Information-Centric Mobile Edge Computing for Connected Vehicle Environments: Challenges and Research Directions

Dennis Grewe*, Marco Wagner*, Mayutan Arumaithurai†, Ioannis Psaras§ and Dirk Kutscher‡
* Robert Bosch GmbH, Germany; †University of Goettingen, Germany
§University College London, UK; ‡Huawei German Research Center, Germany.

ABSTRACT

Connected vehicle systems form the basis for future features of functions and applications within the automotive domain. In order to allow resource intensive services, cloud offloading and especially Mobile Edge Computing is a promising approach. In this paper, we present a detailed futuristic vehicle scenario – Electronic Horizon – and list the challenges. We argue that the resulting challenges are representative of many of the envisioned use-cases of Mobile Edge Computing. We then present how Information-Centric Networking in combination with Mobile Edge Computing has the potential to support such a futuristic scenario. Finally, we present research directions that could enhance the solution space.

ACM Reference format:

DOI: 10.475/123.4

1 INTRODUCTION

The vision of connected automated driving is one of the major technological drivers in the automotive domain today. It already influences the development of today’s automotive services which are expected to form the basis for automated systems in the future. Such vehicle systems will heavily rely on information from cloud backends such as real-time maps, street condition information, from in-vehicle systems such as sensors, and personalized data from other domains such as smart city or smart home environments. However, the high degree of mobility as well as the processing and fusion of data from multiple sources within cloud based services challenges data delivery in a scale.

Research activities in academia and industry are investigating the Edge Computing (EC) paradigm towards high dynamic mobile networks. The EC paradigm brings computational resources, storage and services from the cloud to the edge of the network and therefore closer to the consumers [1–3], in order to reduce latency. Mobile Edge Computing (MEC) defines a specialization of edge computing to deal with characteristics of wireless networks, while providing cloud-computing capabilities at the edge of the network. Examples for such capabilities are resources for computational-intensive and time-sensitive operations or flexible deployment of applications and services.

However, the Mobile Edge Computing approach as it is deployed today, is based on coarse-grained virtual machines (VM) and heavily relies on the underlying host-centric networking model. This sets up challenges in data dissemination between the highly mobile participants as well as in address interception issues in domain name systems (DNS) due to constantly joining and leaving the network. Furthermore, it raises additional questions such as finding the closest instance of an edge cloud in dynamic networks.

In recent years, research activities are investigating communication models towards data-oriented approaches such as Information-Centric Networks (ICN) [4]. Based on a loosely coupled communication model, ICN addresses data directly using content identifiers, instead of addressing the host in the network providing the data. This fact allows mobility support by nature, while not maintaining network addresses of hosts and also facilitates features such as in-network processing and caching of data.

This paper contributes the vision of Information-Centric Networking in combination with Mobile Edge Computing in the context of connected vehicle environments. An exemplary futuristic automobile scenario – electronic horizon – is presented (see Section 2) and the potential challenges are derived (see Section 3). We then present how a combination of Information Centric Networking and Mobile Edge Computing have the potential to tackle these challenges (see Section 4). Finally, we discuss the open research directions that would enhance the solution space (see Section 5).
2 USE-CASE: ELECTRONIC HORIZON FOR VEHICLE SYSTEMS

The electronic horizon describes a cloud based virtual sensor that includes map data, vehicle’s mobility model as well as additional data regarding the road ahead. Figure 1 illustrates the idea of the electronic horizon. The vehicle utilizes data from in-vehicle systems such as the adaptive cruise control (ACC), the electronic stability program (ESP) or radar sensors as well as from external systems. Such data may include topographical information, traffic infrastructure, traffic or hazard information [5]. Based on the fusion of all data, an environment model is computed in the cloud backend by the electronic horizon function to provide a detailed preview of the road ahead. It allows for new features and functions such as adaptive cruise and predictive power-train control (e.g. gear up and down to reduce fuel or battery consumption), adaptive navigation (e.g. based on the traffic ahead) or adaptive headlight adjustment (e.g. to spot to a hazard). However, the fusion of data and computation of such model requires a high amount of computing power, while the relatively large model needs to be downloaded to the vehicle periodically during the journey. This scenario illustrates major requirements ranging from mobility aspects, network aspects such as bandwidth and latency, computational resources to process expensive operations such as the fusion of large amount of (big) data [6] produced, collected, received and processed in order to get the vital info from several sources, all the way to safety and security regarding personalized data.

3 CHALLENGES: ELECTRONIC HORIZON FOR VEHICLE SYSTEMS

The use-case described in Section 2 is a futuristic concept and there exist a number of open challenges that need to be solved before it can hit the market on a large scale. This Section tries to contribute to this approach by introducing and discussing such challenges.

In the case of an application scenario that involves connected vehicles, mobility is a major issue and includes data consumer as well as data producer. In this particular case, one of the influencing factors is the potential speed of a vehicle. It may lead to situations in which a request for data or a service may not be answered before the vehicle disconnects and re-connects to another network entry point. This leads to unfulfilled requests and potential duplication of efforts. The mobility aspects results in the challenge of coordinating these requests between edge nodes in order to solve this issue.

Another challenge to be addressed is scalability. While system design is relatively simple if the number of users and applications is rather low, it becomes difficult when both numbers rise. This includes computation issues, which can be potentially solved by some cloud technologies named in Section 4 such as containers or serverless technologies. However, such technologies challenge the (local) area network regarding load-balancing and efficient content distribution.

The establishment of suitable deployment strategies has partly things in common with the previous one. On the one hand, this includes questions like where to deploy applications/services and content. On the other hand, it also raises questions on where to place edge clouds or resources with respect to communication bandwidth, computation power or data storage. These issues grow even bigger when such a deployment is not defined statically but can vary over the lifetime of the system. For example, content or services can be moved between nodes, network entities can be added or removed dynamically, or the availability of resources changed over the lifetime of the system.

A huge show stopper regarding cloud-based systems in the automotive domain today, is the fear of having low availability of the services deployed outside the vehicle. This is driven by the constantly changing quality of the wireless connectivity and due to the fact that participants of connected vehicle environments freely join and leave the network. Other issues are for example overload scenarios within the cloud instances or the potential of hacker attacks. One option to solve this issue is the deployment of (potentially simplified) services instances on the vehicle itself, acting as a fall-back solution. However, such concept increases the
costs significantly and hence hinders the adoption of ECs in the cost-sensitive automotive industry.

Another related challenge concerns about handling and processing the large amount of data, popularly known as **Big-Data**. The issues can be classified into two folds: (i) the collection of the necessary data from the various stakeholders such as vehicles, other MEC devices, sensors on the road etc. and (ii) making use of the collected data in a timely manner in order to extract the useful information or interpretation quickly.

The most efficient communication pattern is highly dependable on the use case and the application. For example, traffic update applications relay on query-response communication, emergency notifications require push based pub/sub communication, while other applications such as live-streaming could desire real-time communication and some others could tolerate a more delay-tolerant pattern. The underlying network technology need to support mechanisms for different data dissemination strategies.

Whenever connected computer systems are created, security and privacy are a challenge. This applies especially to the automotive domain where recent hacks have decreased the customers’ sense of security. In this special scenario, the functionality of the overall application, the data of the application as well as the consumer’s personal data have to be protected. This is especially challenging due to the distributed style such EC-based systems are designed.

Another challenge occurring whenever parts of a function or system are deployed outside the vehicle is Quality of Service or Quality of Experience respectively. Depending on the application, those issues could lead to reduced functionality, the failure of a function or service, and in worst case to a danger to life for the vehicle’s passengers or the people surrounding the vehicle.

A last, but certainly an important challenge is to ensure fairness within such systems. This issue can be discussed in several dimensions. One view could be, that when using such systems, the resources available have to be shared fairly among the users or applications utilizing them. On the other hand, there have to be business models that make supplying resources such as computing power or storage interesting to potential providers. Lastly, fairness also applies to the network itself which may favor specific parties offering services or content by routing requests to them regularly.

4 POTENTIAL SOLUTION: MEC ENHANCED WITH ICN

In recent years, the challenges set by connected vehicle environments have been illuminated by several research activities. While some of these activities focused on communication technologies or protocols, others brought up proposals for designing efficient network architectures. One of the promising architectural paradigms discussed recently is Edge Computing and its specialization Mobile Edge Computing. MEC can be seen as a specific shape of cloud computing. While cloud computing tries to offload resource intensive algorithms from small client devices to more powerful servers, it raises questions regarding latency and network load. These issues are tackled by MEC: applications, characterized by specific requirements such as low-latency and bandwidth-intensive content, are executed closer to the consumers, by running on cloud components, which are installed at the edge of the network. In connected vehicle scenarios such network edges are usually given in form of cellular base transceiver stations (BTS) or road-side units (RSU) [2, 7].

The authors of [1, 3] introduce the edge computing paradigm as one of the major enabler for the Internet of Things (IoT) such as future connected vehicle environments. Furthermore, research activities such as [2, 8] provide detailed architectural approaches of edge computing in the context of connected vehicles. Besides these academic research activities, the combination of EC and connected vehicles has also been deployed recently in practical field tests [9, 10].

One example for connected vehicle applications that could benefit from the MEC is the electronic horizon introduced earlier in this paper. It combines resource-heavy algorithms that would benefit from being deployed outside the vehicle with high latency requirements. In order to fulfill both demands, instances of the electronic horizon application are running on edge components placed at RSUs or BTSs.

While MEC, as an architectural style offers several benefits, it does not define how a client communicates with a cloud instance. In many cases this is done using the Internet Protocol (IP) and hence a host-centric communication model. However, as soon as the client becomes mobile, this paradigm introduces several difficulties such as IP address maintenance. In order to overcome these issues, researchers have combined MEC with the Information-Centric Networking paradigm such as [11-14]. ICN preliminarily is used to set up a communication model for accessing named data. In ICN, content is introduced as the first class citizen of the network by separating content from its location using content identifiers, resulting in a loosely coupled communication model. Therefore, content is independent of a certain physical location which facilitates mobility, in-network caching/processing and multicast communication [4]. Below, due to space constraints, we describe two means by which ICN could enhance MEC to address the challenges of the presented futuristic automotive scenario.

4.1 Data Retrieval in Spite of Vehicle Mobility

In an MEC scenario that includes mobile nodes, turning away from depending on physical locations of data simplifies the communication. In ICNs, a vehicle sends an Interest packet to the network asking for some data independently of its
current physical location, rather than resolving the address of a specific network, node providing the data. The actual content resolution and transport to the vehicle is performed by the network. In a connected vehicle scenario, this concept ensures that the data is retrieved from the closest producer or cache and thereby natively finds the nearest MEC instance.

Figure 2 illustrates the operational steps of the electronic horizon making use of the edge computing paradigm. A vehicle requests for data about a detailed preview of the road ahead (Figure 2 step 1). This includes static named common data (e.g. traffic situation, speed limit or hazards) as well as dynamic computed results including personalized data (e.g. fuel consumption or point of interests). The electronic horizon service instance aggregates such data from multiple sources in the network and computes the results of the electronic horizon individually (Figure 2 step 2). Finally, the result is delivered back to the vehicle (Figure 2 step 3). The main benefits of this architectural design lies in receiving the result in time, while the vehicle is moving through the communication range as well as decreasing the communication in the core network. From the networking point of view, there are two different segments to be examined. One part is the exchange of data between the vehicle and the edge node. Here, ICN helps by implicitly gathering data from the closest MEC component providing it. This is due to the detachment from a host-centric communication model and the network-based content resolution of ICN. In doing so, issues caused by the mobility of the vehicle, such as frequent disconnects and re-connects to different network access points are solved implicitly. The second segment to be looked at is the connector of the edge node to the Internet. On this interface, the electronic horizon application gathers additional information. Here, ICN supports by allowing a direct addressing of the content the node is interested in, rather than exploring the network on a host-base when searching for data.

4.2 Virtualized Services

Recently, there have been numerous works that propose means to provide more efficient virtualized services. This includes for example containers (e.g. Docker) or serverless computing technologies such as unikernels or serverless computing in the cloud.

Although ICN in combination with MEC is promising, the following research challenges exist.

5 RESEARCH DIRECTIONS

5.1 Mobility

While it is to be expected that automotive services will vary in their use cases, there is one common denominator in connected vehicle environment: the high degree of mobility of network participants. While consumer mobility in ICNs is naturally supported, producer mobility (data-source is moving) still defines a research direction. While there are some first publications dealing with producer mobility such as [21, 22], most of these work cover fixed networks. There need to be further investigations regarding producer mobility (such as vehicles).

5.2 Naming

The concept of identifying content (addressing) in ICNs using mechanisms such as naming schemes describes one of the core concepts in ICN to access applications/services and their data. From a consumer perspective, there are multiple options to request for data such as query for data objects or chunks using their name or sequence number. When talking about services, querying for results (e.g. personalized function results) becomes difficult. Personalized information or parameters need to be provided by a consumer, for example
as part of the naming scheme such as the NFN [14] approach. On the one hand, research activities need to investigate the options to querying the network for computational expensive and context-sensitive service results such as [23]. On the other hand, mechanisms such as name-based routing and forwarding [24] have to cope with the mobility of network participants to deliver requests and responses of such results, especially in fast changing networks such as vehicle environments.

5.3 Routing & Forwarding

Depending on the mobility model and the service function to be executed in a MEC component, requests may not be answered in time, while a vehicle has left the network. Such node need a mechanism to access already computed results after re-connecting to the network, instead of starting the entire process of querying for results again. This also affects forwarding of data. Forwarding strategies of ICNs need to react to the mobility model of the requesting node dynamically, which is a crucial aspect when context-aware communication has to be supported. Context-aware routing and forwarding mechanisms need to ensure that network queries are requested by service instance aware of the up-to-date context about the current communication as well as the mobility model of the mobile node.

5.4 Deployment Strategies and Orchestration

Looking at the MEC components in detail, cloud technologies such as containers orunikernels simplifies the deployment of functions and services and therefore contribute significantly to provide efficient data processing and dissemination. The loosely coupled addressing concept of ICN simplifies both the access to and the placement of functions and services in the network. This is due to the fact that addresses of physical components need not be maintained and well-known by consumers as it is required in today’s host-centric networks. However, resource allocation and management of functions and services defines a non-trivial task and requires function distribution strategies. On the one hand, pervasive deployment of functions and services increases availability. On the other hand, resources are unnecessarily occupied and thus decreases network efficiency [25].

This situation becomes more challenging in terms of context-aware communication, in which certain service instances in the EC keep track of contextual information. For efficient data dissemination, such information need to be transferred from service instance to instance according to the mobility model of the mobile node. Strategies for function mobility and migration can support such transfer of context information. Such mechanisms need to be supported by QoS mechanisms, required to distinct between different types of data and services and to ensure efficient data dissemination with respect to time- and safety-critical vehicular applications. Additionally, such strategies have to ensure fairness among the different consumers and providers of the services.

5.5 Caching

Besides the distribution of functions and services, availability of data, needed for successful execution of such functions, describes another aspect. Based on the in-network caching capabilities of ICNs, predictive (proactive) caching strategies prefetching data and functions (e.g. for function chaining) from the network and storing at strategically valuable nodes can speed up computation of service results on the edge component (e.g. the result of the electronic horizon for a road segment ahead). Data is already available at the nodes and need not be collected from the network on demand (reactive approach), which reduces computation time and therefore latency. While some first publications showed caching approaches to place data in ICN networks proactively (e.g. [26] or [27]), strategies to serve multiple service instances accessing cached data concurrently are required, while the number of cached copies is low to avoid cache pollution. Regarding connected vehicle environments, cached copies of data at multiple edge components (multi-homing) using multiple available communication technologies concurrently (multi-channel - e.g. using Wifi and cellular) are aspiring to ensure Quality of Service (QoS) and Quality of Experience (QoE) of automotive services.

5.6 Data Dissemination

Data dissemination patterns that could better support high mobility scenarios including and is a desired feature. Current ICN solutions are designed for fixed networks, but might be inefficient in connected vehicle scenarios. For example, the ICN approaches Content-Centric Network (CCN) [4] or Named-Data Networking (NDN) [28] relies on a pull-based communication and a reverse path forwarding pattern. Such approach questions efficiency while highly mobile consumers freely join and leave the network and forwarding paths are changing frequently. For example, [29] discusses the need of different dissemination patterns in the IoT. Other approaches such as COPSS [30] introducing a push-based approach for ICNs. COPSS is designed as a subscription tree based approach introducing Rendezvous Points to deal with high mobility of consumers.

5.7 Security & Privacy

Another field affected by the loose coupled communication model of ICNs is security. In ICNs, security features are directly introduced as part of the content itself [4, 28], instead of the transport layer as given in today’s connection-oriented networks. Regarding in-network caching, security is intensifies by the fact that data is expected to stay within untrusted caching nodes. This also includes privacy concerns while requesting personalized service results. In recent years,
6 CONCLUSION

The work in this paper describes the vision of combining both introduced network paradigms Information-Centric Networking and Mobile Edge Computing together in the context of connected vehicle environments. The resulting advantages are described by providing access to resource intensive services in fast changing networks and characterized by specific requirements such as low-latency and bandwidth-intensive content.

In consideration of the electronic horizon automotive use case scenario, the paper discusses the different elements of the use case, introduces challenges such as mobility, availability, privacy and security and contributes in open research directions for the combination of Information-Centric Networking and Fog Computing.

The paper has shown that there are still open research directions such as context-based selection and dynamic orchestration of automotive services, naming strategies to discover services and to deal with the high degree of mobility as well as disseminate data efficiently through the core network to decrease computation time. Especially in the automotive industry, safety and security mechanisms define important roles to provide a maximum of protection for passengers and its surroundings.

Future work needs to address the open research points. By setting up the use case scenario, for example by both simulation and proof of concept prototypes, further investigations and evaluations of mechanisms have to be considered with respect to the introduced elements of the use case and the open research directions.

REFERENCES

[17] [Online]. Available: https://linuxcontainers.org/