Automotive SOC: From concept to realization," in *Proc. 6th Escar USA*, Jun. 2018.

[55] C. Olt. (2018). Cybersecurity for connected cars: Pathways to a security operation center for the automotive industry. T-Systems International GmbH. Accessed: May 19, 2021. [Online]. Available: https://www.t-systems.com/de/en/industries/automotive/automotive-security

[56] F. Langer, F. Schüppel, and L. Stahlbock, "Incident response for vehicular systems: More than online updates," in *Proc. 18th Escar Eur., World's Lead. Automot. Cyber Secur. Conf.* Bochum, Germany: Ruhr-Univ. Bochum, 2020, pp. 117–122.

[57] H. Guissouma, C. P. Hohl, H. Stoll, and E. Sax, "Variability-aware process extension for updating cyber physical systems over the air," in *Proc. 9th Medit. Conf. Embedded Comput. (MECO)*, Budva, Montenegro, Jun. 2020, p. 1, doi: 10.1109/MECO49872.2020.9134339.

[58] A. K. Dey and G. D. Abowd, "Towards a better understanding of context and context-awareness," in *Proc. 1st Int. Symp. Handheld Ubiquitous Comput. (HUC)*, 1999, pp. 304–307.

[59] V. Jovanovikj, D. Gabrijelčič, and T. Klobučar, "A conceptual model of security context," *Int. J. Inf. Secur.*, vol. 13, no. 6, pp. 571–581, Nov. 2014, doi: 10.1007/s10207-014-0229-x.

[60] Gartner Information Technology Glossary. (Apr. 11, 2020). *Definition of Context-Aware Security*. [Online]. Available: https://www.gartner.com/en/information-technology/glossary/context-aware-security

[61] R. McMillan. (May 16, 2013). Definition: Threat intelligence: ID: G00249251. Gartner Research. Accessed: Nov. 4, 2020. [Online]. Available: https://www.gartner.com/en/documents/2487216/definition-threat-intelligence

[62] Auto-ISAC. (2015). *Automotive Information Sharing & Analysis Center*. Accessed: Aug. 23, 2020. [Online]. Available: https://automotiveisac.com/

[63] Upstream Security. (Aug. 23, 2020). *AutoThreat Intelligence Cyber Incident Repository*. [Online]. Available: https://www.upstream.auto/research/automotive-cybersecurity/

[64] G. Harpak and Y. Chen, "Mercedes-Benz and 360 group: Defending a luxury fleet with the community," in *Proc. RSA Conf.*, San Francisco, CA, USA, Feb. 2020. [Online]. Available: https://www.rsaconference.com/usa/us-2020/agenda/mercedes-benz-and-360-group-defending-a-luxury-fleet-with-the-community

[65] S. Sinha, M. Bailey, and F. Jahanian, "One size does not fit all: 10 years of applying context-aware security," in *Proc. IEEE Conf. Technol. Homeland Secur.*, May 2009, pp. 14–21, doi: 10.1109/THS.2009.5168009.

[66] K. Henricksen, "A framework for context-aware pervasive computing applications," Ph.D. dissertation, Univ. Queensland, Brisbane, QLD, Australia, Jan. 2003. [Online]. Available: https://espace.library.uq.edu.au/view/UQ:106832

[67] P. Hu, J. Indulska, and R. Robinson, "An autonomic context management system for pervasive computing," in *Proc. 6th Annu. IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*, Los Alamitos, CA, USA: IEEE Computer Society, Mar. 2008, pp. 213–223, doi: 10.1109/PERCOM.2008.56.

[68] M. Baldauf, S. Dustdar, and F. Rosenberg, "A survey on context-aware systems," *Int. J. Ad Hoc Ubiquitous Comput.*, vol. 2, no. 4, p. 263, 2007, doi: 10.1504/IJAHUC.2007.014070.

[69] C. Bettini, O. Brdiczka, K. Henricksen, J. Indulska, D. Nicklas, A. Ranganathan, and D. Riboni, "A survey of context modelling and reasoning techniques," *Pervasive Mobile Comput.*, vol. 6, no. 2, pp. 161–180, 2010, doi: 10.1016/j.pmcj.2009.06.002.

[70] T. Strang and C. Linnhoff-Popien, "A context modeling survey," in *Proc. 1st Int. Workshop Adv. Context Modelling, Reasoning Manage., 6th Int. Conf. Ubiquitous Comput.*, Nottingham, U.K., Sep. 2004.

[71] A. Khattak, N. Akbar, M. Aazam, T. Ali, A. Khan, S. Jeon, M. Hwang, and S. Lee, "Context representation and fusion: Advancements and opportunities," *Sensors*, vol. 14, no. 6, pp. 9628–9668, May 2014, doi: 10.3390/s140609628.

[72] M. Hoda, V. Montaghami, H. Al Osman, and A. El Saddik, "ECOPPA: Extensible context ontology for persuasive physical-activity applications," in *Proc. Int. Conf. Inf. Technol. Syst. (ICITS)*, in Advances in Intelligent Systems and Computing, vol. 721, Á. Rocha and T. Guarda, Eds. Cham, Switzerland: Springer, 2018, pp. 309–318, doi: 10.1007/978-3-319-73450-7_30.

[73] AUTOSAR Foundation. (Nov. 28, 2019). *Specification of Persistency*. [Online]. Available: https://www.autosar.org/fileadmin/user_upload/standards/adaptive/19-11/AUTOSAR_SWS_Persistency.pdf

[74] M. Knappmeyer, S. L. Kiani, C. Fra, B. Moltchanov, and N. Baker, "ContextML: A light-weight context representation and context management schema," in *Proc. 5th IEEE Int. Symp. Wireless Pervas. Comput. (ISWPC)*, M. L. Merani, Ed., Piscataway, NJ, USA, May 2010, pp. 367–372, doi: 10.1109/ISWPC.2010.5483753.

[75] S. Auer, C. Bizer, G. Kobilarov, J. Lehmann, R. Cyganiak, and Z. Ives, "DBpedia: A nucleus for a web of open data," in *The Semantic Web* (Lecture Notes in Computer Science), vol. 4825, K. Aberer, Ed. Berlin, Germany: Springer, 2007, pp. 722–735, doi: 10.1007/978-3-540-76298-0_52.

[76] L. Sanchez, J. Lanza, R. Olsen, M. Bauer, and M. Girod-Genet, "A generic context management framework for personal networking environments," in *Proc. 3rd Annu. Int. Conf. Mobile Ubiquitous Syst., Netw. Services*, Piscataway, NJ, USA, Jul. 2006, pp. 1–8, doi: 10.1109/MOBIQW.2006.361743.

[77] K. Wan and V. Alagar, "Context-aware security solutions for cyber-physical systems," *Mobile Netw. Appl.*, vol. 19, no. 2, pp. 212–226, 2014, doi: 10.1007/s11036-014-0495-x.

[78] M. Feld and C. Müller, "The automotive ontology: Managing knowledge inside the vehicle and sharing it between cars," in *Proc. 3rd Int. Conf. Automot. User Interfaces Interact. Veh. Appl.*, M. Tscheligi, Ed., New York, NY, USA, 2011, p. 79, doi: 10.1145/2381416.2381429.

[79] H. Lee, S. H. Jeong, and H. K. Kim. (2017). CAN-intrusion-dataset (OTIDS). Hacking and Countermeasure Research Lab. [Online]. Available: https://ocslab.hksecurity.net/Dataset/CAN-intrusion-dataset

[80] H. M. Song, J. Woo, and H. K. Kim. (2018). Car-hacking dataset. Hacking and Countermeasure Research Lab. [Online]. Available: https://ocslab.hksecurity.net/Datasets/CAN-intrusion-dataset

[81] The European Parliament and the Council of the European Union. (May 4, 2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons With Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC: General Data Protection Regulation*. [Online]. Available: http://data.europa.eu/eli/reg/2016/679/oj

[82] G. K. Rajbahadur, A. J. Malton, A. Walenstein, and A. E. Hassan, "A survey of anomaly detection for connected vehicle cybersecurity and safety," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Piscataway, NJ, USA, Jun. 2018, pp. 421–426, doi: 10.1109/IVS.2018.8500383.

[83] S. Sharma and A. Kaul, "A survey on intrusion detection systems and honeypot based proactive security mechanisms in VANETs and VANET cloud," *Veh. Commun.*, vol. 12, pp. 138–164, Apr. 2018, doi: 10.1016/j.vehcom.2018.04.005.

[84] F. A. Ghaleb, M. A. Maarof, A. Zainal, B. A. S. Al-Rimy, F. Saeed, and T. Al-Hadhrami, "Hybrid and multifaceted context-aware misbehavior detection model for vehicular ad hoc network," *IEEE Access*, vol. 7, pp. 159119–159140, 2019, doi: 10.1109/ACCESS.2019.2950805.

[85] Argus Cyber Security. (Dec. 11, 2020). *Argus Fleet Protection*. [Online]. Available: https://argus-sec.com/argus-fleet-protection/

[86] Arilou. (Dec. 11, 2020). *SIEM/SOC Backend Solution*. [Online]. Available: https://ariloutech.com/solutions/siem-soc-backend-solution/

[87] Harman Automotive. (Dec. 11, 2020). *Cybersecurity for Automotive | Harman Shield*. [Online]. Available: https://car.harman.com/solutions/cybersecurity/cybersecurity-automotive-harman-shield

[88] Upstream Security. (2020). *Setting the Standard for Automotive Cybersecurity: White Paper—Understanding the ISO/SAE 21434 Draft Standard and Upstream Security's Solutions for Effective Compliance*. Accessed: Nov. 12, 2020. [Online]. Available: https://upstream.auto/lp-setting-the-standard-for-automotive-cybersecurity/

[89] D. Sahar. (Oct. 2, 2018). Upstream's contextually aware security architecture. Upstream Security. Accessed: Nov. 12, 2020. [Online]. Available: https://upstream.auto/resources/upstreams-contextually-aware-security-architecture/

[90] T. Haga, R. Takahashi, T. Sasaki, T. Kishikawa, J. Tsurumi, and H. Matsushima, "Automotive SIEM and anomaly detection using sand-sprinkled isolation forest," in *Proc. Escar Eur.*, 2017, pp. 1–10.

[91] O. Berlin, A. Held, M. Matousek, and F. Kargl, "POSTER: Anomaly-based misbehaviour detection in connected car backends," in *Proc. IEEE Veh. Netw. Conf. (VNC)*, Dec. 2016, pp. 1–2, doi: 10.1109/VNC.2016.7835978.

[92] M. Levi, Y. Allouche, and A. Kontorovich, "Advanced analytics for connected cars cyber security," Nov. 2017, *arXiv:1711.01939*. [Online]. Available: http://arxiv.org/abs/1711.01939

[93] S. Iqbal, A. Haque, and M. Zulkernine, "Towards a security architecture for protecting connected vehicles from malware," in *Proc. VTC Spring*. Piscataway, NJ, USA: Institute of Electrical and Electronics Engineers, 2019, pp. 1–5, doi: 10.1109/VTCSpring.2019.8746516.

[94] M. R. Moore, R. A. Bridges, F. L. Combs, and A. L. Anderson, "Data-driven extraction of vehicle states from CAN bus traffic for cyber-protection and safety," *IEEE Consum. Electron. Mag.*, vol. 8, no. 6, pp. 104–110, Nov. 2019, doi: 10.1109/MCE.2019.2928577.

[95] S. N. Narayanan, S. Mittal, and A. Joshi, "Using semantic technologies to mine vehicular context for security," in *Proc. IEEE 37th Sarnoff Symp.*, Sep. 2016, pp. 124–129, doi: 10.1109/SARNOF.2016.7846740.

[96] M. Bermudez-Edo, T. Elsaleh, P. Barnaghi, and K. Taylor, "IoT-lite: A lightweight semantic model for the Internet of Things," in *Proc. UIC-ATC-ScalCom-CBDCom-IoP-SmartWorld*, D. E. Baz and J. Bourgeois, Eds., Piscataway, NJ, USA, 2016, pp. 90–97, doi: 10.1109/UIC-ATC-ScalCom-CBDCom-IoP-SmartWorld.2016.0035.

[97] S. N. Narayanan, "A framework for detecting anomalous behaviors in smart cyber-physical systems," Ph.D. dissertation, Univ. Maryland, College Park, MD, USA, 2019.

[98] A. R. Wasicek, M. D. Pesé, A. Weimerskirch, Y. Burakova, and K. Singh, "Context-aware intrusion detection in automotive control systems," in *Proc. 5th ESCAR USA*, Ypsilanti, MI, USA, Jun. 2017, pp. 1–14.

[99] J. Jiang, C. Wang, S. Chattopadhyay, and W. Zhang, "Road context-aware intrusion detection system for autonomous cars," Aug. 2019, *arXiv:1908.00732*. [Online]. Available: http://arxiv.org/abs/1908.00732, doi: 10.1007/978-3-030-41579-2_8.

[100] H. K. Kalutarage, M. O. Al-Kadri, M. Cheah, and G. Madzudzo, "Context-aware anomaly detector for monitoring cyber attacks on automotive CAN bus," in *Proc. ACM Comput. Sci. Cars Symp. (CSCS)*, H.-J. Hof, M. Fritz, C. Krauß, and O. Wasenmüller, Eds., New York, NY, USA, 2019, pp. 1–8, doi: 10.1145/3359999.3360496.

[101] M. Rumez, A. Duda, P. Grunder, R. Kriesten, and E. Sax, "Integration of attribute-based access control into automotive architectures," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2019, pp. 1916–1922, doi: 10.1109/IVS.2019.8814265.

[102] V. Hugot, A. Jousse, C. Toinard, and B. Venelle, "oMAC: Open model for automotive cybersecurity," in *Proc. 17th Escar Eur.* Bochum, Germany: Ruhr-Univ. Bochum, 2019, pp. 170–184, doi: 10.13154/294-6674.

[103] S. Gansel, S. Schnitzer, A. Gilbeau-Hammoud, V. Friesen, F. Dürr, K. Rothermel, C. Maihöfer, and U. Krämer, "Context-aware access control in novel automotive HMI systems," in *Information Systems Security* (Lecture Notes in Computer Science), vol. 9478, S. Jajoda and C. Mazumdar, Eds. Cham, Switzerland: Springer, 2015, pp. 118–138, doi: 10.1007/978-3-319-26961-0_8.

[104] A. Armand, D. Filliat, and J. Ibanez-Guzman, "Ontology-based context awareness for driving assistance systems," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2014, pp. 227–233, doi: 10.1109/IVS.2014.6856509.

[105] S. Ulbrich, T. Nothdurft, M. Maurer, and P. Hecker, "Graph-based context representation, environment modeling and information aggregation for automated driving," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2014, pp. 541–547, doi: 10.1109/IVS.2014.6856556.

[106] S. Al-Sultan, A. H. Al-Bayatti, and H. Zedan, "Context-aware driver behavior detection system in intelligent transportation systems," *IEEE Trans. Veh. Technol.*, vol. 62, no. 9, pp. 4264–4275, Nov. 2013, doi: 10.1109/TVT.2013.2263400.

[107] A. Groza, A. Marginean, and V. Muresan, "An ontology-based model for vehicular ad-hoc networks," in *Proc. IEEE 18th Int. Conf. Intell. Eng. Syst. INES*, Jul. 2014, pp. 83–88, doi: 10.1109/INES.2014.6909346.

[108] D. Zhang, X. H. Wang, and K. Hackbarth, "OSGi based service infrastructure for context aware automotive telematics," in *Proc. IEEE 59th Veh. Technol. Conf. (VTC-Spring)*, May 2004, pp. 2957–2961, doi: 10.1109/VETECS.2004.1391466.

[109] X. H. Wang, D. Q. Zhang, T. Gu, and H. K. Pung, "Ontology based context modeling and reasoning using OWL," in *Proc. 2nd IEEE Annu. Conf. Pervasive Comput. Commun. Workshops*, Los Alamitos, CA, USA: IEEE Computer Society, Mar. 2004, pp. 18–22, doi: 10.1109/PERCOMW.2004.1276898.

[110] P. Obergfell, C. Segler, E. Sax, and A. Knoll, "Synchronization between run-time and design-time view of context-aware automotive system architectures," in *Proc. 4th IEEE Int. Symp. Syst. Eng.*, Piscataway, NJ, USA, Oct. 2018, pp. 1–3, doi: 10.1109/SysEng.2018.8544454.

[111] C. Segler, S. Kugele, P. Obergfell, M. H. Osman, S. Shafaei, E. Sax, and A. Knoll, "Evaluation of feature selection for anomaly detection in automotive E/E architectures," in *Proc. IEEE/ACM 41st Int. Conf. Softw. Eng., Companion (ICSE-Companion)*, Piscataway, NJ, USA, May 2019, pp. 260–261, doi: 10.1109/ICSE-Companion.2019.00104.

[112] C. Segler, S. Kugele, and A. Knoll, "Context discovery for personalised automotive functions," in *Proc. IEEE Intell. Transp. Syst. Conf. (ITSC)*, Piscataway, NJ, USA, Oct. 2019, pp. 2470–2476, doi: 10.1109/ITSC.2019.8917161.

[113] C. Segler, S. Kugele, P. Obergfell, M. H. Osman, S. Shafaci, E. Sax, and A. Knoll, "Anomaly detection for advanced driver assistance systems using online feature selection," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2019, pp. 578–585, doi: 10.1109/IVS.2019.8814175.

[114] PEGASUS Project, "Scenario description: Requirements & conditions," in *Proc. Pegasus Symp.*, Aachen, Germany, Nov. 2017. [Online]. Available: https://www.pegasusprojekt.de/files/tmpl/PDF-Symposium/04_Scenario-Description.pdf

[115] G. Bagschik, T. Menzel, and M. Maurer, "Ontology based scene creation for the development of automated vehicles," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2018, pp. 1813–1820, doi: 10.1109/IVS.2018.8500632.

[116] GENIVI. (Oct. 11, 2020). *Vehicle Signal Specification—Standardized Way to Describe Automotive Data*. [Online]. Available: https://genivi.github.io/vehicle_signal_specification/

[117] B. Klotz, R. Troncy, D. Wilms, and C. Bonnet, "Generating semantic trajectories using a car signal ontology," in *Proc. Companion Web Conf. (WWW)*, P.-A. Champin, F. Gandon, M. Lalmas, and P. G. Ipeirotis, Eds., New York, NY, USA, 2018, pp. 135–138, doi: 10.1145/3184558.3186962.

[118] B. Klotz, R. Troncy, D. Wilms, and C. Bonnet, "VSSo—A vehicle signal and attribute ontology," in *Proc. 9th Int. Semantic Sensor Netw. Workshop (SSN)*, 2018, pp. 1–8. [Online]. Available: https://ssn2018.github.io/submissions/SSN2018_paper_4_submitted.pdf

[119] T. P. Nogueira, R. B. Braga, C. T. de Oliveira, and H. Martin, "FrameSTEP: A framework for annotating semantic trajectories based on episodes," *Expert Syst. Appl.*, vol. 92, pp. 533–545, Feb. 2018, doi: 10.1016/j.eswa.2017.10.004.

[120] F. Terroso-Saenz, M. Valdes-Vela, and A. F. Skarmeta-Gomez, "Design of an event-based architecture for the intra-vehicular context perception," in *Proc. 17th Int. Conf. Inf. Fusion (FUSION)*, 2014, pp. 1–7.

[121] F. Terroso-Sáenz, M. Valdés-Vela, F. Campuzano, J. A. Botia, and A. F. Skarmeta-Gómez, "A complex event processing approach to perceive the vehicular context," *Inf. Fusion*, vol. 21, pp. 187–209, Jan. 2015, doi: 10.1016/j.inffus.2012.08.008.

[122] R. Copeland, "Automotive context-aware policy system for car connectivity requests," in *Proc. 18th Int. Conf. Intell. Next Gener. Netw. (ICIN)*, Piscataway, NJ, USA, 2015, pp. 128–135, doi: 10.1109/ICIN.2015.7073818.

[123] L. Ries, M. Stumpf, J. Bach, and E. Sax, "Semantic comparison of driving sequences by adaptation of word embeddings," in *Proc. IEEE 23rd Int. Conf. Intell. Transp. Syst. (ITSC)*, Sep. 2020, pp. 1–7, doi: 10.1109/ITSC45102.2020.9294364.

[124] T. Taniguchi, S. Nagasaka, K. Hitomi, N. P. Chandrasiri, and T. Bando, "Semiotic prediction of driving behavior using unsupervised double articulation analyzer," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2012, pp. 849–854, doi: 10.1109/IVS.2012.6232243.

[125] T. Bando, K. Takenaka, S. Nagasaka, and T. Taniguchi, "Unsupervised drive topic finding from driving behavioral data," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2013, pp. 177–182, doi: 10.1109/IVS.2013.6629467.

[126] W. Meng, W. Li, and L.-F. Kwok, "EFM: Enhancing the performance of signature-based network intrusion detection systems using enhanced filter mechanism," *Comput. Secur.*, vol. 43, pp. 189–204, Jun. 2014, doi: 10.1016/j.cose.2014.02.006.

[127] E. Hayashi, S. Das, S. Amini, J. Hong, and I. Oakley, "CASA: Context-aware scalable authentication," in *Proc. 9th Symp. Usable Privacy Secur.*, L. F. Cranor, Ed., New York, NY, USA, 2013, p. 1, doi: 10.1145/2501604.2501607.

[128] P. Bhandari and M. Singh, "OntoSecure: A semantic web based tool for network security status prediction," in *Proc. IEEE 6th Int. Conf. Adv. Comput. (IACC)*, Feb. 2016, pp. 551–555, doi: 10.1109/IACC.2016.108.

[129] M. S. Iqbal and M. Zulkernine, "Flamingo: A framework for smartphone security context management," in *Proc. Symp. Appl. Comput. (SAC)*, S. Y. Shin, D. Shin, and M. Lencastre, Eds., New York, NY, USA, 2017, pp. 563–568, doi: 10.1145/3019612.3019726.

[130] M. S. Iqbal and M. Zulkernine, "Droid mood swing (DMS): Automatic security modes based on contexts," in *Information Security* (Lecture Notes in Computer Science), vol. 10599, P. Q. Nguyen and J. Zhou, Eds. Cham, Switzerland: Springer, 2017, pp. 329–347, doi: 10.1007/978-3-319-69659-1_18.

[131] H. Dreger, C. Kreibich, V. Paxson, and R. Sommer, "Enhancing the accuracy of network-based intrusion detection with host-based context," in *Proc. 2nd Int. Conf. Detection Intrusions Malware, Vulnerability Assessment (DIMVA)*, in Lecture Notes in Computer Science, Vienna, Austria, vol. 3548, K. Julisch and C. Kruegel, Eds. Berlin, Germany: Springer, Jul. 2005, pp. 206–221, doi: 10.1007/11506881_13.

[132] M. Miettinen, S. Heuser, W. Kronz, A.-R. Sadeghi, and N. Asokan, "ConXsense—Automated context classification for context-aware access control," in *Proc. 9th ACM Symp. Inf., Comput. Commun. Secur.*, S. Moriai, T. Jaeger, and K. Sakurai, Eds. New York, NY, USA: Association for Computing Machinery, 2014, pp. 293–304, doi: 10.1145/2590296.2590337.

[133] S. More, M. Matthews, A. Joshi, and T. Finin, "A knowledge-based approach to intrusion detection modeling," in *Proc. IEEE Symp. Secur. Privacy Workshops (SPW)*, Piscataway, NJ, USA, May 2012, pp. 75–81, doi: 10.1109/SPW.2012.26.

[134] Recorded Future. (Nov. 3, 2020). *The Security Intelligence Graph: White Paper*. [Online]. Available: https://go.recordedfuture.com/security-intelligence-graph

[135] M. Iannacone, S. Bohn, G. Nakamura, J. Gerth, K. Huffer, R. Bridges, E. Ferragut, and J. Goodall, "Developing an ontology for cyber security knowledge graphs," in *Proc. 10th Annu. Cyber Inf. Secur. Res. Conf. (CISR)*, J. P. Trien, S. J. Prowell, R. A. Bridges, and J. R. Goodall, Eds., New York, NY, USA, 2015, pp. 1–4, doi: 10.1145/2746266.2746278.

[136] A. Sadighian, J. M. Fernandez, A. Lemay, and S. T. Zargar, "ONTIDS: A highly flexible context-aware and ontology-based alert correlation framework," in *Foundations and Practice of Security* (Lecture Notes in Computer Science), vol. 8352, J. L. Danger, M. Debbabi, J.-Y. Marion, and N. Z. Heywood, Eds. Cham, Switzerland: Springer, 2014, pp. 161–177, doi: 10.1007/978-3-319-05302-8_10.

[137] A. Sadighian, S. T. Zargar, J. M. Fernandez, and A. Lemay, "Semantic-based context-aware alert fusion for distributed intrusion detection systems," in *Proc. Int. Conf. Risks Secur. Internet Syst. (CRiSIS)*, Oct. 2013, pp. 1–6, doi: 10.1109/CRiSIS.2013.6766352.

[138] S. Saad and I. Traore, "Semantic aware attack scenarios reconstruction," *J. Inf. Secur. Appl.*, vol. 18, no. 1, pp. 53–67, Jul. 2013, doi: 10.1016/j.jisa.2013.08.002.

[139] S. Saad and I. Traore, "Method ontology for intelligent network forensics analysis," in *Proc. 8th Int. Conf. Privacy, Secur. Trust (PST)*, Piscataway, NJ, USA, Aug. 2010, pp. 7–14, doi: 10.1109/PST.2010.5593235.

[140] Y. Jia, Y. Qi, H. Shang, R. Jiang, and A. Li, "A practical approach to constructing a knowledge graph for cybersecurity," *Engineering*, vol. 4, no. 1, pp. 53–60, 2018, doi: 10.1016/j.eng.2018.01.004.

[141] Z. Syed, A. Padia, M. L. Mathews, T. Finin, and A. Joshi, "UCO: A unified cybersecurity ontology," in *Proc. AAAI Workshop Artif. Intell. Cyber Secur.* AAAI Press, 2015, pp. 14–21.

[142] S. Barnum, "Standardizing cyber threat intelligence information with the structured threat information expression (STIX), version 1.1, revision 1," Mitre Corp., McLean, VA, USA, White Paper, Feb. 2014, pp. 1–22, vol. 11. [Online]. Available: http://www.standardscoordination.org/sites/default/files/docs/STIX_Whitepaper_v1.1.pdf

[143] F. M. Suchanek, G. Kasneci, and G. Weikum, "YAGO: A large ontology from Wikipedia and WordNet," *J. Web Semantics*, vol. 6, no. 3, pp. 203–217, 2008, doi: 10.1016/j.websem.2008.06.001.

[144] A. Pingle, A. Piplai, S. Mittal, A. Joshi, J. Holt, and R. Zak, "RelExt: Relation extraction using deep learning approaches for cybersecurity knowledge graph improvement," May 2019, *arXiv:1905.02497*. [Online]. Available: http://arxiv.org/abs/1905.02497

[145] E. Kiesling, A. Ekelhart, K. Kurniawan, and F. Ekaputra, "The SEPSES knowledge graph: An integrated resource for cybersecurity," in *The Semantic Web—ISWC*, C. Ghidini, O. Hartig, M. Maleshkova, V. Svátek, I. Cruz, A. Hogan, J. Song, M. Lefrançois, and F. Gandon, Eds. Cham, Switzerland: Springer, 2019, pp. 198–214.

[146] P. Najafi, A. Mühle, W. Pünter, F. Cheng, and C. Meinel, ''MalRank: A measure of maliciousness in SIEM-based knowledge graphs,'' in *Proc. 35th Annu. Comput. Secur. Appl. Conf. (ACSAC)*. New York, NY, USA: Association for Computing Machinery, 2019, pp. 417–429, doi: 10.1145/3359789.3359791.

[147] T. Yu, V. Sekar, S. Seshan, Y. Agarwal, and C. Xu, ''Handling a trillion (unfixable) flaws on a billion devices: Rethinking network security for the Internet-of-Things,'' in *Proc. 14th ACM Workshop Hot Topics Netw. (HotNets)*, J. D. Oliveira, J. Smith, K. Argyraki, and P. Levis, Eds., New York, NY, USA, 2015, pp. 1–7, doi: 10.1145/2834050.2834095.

[148] B. A. Mozzaquatro, R. Melo, C. Agostinho, and R. Jardim-Goncalves, ''An ontology-based security framework for decision-making in industrial systems,'' in *Proc. MODELSWARD*, S. Hammoudi, Ed. Setúbal, Portugal: SCITEPRESS-Science and Technology Publications Lda, 2016, pp. 779–788, doi: 10.5220/0005853107790788.

[149] B. A. Mozzaquatro, R. Jardim-Goncalves, and C. Agostinho, ''Towards a reference ontology for security in the Internet of Things,'' in *Proc. IEEE Int. Workshop Meas. Netw. (M&N)*, Piscataway, NJ, USA, 2015, pp. 1–6, doi: 10.1109/IWMN.2015.7322984.

[150] G. Xu, Y. Cao, Y. Ren, X. Li, and Z. Feng, ''Network security situation awareness based on semantic ontology and user-defined rules for Internet of Things,'' *IEEE Access*, vol. 5, pp. 21046–21056, 2017, doi: 10.1109/ACCESS.2017.2734681.

[151] S. D. Anton, D. Fraunholz, C. Lipps, K. Alam, and H. D. Schotten, ''Putting things in context: Securing industrial authentication with context information,'' *Int. J. Cyber Situational Awareness*, vol. 4, no. 1, pp. 98–120, Dec. 2018, doi: 10.22619/IJCSA.2018.100122.

[152] S. D. Anton, D. Fraunholz, J. Zemitis, F. Pohl, and H. D. Schotten, ''Highly scalable and flexible model for effective aggregation of context-based data in generic IIoT scenarios,'' in *Proc. 9th Central Eur. Workshop Services Their Composition (ZEUS)*, O. Kopp, J. Lenhard, and C. Pautass, Eds., 2017, pp. 51–58.

[153] A.-P. Lohfink, S. D. D. Anton, H. D. Schotten, H. Leitte, and C. Garth, ''Security in process: Visually supported triage analysis in industrial process data,'' *IEEE Trans. Vis. Comput. Graphics*, vol. 26, no. 4, pp. 1638–1649, Apr. 2020, doi: 10.1109/TVCG.2020.2969007.

[154] J. L. H. Ramos, J. B. Bernabe, and A. F. Skarmeta, ''Managing context information for adaptive security in IoT environments,'' in *Proc. IEEE 29th Int. Conf. Adv. Inf. Netw. Appl. Workshops (WAINA)*, Piscataway, NJ, USA, Mar. 2015, pp. 676–681, doi: 10.1109/WAINA.2015.55.

[155] A. Lekidis and I. Barosan, ''Model-based simulation and threat analysis of in-vehicle networks,'' in *Proc. 15th IEEE Int. Workshop Factory Commun. Syst. (WFCS)*, Piscataway, NJ, USA, May 2019, pp. 1–8, doi: 10.1109/WFCS.2019.8757968.

[156] K. Tong, Z. Ajanovic, and G. Stettinger, ''Overview of tools supporting planning for automated driving,'' Mar. 2020, *arXiv:2003.04081*. [Online]. Available: http://arxiv.org/abs/2003.04081

[157] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, ''CARLA: An open urban driving simulator,'' 2017, *arXiv:1711.03938*. [Online]. Available: http://arxiv.org/abs/1711.03938

[158] T. Toyama, T. Yoshida, H. Oguma, and T. Matsumoto, ''PASTA: Portable automotive security testbed with adaptability,'' in *Proc. BlackHat Asia*, Singapore, Mar. 2019.

[159] D. Zelle, R. Rieke, C. Plappert, C. Kraus, D. Levshun, and A. Chechulin, ''SEPAD—Security evaluation platform for autonomous driving,'' in *Proc. 28th Euromicro Int. Conf. Parallel, Distrib. Netw.-Based Process. (PDP)*, Mar. 2020, pp. 413–420, doi: 10.1109/PDP50117.2020.00070.

[160] Offensive Security. (Jun. 17, 2020). *Kali Linux*. [Online]. Available: https://www.kali.org/

[161] Rapid7. (Jun. 17, 2020). *Metasploit: Penetration Testing Software, Pen Testing Security*. [Online]. Available: https://www.metasploit.com/

[162] M. Bertoncello and G. Camplone. (Sep. 2016). Monetizing car data: New service business opportunities to create new customer benefits. McKinsey & Company. Accessed: Aug. 25, 2020. [Online]. Available: https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/monetizing-car-data

[163] X. Gao, B. Firner, S. Sugrim, V. Kaiser-Pendergrast, Y. Yang, and J. Lindqvist, ''Elastic pathing: Your speed is enough to track you,'' in *Proc. ACM Int. Joint Conf. Pervas. Ubiquitous Comput.*, A. J. Brush, Ed., New York, NY, USA, 2014, pp. 975–986, doi: 10.1145/2632048.2632077.

[164] M. U. Hassan, M. H. Rehmani, and J. Chen, ''Differential privacy techniques for cyber physical systems: A survey,'' *IEEE Commun. Surveys Tuts.*, vol. 22, no. 1, pp. 746–789, 1st Quart., 2020, doi: 10.1109/COMST.2019.2944748.

[165] L. Sweeney, ''K-anonymity: A model for protecting privacy,'' *Int. J. Uncertainty, Fuzziness Knowl.-Based Syst.*, vol. 10, no. 5, pp. 557–570, 2002, doi: 10.1142/S0218488502001648.

[166] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, ''L-diversity: Privacy beyond k-anonymity,'' *ACM Trans. Knowl. Discovery From Data*, vol. 1, no. 1, p. 3, 2007, doi: 10.1145/1217299.1217302.

[167] N. Li, T. Li, and S. Venkatasubramanian, ''T-closeness: Privacy beyond k-anonymity and l-diversity,'' in *Proc. IEEE 23rd Int. Conf. Data Eng.*, Piscataway, NJ, USA, Apr. 2007, pp. 106–115, doi: 10.1109/ICDE.2007.367856.

[168] B. Nelson and T. Olovsson, ''Introducing differential privacy to the automotive domain: Opportunities and challenges,'' in *Proc. IEEE 86th Veh. Technol. Conf. (VTC-Fall)*, Piscataway, NJ, USA, Sep. 2017, pp. 1–7, doi: 10.1109/VTCFall.2017.8288389.

[169] E. G. Kalaycı, I. G. González, F. Lösch, G. Xiao, A. U. Mehdi, E. Kharlamov, and D. Calvanese, ''Semantic integration of Bosch manufacturing data using virtual knowledge graphs,'' in *The Semantic Web—ISWC* (Lecture Notes in Computer Science), vol. 12507, J. Z. Pan, Ed. Cham, Switzerland: Springer, 2020, pp. 464–481, doi: 10.1007/978-3-030-62466-8_29.

[170] K. Allix, T. F. Bissyandé, J. Klein, and Y. Le Traon, ''AndroZoo: Collecting millions of Android apps for the research community,'' in *Proc. 13th Work. Conf. Mining Softw. Repositories (MSR)*, M. Kim, R. Robbes, and C. Bird, Eds., Piscataway, NJ, USA, 2016, pp. 468–471, doi: 10.1145/2901739.2903508.

DANIEL GRIMM received the B.Sc. and M.Sc. degrees in electrical engineering and information technology from the Karlsruhe Institute of Technology (KIT), Germany, in 2015 and 2017, respectively, where he is currently pursuing the Ph.D. degree with the Group of Systems Engineering, Institute for Information Processing Technologies (ITIV). Since 2018, he has been a Research Assistant with the Group of Systems Engineering, ITIV, KIT.

MARCO STANG received the B.Sc. and M.Sc. degrees in electrical engineering and information technology from the Karlsruhe Institute of Technology (KIT), Germany, in 2011 and 2015, respectively, where he is currently pursuing the Ph.D. degree with the Research Center for Computer Science (FZI). He is currently a Research Assistant with the FZI, KIT. Since 2017, he has been with the Group of Systems Engineering, Institute for Information Processing Technologies (ITIV), KIT.

ERIC SAX is currently the Head of the Institute for Information Processing Technology (itiv.kit.edu), Karlsruhe Institute of Technology. A tight link to industry derives from the fact that he was responsible for E/E at Daimler Buses, from 2009 to 2014, and before he was the Head of test engineering at the MBtech Group. His main area of research interests include processes, methods, and tools in systems engineering, data driven, and service-oriented architectures supported by the idea of machine learning.