Interconnecting Smart Cities by Vehicles: How feasible is it?

Florian Hagenauser†, Christoph Sommer†, Ryokichi Onishi*, Matthias Wilhelm*, Falko Dressler†, and Onur Altintas

*Toyota InfoTechnology Center, Tokyo, Japan
†Dept. of Computer Science, Paderborn University, Germany

{florian,onishi,matthias,onur}@jp.toyota-itc.com
{hagenauser,sommer,dressler}@ccs-labs.org

Abstract—In future smart cities cars will play an important role not only by transporting people and goods, but also as information hubs. Cars will be equipped with various communication technologies and will be able to offer their own resources like data storage, processing power, and sensor data (e.g., camera pictures or temperature). In Car4ICT, cars in smart cities connect service producers and consumers by offering service discovery for them and transferring the data in-between. But, the proposed architecture cannot be used as-is for interconnecting smart cities or transferring data from and to rural areas. This is due to mobility patterns outside of cities being different and requiring specialized algorithms. Therefore, in this paper we extend Car4ICT to interconnect smart cities and rural areas. The extension still relies on IVC as cellular coverage in rural areas is often sparse. By covering additional areas, the number of available services increases and a redundant communication link between smart cities is added. To support long-distance communication, we changed core parts of the architecture and evaluated the changes with extensive simulations. For this, we used a segment from the freeway between Tokyo and Osaka in Japan and evaluated the transmission delay for various distances. The foundation for these simulations was real world traffic data provided by Japanese authorities. Based on these simulations we can show for which distance Car4ICT in a freeway scenario is still feasible.

I. INTRODUCTION

Future smart cities will generate and process a lot of data. As vehicles are continuously improving in terms of processing power and networking capacities, they are one of the most promising building blocks for these smart cities. We envisioned the concept of Car4ICT [1], which puts cars at the center of a service discovery oriented architecture. Various non-safety critical services (e.g., traffic information, sharing video files, weather forecasts, processing power) can be offered and consumed using the architecture by various entities: humans (usually connected to the Car4ICT network via a smartphone) and machines (e.g., cars or computers). Vehicles support the architecture by enabling service discovery and data transfer between consumers and producers. Service Providers (producers) offer their services over the car-based network, while Service Consumers (consumers) are able to search and use their desired services. Previously the concept was designed and evaluated with a focus on urban scenarios [1] which cover already a wide range of different use cases. But not all use cases can be covered by this city-centric design – especially rural areas and connections between cities are not covered. Such intercity and rural connections have multiple advantages. First, they allow users to reach services farther away and in turn greatly increase the number of available services. An example of a non-time critical service enabled by these links is the gathering of traffic information for a route ahead. In this case, a user can react to traffic changes on the road ahead in a timely manner. Another possible service is to gather highly localized weather data from along the route or from the destination in order to generate a personalized weather forecast. Such data can be gathered from sensor readings (e.g., humidity, temperature) offered by cars or other sensors along the road. Second, if there exists a connection between two cities via Car4ICT, this adds an additional, independent communication link. This link is independent of existing infrastructure and even exists in case others are broken (e.g., in a disaster case or power outages). Therefore, adding this link makes the communication between multiple smart cities more resilient.

Additionally, in many countries, cellular networks cover cities, but barely exist outside of metropolitan areas. This is especially true for third world countries, but also holds for more developed nations. A report covering east Africa, stated that, while 80% of the population are covered by a mobile network, only 50% of the land area are covered [2]. The lowest fraction of covered land, 35%, was reported in Kenya. Even if the area is covered by a cellular network, this does not necessarily imply that data communication is possible as a report investigating the connectivity in Frankfurt am Main, Germany [3] discovered. While driving through the city, the authors were not able to establish a connection with a central server in 8% of the cases. Therefore, to reach areas without cellular coverage, it is easier to rely on IVC instead of building new infrastructure for cellular networks.

In this paper, we extend the Car4ICT framework to operate between city boundaries, with a focus on freeways. The vehicle mobility on such freeways and in rural areas is different from city scenarios which means that the Car4ICT architecture can not be used as-is for these new scenarios. We changed the specific protocol for data transfer and updated the service discovery to accommodate this different mobility on a freeway. Finally, we enhanced our proposed data structure to identify
services and support the aforementioned changes. With these updates, Car4ICT enables consumers and providers to provide and use services in rural areas or distant cities. To investigate how feasible these extensions are, we performed simulations based on real world traffic data for a freeway segment in Japan. This will tell us what delays are achievable via Car4ICT and what kind of services can still be used over such long distances.

II. RELATED WORK

Noguchi et al. [4] propose a Service Discovery Protocol (SDP) based on IPv6 and geonetworking. To implement their approach, an additional layer is added between network and link layer, and a mapping of IPv6 addresses to coordinates and radii is introduced. Such additional layers and mappings increase the complexity of the IPv6 stack even further. Lakas et al. [5] propose another SDP for VANETs and introduce mobile directories. Their request scheme is broadcast based and includes a time-to-live as well as some custom caching of service replies. Abrougui et al. [6] present another SDP protocol which aims to offer three types of services: fixed, moving, and migratory services. The last type denotes services that stay in a certain area but are moved between vehicles when one of them leaves the area. Their system relies on clustered roadside units (denoted RR's) to work properly, which makes it impracticable for Car4ICT.

There exist already a multitude of routing algorithms for vehicular networks, but none which successfully combine urban and rural scenarios. Fonseca and Vazão [7] compare various routing algorithms designed for VANETs and survey whether there is a single algorithm usable in both urban and freeway scenarios. Their conclusion is that none of the existing ones is suitable for all scenarios. Therefore, either new ones should be devised (preferable) or algorithms which are unsuitable in one of the two scenarios have to be used (current reality).

The architecture of Car4ICT for urban scenarios builds on the concepts of SDPs as well as Information-Centric Networking (ICN) and Named Data Networking (NDN). First, Car4ICT uses a service discovery approach to enable users to provide and request services. In Car4ICT’s SDP different entities become the providers and consumers and cars take the role of directories. Second, to make it possible to describe services, we introduced the concept of identifiers [1], which is based on recent developments related to ICN [8] and NDN [9].

Due to the differences in vehicular mobility between urban and rural scenarios, the protocols envisioned for Car4ICT in urban scenarios cannot be used directly and another method for service discovery and data transfer is required. The presented protocol is mainly based on Contention Based Forwarding (CBF), which was first introduced by Füßler et al. [10]. CBF is a receiver-centric routing protocol where the buffer time of messages is inversely proportional to the message progress towards the destination. When one car forwards the packets, others overhear this and understand it as an implicit acknowledgment and do not further forward the message themselves.

Although Füßler et al. [10] propose to use the protocol as-is and do not introduce further changes to make it more suitable for vehicular networks, others, e.g., [11], [12], tried to increase the protocol’s performance. Al-Kubati et al. [11] assume that CBF is not well suited for city scenarios as crowded intersections will lead to packet collisions. In their protocol, the vehicles with the highest capability to forward in multiple directions, are in charge of transmitting the data further. They claim that their protocol works in urban and freeway scenarios but evaluate it only in a Manhattan scenario. Salvo et al. [12] propose to add the angle the message was received to the deciding parameters for forwarding a message. Their Triangle Forwarding Rule can be used in three different ways, each one differing in complexity of the calculation and in the way which angles are used. Again, their protocol is only evaluated in a Manhattan scenario and not in a freeway scenario.

Finally, the European ETSI ITS-G5 standard also includes a variation of CBF [13]. The essential difference is that, in case of the ETSI version, packets that travel farther than a defined maximum distance are also forwarded and not discarded. In addition, it adds a minimal buffer time, which increases the flexibility for configuring the parameters.

III. INTERCITY CONNECTIONS VIA CAR4ICT

When using Car4ICT in city scenarios, cars exchange their service tables via beacons. Therefore, consumers are able to find an offered service quickly [1] and exchanging data can be done in a similar or georouting based fashion. While urban environments are a good starting point, we want to extend the concept to connect smart cities with each other and link rural areas to the Car4ICT network. In the case of Car4ICT, services can be offered to far-away consumers via long distance connections (i.e. longer than a few hops). This greatly increases the number of available services and makes the link interconnecting the cities more stable. In addition, by using IVC, it is possible to reach remote areas without any cellular infrastructure. Such an area could be on the countryside where no infrastructure exists or an area hit by a disaster. To make this possible, a contribution of this paper is the addition of georouting to Car4ICT in freeway scenarios. This in turn allows consumers to search for services farther away.

The used routing is based on two different concepts, namely Contention Based Forwarding (CBF) and Store-Carry-Forward (SCF). First, CBF is the basic routing protocol where every receiver of a message decides for itself if the message should be forwarded. Second, SCF is used to store messages and forward them later in case no suitable forwarder is found. Due to implicit acknowledgments used in CBF it is possible to use SCF without explicit acknowledgments.

As already discussed, CBF is not an ideal solution for city scenarios, but is supposed to perform well on freeways. The original CBF formula [10] calculated the buffer time as

\[ t = \begin{cases} 
   t_{max}(1 - p_t/p_{max}) & \text{if } 0 \leq p_t < p_{max} \\
   \infty & \text{otherwise}
\end{cases} \quad (1) \]

Although Füßler et al. [10] propose to use the protocol as-is and do not introduce further changes to make it more suitable for vehicular networks, others, e.g., [11], [12], tried to increase the protocol’s performance. Al-Kubati et al. [11] assume that CBF is not well suited for city scenarios as crowded intersections will lead to packet collisions. In their protocol, the vehicles with the highest capability to forward in multiple directions, are in charge of transmitting the data further. They claim that their protocol works in urban and freeway scenarios but evaluate it only in a Manhattan scenario. Salvo et al. [12] propose to add the angle the message was received to the deciding parameters for forwarding a message. Their Triangle Forwarding Rule can be used in three different ways, each one differing in complexity of the calculation and in the way which angles are used. Again, their protocol is only evaluated in a Manhattan scenario and not in a freeway scenario.

Finally, the European ETSI ITS-G5 standard also includes a variation of CBF [13]. The essential difference is that, in case of the ETSI version, packets that travel farther than a defined maximum distance are also forwarded and not discarded. In addition, it adds a minimal buffer time, which increases the flexibility for configuring the parameters.
where \( t_{\text{max}} \) is the largest possible time a packet is buffered before being sent, \( p_i \) is the packet progress and \( p_{\text{max}} \) the assumed maximum transmission distance. This allows a car to calculate the buffer time \( t \) of a message inversely proportional to the transmission progress. In the CBF version standardized by ETSI, the buffer time is calculated as

\[
t = \begin{cases} 
    t_{\text{max}} + \frac{t_{\text{min}} - t_{\text{max}}}{p_{\text{max}}} \times p_i & \text{if } 0 \leq p_i \leq p_{\text{max}} \\
    t_{\text{min}} & \text{otherwise}
\end{cases}
\]

(2)

which adds the parameter \( t_{\text{min}} \). In case of \( t_{\text{min}} = 0 \), the first two cases of the formulas are equal, but generally the addition of \( t_{\text{min}} \) gives Equation (2) more flexibility when configuring the buffer interval for all cars. The bigger difference between the two formulas is in the handling of messages which have been transmitted farther than the maximum expected transmission distance (i.e., when \( p_i > p_{\text{max}} \)). While Equation (1) discards such messages, Equation (2) actually transmits them after the smallest possible buffer time \( t_{\text{min}} \). The second case is more useful, as it allows the protocol to still work in case \( p_{\text{max}} \) was configured too low. Still, setting \( p_{\text{max}} \) far too low could lead to packet collisions because too many cars send the packet after buffering it for \( t_{\text{min}} \).

After waiting for the calculated buffer time \( t \) and not receiving the message from anyone else, cars broadcast the message. Other cars receiving this message cancel their buffer timer and will not send the message, following the concept of an implicit acknowledgment. In addition, cars which sent the message store it and start a timer to potentially perform simple SCF. In case the car forwarded the message (i.e., the buffer time expired) and no one forwarded the message afterwards (i.e., the SCF timer expired before receiving an acknowledgment) the car sends the message again. This method allows the message to traverse sections without forwarding nodes. This is especially crucial on a freeway as the network topology might not change as much as in a city and the traffic density might be smaller for longer distances.

With these additions, it is possible to send requests for distant services. There are now two options: service discovery is performed in advance, or, if the service is already known (e.g., in case of a weather station or a traffic information provider), the known identifier can be used immediately to reduce the overall delay for service discovery and data transfer. The second case helps for example when looking for weather forecasts and the providing sensors are already known from a previous discovery. Figure 1 gives an overview of Car4ICT and shows where CBF and SCF fit into the whole picture. In this case, a car offers some of its sensor readings (e.g., environmental temperature), therefore acting as a provider. There are then two different consumers interested in this data, a user who is traveling to the area of the provider and a computer that gathers various sensor readings to calculate weather forecasts. Both send a request to a neighboring Car4ICT member, which then forwards the message towards the requested destination (if no local entry fits the request). For freeway scenarios, this is done as outlined via CBF and SCF. If a suitable service is found, a reply is sent back to the consumer, which can then choose which provider to use. The data transfer between consumer and provider is handled by the Car4ICT members as well.

To clearly identify services, we [1] introduced identifiers. These identifiers consist of a hash and a list of key-value pairs, the metadata, to describe them further. While the hash is mandatory for offers, there are no restrictions on metadata. This allows to use identifiers as queries when searching for a service. For example, in the case of searching for a video, the consumer can omit the hash and only rely on metadata (e.g., using only fileType=video, size=1GB). Such an identifier is only matched to other metadata (and not to any hash value) in a service table, and subsequently a list of fitting entries is sent back.

In order to make these identifiers more powerful, we add two more mandatory fields: Position and Validity. Position is introduced to support georouting; without it, it would not be possible to use SCF. In addition, it can also be used for restricting requests to certain areas. This can be useful if the provided service is only interesting for a city district and the service offer should stay in this area. Validity moves the task of removing service table entries from the directories (i.e., the car) to the provider, which then is in charge of defining a validity of its offered services.

IV. SIMULATION STUDY

For some services, like weather forecasts, longer delays (i.e., range of minutes) are acceptable, but for others (e.g., traffic information) shorter delays are more suitable. To see which of Car4ICT's services are still possible on a freeway scenario we performed simulations in a realistic freeway scenario with traffic densities based on real world data.

For this, we perform realistic simulations on a freeway segment between Tokyo and Osaka, Japan. As can be seen in Figure 2, the chosen segment of the freeway is on the Tomei
Expressway and is located between Gotemba and Mikkabi in Shizuoka prefecture. One of the main reasons for choosing this segment was the good availability of real traffic data for it. The map data was taken from OpenStreetMap\(^1\) and converted using the provided tools of the mobility simulator SUMO \cite{SUMO}. The simulated segment is roughly 150 km in length and mostly consists of two (sometimes three) lanes per direction.

Figure 3 shows the vehicle densities in both directions on the freeway from east to west. We obtained two kinds of statistics for every intersection from traffic authorities: the number of cars using it \cite{traffic_authorities} and the number of cars driving towards it \cite{traffic_authorities}. We choose to take the hours with the highest (i.e. 5pm to 6pm) and lowest (i.e. 5am to 6am) vehicle density as well as one hour (i.e. 2pm to 3pm) with an average traffic density. The small spikes in the data are freeway intersections where more traffic occurs compared to the lanes in between intersections. This happens if there is an additional lane which is used by vehicles entering and leaving the freeway at the same time.

### A. Traffic Generation

For simulations the routes people took are required as an input. Such routes are usually in the form of Origin-Destination (O/D) matrices which describe how many cars are driving from a certain origin to a certain destination. We denote these values as the number of cars \(f_{i \rightarrow j}\) arriving on the freeway at intersection \(i\) and exiting the freeway at a later intersection \(j\).

Unfortunately, the available statistics contain only the total amount of cars that use a certain freeway segment, and/or the amount of cars that enter and exit at a certain intersection. Such data is usually gathered by induction loops installed on the street. Thus, only two parameters are known for each intersection \(i\): first, the combined number of cars \(T_i\) accessing and exiting the freeway at this intersection; and second, the number of cars \(c_i\) arriving from the previous intersection. This available data for the selected freeway does not directly translate to such matrices.

The following concept (illustrated in Figure 4) explains how we translated the available data to get O/D matrices. We first calculate how many cars are arriving on (or exiting from) the freeway at each intersection as \(a_i\) and \(e_i\), respectively:

\[
\begin{align*}
    a_0 &= c_0, \\
    a_i &= \frac{1}{2} (c_{i+1} - c_i + T_i), \\
    e_i &= T_i - a_i.
\end{align*}
\]

Based on these numbers we estimate \(f_{i \rightarrow j}\), the component values of our O/D matrices, as

\[
    f_{i \rightarrow j} = \begin{cases} 
    e_j / c_j \times \left( a_i - \sum_{x=i+1}^{j-1} f_{x \rightarrow j} \right) & \text{if } j > i + 1 \\
    e_j / c_j \times a_i & \text{if } j = i + 1.
    \end{cases}
\]

If input values are known for each hour of a typical 24 h workday, we are able to derive one O/D matrix per hour and thus generate time-varying traffic flows for our freeway.

As a simulation tool we use the mature Veins simulation environment \cite{Veins}, which couples the network simulator OMNeT++ to the mobility simulator SUMO. Veins provides IEEE 802.11p, on top of which we simulated Car4ICT. As Car4ICT does not depend on a specific physical layer access protocol and the model for 5.9 GHz is well established we also used this for our simulations instead of the 700 MHz used by the Japanese ARIB T109 standard. SUMO allows us to generate a realistic map including buildings around the freeway, which in turn is used by Veins to calculate realistic shadowing effects for the radio transmissions.

These three components, a realistic map, real traffic data, and a well established simulation environment, allow us to perform detailed simulations. Further simulation parameters can be found in Table I.

### B. Parameter Study

As an initial step we performed a parameter study for the CBF formula. We modified the values of \(p_{\text{max}}\), the expected maximum transmission distance, and \(t_{\text{max}}\), the maximum buffer time, according to Table I. These simulations were performed for data transfer over a distance of 100 km, with traffic densities corresponding to the traffic volume between 2pm and 3pm.

As can be seen in Figure 5, increasing \(t_{\text{max}}\) leads to longer overall message delays. The same can be said for increasing \(p_{\text{max}}\), but for different reasons. In the case of increasing \(p_{\text{max}}\),

---

1\text{http://www.openstreetmap.org}
Figure 5. Results of the parameter study (delays for a distance of 100 km) for varying $p_{\text{max}}$ and $t_{\text{max}}$.

the range for the buffer time becomes larger, which in turn leads to longer buffered messages. By having a small $p_{\text{max}}$, some message might cover larger distances than the intended $p_{\text{max}}$. This does not happen anymore when $p_{\text{max}}$ is increased. On the contrary, it is even more likely that messages are transmitted after buffering them for a longer time.

Four different relations are possible between $t$ and $p_i$. A case of $p_i = 0$ with $t = t_{\text{max}}$ will rarely happen, as this means that a message makes no progress at all. Most of the time a message will be buffered in a case like $p_i \in [0, p_{\text{max}}]$ with $t \in [t_{\text{min}}, t_{\text{max}}]$. But, if $p_i$ is greater than or equal to $p_{\text{max}}$, $t_{\text{min}}$ is chosen as a buffer time. Thus, if many vehicles receive a message where the one-hop distance $p_i$ is at least $p_{\text{max}}$, the channel might get congested. On the other hand, if $p_{\text{max}}$ is much greater than the highest achieved $p_i$, then buffer times will never be close to $t_{\text{min}}$ and therefore the overall delay will be greater. This can be observed in Figure 5, especially by comparing the different values for $p_{\text{max}}$. Furthermore, for all three densities, the average distance per hop is between 400 m and 500 m. This indicates that setting $p_{\text{max}}$ to 500, instead of the proposed 1000 by ETSI, leads to better results in a freeway scenario. In this parameter study we can see that CBF provides reasonable delays even over a distance of 100 km if the right parameters are used.

Overall, we can see from these results that most of the time the delay for messages will be small enough to provide all services except safety related ones.

C. Penetration Rate Study

Finally, we study what effect the penetration rate $r$ of Car4ICT equipped cars has on the delay. This means we reduce the rate of cars equipped with Car4ICT module from 100% to 75%, 50%, and 25%. Figure 6 shows the same plot for four different penetration rates. Plots are shown for all four simulated distances with a traffic density from 2pm and the ETSI CBF parameters. To make sure data can traverse the whole distance of 100 km we disabled the maximum number of hops and simulated the scenario for a duration of 3600 s. This was done so that we can observe also longer delays in case of small penetration rates. As can be seen, if every car is equipped with a Car4ICT module (i.e., $r = 1$), all messages arrive without any problems and the delay is at maximum 12 s for a distance of 100 km. For a rate of 0.75, the results are pretty similar, indicating such a penetration rate is big enough to keep the architecture working. In case of $r = 0.5$ messages still cover the distance with a very short delay, only some outliers for 50 km (less than 20%) and for 100 km (less than 25%). Finally, for a penetration rate of 0.25 all messages for 10 km arrive and a high number for a distance of 30 km and 50 km (more than 90%). Only for the longest distance of 100 km barely any message is able to cover the distance in an hour. This is mostly related to the simple SCF algorithm used in which some cases lets cars in the wrong direction carry the message and therefore moving it away from the destination. We consider a more sophisticated algorithm for such cases as interesting future work. To summarize the penetration rate study, slightly smaller rates do not affect the performance of Car4ICT but if they get too small, the delay might increase to uncomfortable long times.

D. Comparison to other approaches

Beside using the Car4ICT algorithm based on CBF, other ways of communication could be used. If the system would rely only on cellular communications, data exchange between two cities would work in nearly all cases. But if the user searches for a service along a freeway, e.g., for traffic information, cellular networks are not a good choice. There might be occasional coverage along a road, but as discussed, rural areas are usually not well covered. Therefore, relying on IVC gives an advantage when trying to reach rural areas.

If the CBF-based approach is compared to a data mule approach other things can be observed. When relying on a data mule, a car never forwards the message, but carries it to the destination. While this is an impractical approach for a real-world deployment, it helps understanding how the presented results relate to a baseline. In our simulations all vehicles traveling over 100 km had an average speed of 97 km/h. Therefore, it would take them roughly an hour to carry the message from the source to the destination. If this time is compared to the penetration rate study in Figure 6, in all cases the CBF approach is faster. Only for the latest case, for a penetration rate of 25%, the data mule approach can help achieving a higher success rate. But, as already mentioned, a more sophisticated algorithm probably improves the success rate of the CBF-based approach.

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>simulation duration</td>
<td>100–900 s, 3600 s</td>
</tr>
<tr>
<td>distance consumer→provider</td>
<td>10, 30, 50 and 100 km</td>
</tr>
<tr>
<td>$t_{\text{min}}$</td>
<td>1 ms</td>
</tr>
<tr>
<td>$t_{\text{max}}$</td>
<td>100 ms, 200 ms, 300 ms</td>
</tr>
<tr>
<td>$p_{\text{max}}$</td>
<td>500 m, 750 m, 1000 m</td>
</tr>
<tr>
<td>penetration rate $r$</td>
<td>0.25, 0.5, 0.75, 1</td>
</tr>
<tr>
<td>SCF interval</td>
<td>2 s</td>
</tr>
<tr>
<td>physical layer</td>
<td>IEEE 802.11p</td>
</tr>
<tr>
<td>max transmission power</td>
<td>20 mW</td>
</tr>
</tbody>
</table>
V. CONCLUSION

In this paper we extend the Car4ICT concept to interconnect smart cities and rural areas. This extension increases the number of available services and makes the connecting link between smart cities much more resilient if other infrastructure fails. We evaluate the extension with an extensive simulation study regarding the usage of CBF in a freeway scenario. To simulate a realistic freeway, we developed a new scheme to generate realistic traffic for a freeway based on data provided by Japanese authorities. The generated traffic was then used to simulate traffic for a 150 km long segment on a freeway in Japan. Our results indicate that CBF works well for such a freeway scenario. This in turn supports our approach to use CBF in Car4ICT to interconnect smart cities as well as urban areas. When using different parameters for CBF instead of the ones proposed by ETSI, our results indicate that the performance might get better if the parameters are adapted to more realistic values for the maximum transmission distance. By varying the penetration rate, we could also show that only a very low penetration rate leads to service disruptions. In such a case it could be an option to rely on cellular connections for reducing the delay. Such an addition of cellular connections, as well as a unified option to rely on cellular connections for reducing the delay, leads to service disruptions. In such a case it could be an option to rely on cellular connections for reducing the delay.

VI. ACKNOWLEDGMENTS

We want to thank Makiko Matsumoto for her work and support during the work on this paper.

REFERENCES


