On the Serviceability of Mobile Vehicular Cloudlets in a Large-Scale Urban Environment

Chuanmeizhi Wang, Yong Li, Member, IEEE, Depeng Jin, Member, IEEE, and Sheng Chen, Fellow, IEEE

Abstract—Recently, cloud computing technology has been utilized to make vehicles on roads smarter and to offer better driving experience. Consequently, the concept of mobile vehicular cloudlet (MVC) was born, where nearby smart vehicles were connected to provide cloud computing services locally. Existing researches focus on MVC system models and architectures, and no work to date addresses the critical question of what is the potential, i.e., level of local cloud computing service, achievable by MVCs in real-world large-scale urban environments. This issue is fundamental to the practical implementation of MVC technology. Answering this question is also challenging because MVCs operate in highly complicated and dynamic environments. In this paper, we directly address this challenging issue and we introduce the concept of serviceability to measure the ability of an MVC to provide cloud computing service. In particular, we evaluate this measure in practical environments through a real-world vehicular mobility trace of Beijing. Using the time-varying graph model for mobile cloud computing under different scenarios, we find that the serviceability has a relationship with the delay tolerance of the undertaken computational task, which can be described by two characteristic parameters. The evolution of serviceability through a day and the influence of network congestion are also analyzed. We also portray the spatial distribution of the serviceability and analyze the influence of connectivity and mobility in both MVC and vehicle levels. Our observations are valuable to assist designing vehicular cloud computing systems and applications, as well as to help make offloading decisions.

Index Terms—Vehicular networks, vehicular cloud computing (VCC), mobile vehicular cloudlets (MVCs), serviceability.

I. INTRODUCTION

The number of vehicles in operation worldwide has surpassed the 1 billion-unit mark in 2010 [1], and it is continuously increasing with a remarkable speed. Such a trend makes the implementation of an intelligent transportation system in urban environment increasingly important. One of the challenging issues caused by the huge amount of vehicles is traffic congestion, which wastes a massive amount of time, fuel and money. Given United States in 2011 as an example, 5.5 billion hours of extra time and fuel enough to fill four New Orleans Superdomes were wasted due to congestion [2]. In addition, drivers and passengers may need many customized services on road, some of which require real-time provisioning [3]. For example, an emerging application called vehicular augmented reality [4] helping drivers see around blind spots may require huge amount of computation.

Fortunately, there are already various solutions developed in the past decade in offering better driving experience [5]. A typical modern car or truck equipped with an on-board computer and other devices is able to interact with the Internet to obtain associated services. As technology advances in embedding sophisticated resources on individual vehicles, it can be predicted that even the low-end vehicles will be equipped with on-board wireless communications and data collection devices in the near future. With the access to the Internet, information collected by vehicles are uploaded to remote data centers to enable cloud-assisted services for vehicles under the traditional client-server communication model, which takes smart vehicles as ordinary mobile devices.

We investigate the power of smart vehicles in terms of their increasing memory and computational capability to improve the performance of vehicular cloud computing, including the convenience on handling both local and real-time queries. In order to fully utilize the nearby computational resources with a timely and efficient manner, a peer-to-peer communication model, which is proposed for mobile cloud computing [6], can be specially applied to smart vehicles. As shown in Fig. 1, a group of nearby smart vehicles can connect to form a mobile vehicular cloudlet (MVC) via dedicated short-range communications (DSRC), which have 75 MHz of licensed spectrum.
in the 5.9 GHz band allocated by the Federal Communications Commissions in support of vehicular networking in the United States [7]. Both vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications can exist in an MVC, since there may exist roadside units (RSUs) connecting to the vehicles to strengthen the connectivity of the MVC. Each smart vehicle in an MVC, referred to as a node, can be a computing service provider as well as a client of the service. By dividing the task among MVC nodes, the initiator can speedup computing while conserving energy consumption [8]. Since cloud computing service is accessed instantly through local interactions in the MVC, the communication latency and remote data roaming charges are eliminated. Applications like vehicular augmented reality can be provided or accelerated by local MVCs.

However, practical difficulties, including limited communication ranges, high speeds and unpredictable moving patterns of vehicles, severely constrain the serviceability of MVCs, i.e., the ability to provide computing service to the initiator. To be more specific, the serviceability that an initiator possesses equals to the maximal possible offloadable amount of computation. Since the serviceability measure considers the process of service, it is better and more relevant than other network measures when dealing with computing service problem. Unfortunately, most existing related researches focus on architectures and systems [5] without clear knowledge and careful considerations on the answer to a fundamental question: what is the level of serviceability of MVCs in the real world, especially in large-scale urban environment? This is a very complicate question due to two major reasons: i) the influence of practical difficulties mentioned above, and ii) the features of serviceability level are currently unknown in both spatial and temporal dimensions.

To answer this fundamental question, we propose a novel model and an algorithm to describe and evaluate the serviceability of MVCs for the first time. Moreover, we use a real-world large-scale urban vehicular mobility trace rather than a synthetic trace to observe the serviceability of MVCs through a comprehensive study. The novelty and contributions of our work can be summarized as follows.

- For the first time we provide the answer to the fundamental question considering the serviceability level of MVCs, which has never been considered before, through using a real-world large-scale urban vehicular mobility trace.
- We exploit the inner relationships among the parameters in our model with the information excavated from the vehicular mobility trace to extend our analysis method to other large cities. In particular, we discover the influence of connectivity and mobility of vehicles on the serviceability of MVCs.

The rest of this paper is organized as follows. After discussing the related work in Section II, we present the relevant models, definitions and algorithms in Section III. Then we introduce our dataset and processing method in Section IV. In Sections V and VI, we analyze the statistical results to evaluate the serviceability of MVCs and we exploit the inner relationships among the parameters of our model, both macroscopically and microscopically. The versatility and significance of our findings are discussed in Section VII, and the conclusion is offered in Section VIII.

II. RELATED WORK

Traditionally, vehicular cloud computing (VCC), as a specific case of mobile cloud computing (MCC) [9], has been introduced as an enabling technology for vehicles equipped with mobile communication devices. MCC integrates remote cloud into mobile environment to overcome the obstacles related to the performance, environment and security in mobile computing [10]. However, because traditional MCC uses the client-server communication model, uploading real-time information to the cloud by using the Internet is costly and time consuming [11]. Therefore, a peer-to-peer communication model for MCC [12]–[15] is proposed to enhance the memory and computational power of mobile devices [6], [8]. In these works, a mobile cloudlet is a group of nearby mobile devices connected by short range communications [8]. However, these works are limited to low capability devices and static scenario, or restricted to distributed computing approaches, and they have serious scalability and reliability issues.

A recent work [16] focuses on the scenario of vehicular ad-hoc networks and apply the peer-to-peer communication model to smart vehicles to fully utilize the nearby computational resources with a timely and efficient manner. However, this approach suffers from several drawbacks, including the high cost of the service constrained communications due to the high mobility of the vehicles [17], [18]. Different from proposing a cloud computing framework, we focus on evaluating the serviceability of MVCs.

The properties of MVCs or vehicular ad-hoc networks, which are heavily influenced by the mobility of vehicles, are highly complex in the real world, and we still lack sufficient understanding of them. Many studies on vehicular networks rely on synthetic mobility traces or analytical tools. For example, the studies on highway traffic [19]–[23] are carried out in this way, and they aim at determining which combinations of conditions, including vehicular density, car speed and communication range, are required to achieve full connectivity. However, these works focus on highway traffic, which is not our interest of urban environment. There are other studies consider the microscopic and macroscopic mobility of vehicles in urban scenario but they are based on small-scale regular road grid [24]–[27], which cannot represent real-world large-scale urban environments.

More recent studies have considered real-world urban areas [28]–[30]. Specifically, Conceia cao et al. [28] focus on the connectivity of vehicles in the city of Porto, Portugal, and Pallis et al. [29] focus on the network topology in Zurich, Switzerland, while Naboulsi et al. [30] focus on the instantaneous topology in Cologne, Germany. However, these works have several important differences from our work. Firstly, these studies focus on the analysis of vehicular network properties, but our work analyzes the computational ability of MVCs. Secondly, these existing studies use synthetically generated vehicular mobility traces, while our study is conducted on a real-world large-scale vehicular mobility trace. Thirdly, the works [28] and [29] are either limited to a small area of a city or over a short time period, which may not be sufficient to analyze the performance of MVCs in a large-scale urban environment.
Another existing work [31] has used real mobility data traces to study dynamic sociality in vehicular networks. However, it focuses on temporal correlations of vehicular sociality to select the most centric vehicles as seeds for mobile advertising, while our work studies the serviceability of whole MVC, rather than only considering which vehicle should be the center.

III. MODELS, DEFINITIONS AND ALGORITHMS

An MVC is generally initiated by a vehicle, denoted by \(v\), that demands a certain computing service from the vehicular network. Obviously, the serviceability of the MVC for \(v\) is related to the connectivity and mobility of vehicles over time in the vehicular network. To give a formal and mathematical definition of the serviceability of MVC, we first model the vehicular network topology over a time span as a time-varying graph (TVG). In [32], Casteigs et al. have proposed a unified framework for TVGs and we will use part of their definitions and notations. In this paper, the set of vehicles is denoted by \(V\), which V2V communication has a packet delivery ratio close to 1, and each vehicle is regarded as a node. Then in a TVG, an MVC at each time point becomes a group of connected nodes in the graph and the serviceability of an MVC can be defined as the feature of node \(v\), namely, the initiator. Based on this viewpoint, the definition of TVG is presented as follows.

**Definition 1 (Time-Varying Graph):** A TVG, denoted by \(G = (V, E, T, \rho)\), describes the dynamics of the system, where:

- \(V\) is the set of nodes, i.e., vehicles in the system. The weight of node \(v_i\) denoted by \(p_i\) represents its computing speed in units.
- \(E \subset V \times V\) is the set of possible edges, and an edge \(e \in E\) indicates that the two vehicles connected by \(e\) may potentially communicate with each other.
- \(T\) is the time span over which the links between nodes are assumed to take place.
- \(Q\) is the set of link conditions, and an element \(q = (R_V, h) \in Q\) indicates a link condition, in which \(R_V\) is the V2V communication range and \(h\) indicates whether multi-hop communication is enabled.
- The relation \(\rho(e, q, t) : E \times Q \times T \to \{0, 1\}\) indicates whether the two nodes connected by \(e\) can actually communicate with each other at time \(t\) under the link condition \(q\).

The RF signal propagation model adopted in our work is a simple disc model, which assumes that successful V2V communication between two vehicles can be established if the distance between the two vehicles is within the V2V communication range \(R_V\). This simple model helps to reduce the computational complexity of analyzing large-scale vehicular networks to a tractable level. The V2V communication range \(R_V\) takes the value of 50 m, 100 m or 200 m in our study as recommended in [30]. Basically, the distance of 100 m is the value found in field tests as a typical reliable distance for V2V communication, and 50 m is the largest distance over which V2V communication has a packet delivery ratio close to one, while 200 m is the maximum distance for achieving V2V communication reception ratio above 0.5 [33]. According to our communication model, only bidirectional edges are present in the graph, which means that the graph \(G = (V, E, T, \rho)\) is undirected. We also assume that there is no RSU.

Because we use a real-world vehicular data trace collected by GPS devices and the data are sampled in discrete time, the time span \(T \subset \mathbb{N}\), where \(\mathbb{N}\) represents the set of natural numbers. We map the processed vehicular trace data onto \(T\) through our preprocessing method given in the next section.

We now describe the process of a vehicle acquiring computing service by initiating a local MVC. Consider the generic case that vehicle \(v_i\) at time \(t_0\) needs computing service from other vehicles. As an initiator, \(v_i\) initiates and manages an MVC by invoking all the vehicles it can communicate with and handling the data flows until the computing task is completed and the result is returned. The initiating time is denoted by \(t_0\) and the duration of the MVC is denoted by \(t_d\), which is also the delay tolerance of the task. Since both \(t_0\) and \(t_d\) are discrete, i.e., mapped onto \(\mathbb{N}\), we set \(T\) to be \(\{t_0, t_0 + 1, \ldots, t_d\}\). As soon as \(v_i\) can communicate with another vehicle \(v_j\) after \(t_0\), it will offload part of its task to \(v_j\). It does not need to be in-contact with \(v_j\) when \(v_j\) is computing the offloaded task, but it has to retrieve the result from \(v_j\) after the task is finished. Because \(t_d\) is finite, there is a limited period of time during which \(v_j\) can compute for \(v_i\) and the serviceability can be measured by the product of the computing speed and the length of maximum possible serving time of \(v_j\). Because we only consider the maximum possible serving time here, we may assume that the delivery of tasks and the retrieve of results between vehicles take little time by comparison, so that we may assume that they can be done instantaneously. Additionally, to take into account nodes’ different computing capacities, we use a weight \(p_j\) to represent the computing capacity of node \(v_j\).

The formal definitions of service and serviceability based on the TVG model are now given below.

**Definition 2 (Service):** A service in \(G\) is an interval \(s(e, q, t_1, t_2) = [t_1, t_2]\) where \(t_1, t_2 \in T\), \(t_1 < t_2\), \(e = (v_i, v_j) \in E\), \(\rho(e, q, t_1) = 1\) and \(\rho(e, q, t_2) = 1\).

Let the length of the interval \(s(e, q, t_1, t_2)\), \(L[s(e, q, t_1, t_2)]\) be the service’s lasting time and \(t_m(e, q, t_0, t_d) = \max_{t_1, t_2 \in [t_0, t_0 + t_d]} L[s(e, q, t_1, t_2)]\) be the maximum possible serving time that \(v_j\) can offer to \(v_i\) under the link condition \(q\) during the time duration of \([t_0, t_0 + t_d]\).

**Definition 3 (Serviceability):** The serviceability of an MVC initiated by node \(v_i\) during \([t_0, t_0 + t_d]\) under the link condition \(q\) is \(S_i(t_0, t_0 + t_d, q) = \sum_j p_j t_m(e = (v_i, v_j), q, t_0, t_d)\). The serviceability of the network under the link condition \(q\) during \([t_0, t_0 + t_d]\) is \(S(t_0, t_0 + t_d, q) = \sum_{v_i} S_i(t_0, t_0 + t_d, q)\), where \(|V|\) denotes the size of \(V\).

How to calculate the serviceability of an MVC initiated by node \(v_i\) is listed in Algorithm 1. In this algorithm, at the end of the outer loop, \(t_{1,n}\) and \(t_{2,n}\) represents the first time and last time when \(v_i\) can communicate with \(v_n\) during \([t_0, t_0 + t_d]\), respectively. Thus \(\sum_{n=1}^{N} p_n(t_{2,n} - t_{1,n})\) indicates the serviceability of the MVC initiated by \(v_i\) during \([t_0, t_0 + t_d]\) under the link condition \(q\). Considering that the numbers of cycles in the inner and outer loops are \(N - 1\) and \(t_d\), respectively, the complexity of the algorithm is on the order of \(t_d(N - 1)\).
Several MVCs may exist during the same time interval. In this case, the average serviceability of \( n \) MVCs initiated by \( n \) different nodes \( \{v^1, v^2, \ldots, v^n\} \) during \([t_0, t_0 + t_d]\) under the link condition \( q \) is \( S_{\text{avg}}(t_0, t_d, q) = (1/n) \sum j p_j \max t^1, t^2 \in [t_0, t_0 + t_d] L[\sum_{i=1}^n s(e = (v^i, v^j), q, t^1, t^2)] \). Furthermore, to describe the congestion degree, the coefficient of congestion denoted by \( z \) is defined as \( z = nS_{\text{avg}}(t_0, t_d, q) / \sum_{i=1}^n S_i(t_0, t_d, q) \). The larger \( z \) is, the more congested the network.

Two related features of MVCs, which influence serviceability, can now be defined.

Definition 4 (Connectivity): The connectivity \( c_i(t_0, t_d, q) \) of an MVC initiated by node \( v_i \) is defined as the average number of nodes connecting to \( v_i \) over the time duration \([t_0, t_0 + t_d]\) under the link condition \( q \). The connectivity of the network \( c(t_0, t_d, q) \) is the average of all the \( c_i(t_0, t_d, q) \) over \( v_i \in V \).

Note that the connectivity measures the average number of other nodes connecting to the initiator in a time interval, which equals the average probability of \( \rho(e = (v_i, v_j), q, t) = 1 \) in the time interval. Besides, if \( p_m = 1 \) for all \( n \), we will always have \( S_i(t_0, t_d, q) / (t_d - t_0) \geq c_i(t_0, t_d, q) \), and this is because during the time interval between the two connections when two nodes are not connected to each other, they may still carry out some computing tasks and therefore contributes to the serviceability.

Definition 5 (Mobility): The mobility \( m_i(t_0, t_d) \) of an MVC initiated by node \( v_i \) is defined as the displacement of \( v_i \) from \( t_0 \) to \( t_0 + t_d \). The mobility of the network \( m(t_0, t_d) \) is the average of all the \( m_i(t_0, t_d) \) over \( v_i \in V \).

Note that if a vehicle has moved back to the same location after \( t_d \), then the value of the mobility of this vehicle is 0, which is reasonable since this kind of mobility has no influence to the serviceability defined in our model.

The connectivity and mobility of an MVC are easier to be measured than the congestion degree, and we will investigate how these two properties influence the serviceability.

### IV. DATA TRACE AND PREPROCESSING

#### A. Data Trace

Beijing trace is the largest vehicular mobility trace available to us. To collect the data, we used the mobility track logs obtained from 27,000 participating Beijing taxis carrying GPS receivers in May 2009 with the duration of one month. Taxis are more sensitive to urban environments in terms of underlying road topology and traffic control than other types of vehicles and they have broader coverage in space and time. This is the main reason why we chose this data set. In the data set, the locations and timestamps of the moving taxis were recorded by fixed time interval. The information we specifically used includes: the taxis’ IDs, the longitude and latitude coordinates of the taxis’ locations and timestamps. There are about twenty million effective records per day from the whole data trace.

#### B. Data Processing

We obtain the taxis’ locations changing with time from the taxis’ moving traces, which are recorded by GPS devices and the coordinates are longitudes and latitudes of which the precision is 0.000001 degree. For the convenience of processing data, we convert the coordinates to meters with a precision of 1 m and set the origin point (0, 0) at (40.0° N, 116.4° E) near the center of Beijing. Since the location data were measured by the GPS devices, the noise may exist in the collected data due to the inaccuracy of GPS device. Because taxis may not report their locations at the same time slots with the same fixed frequency, we need to process the data trace to obtain the accurate locations of all the taxis in the same time slots and frequency. Thus we first use the city map of Beijing to correct the taxis’ locations so that they are all in the city’s roads, and this location adjustment is shown in Fig. 2(a).

We then use the method of interpolation to insert the location points at the time slots we need so that all the taxis have location information at every one-minute interval, which becomes the time measurement accuracy in our data processing. We explain how we carry out this interpolation, as illustrated in Fig. 2(b). Consider that we only have the location information \((x_1, y_1)\) and \((x_2, y_2)\) of a taxi at time points \( t_1 \) and \( t_2 \), respectively, but we do not have any information of the taxi between \( t_1 \) and \( t_2 \).
In order to get the location of the taxi at any time \( t \in (t_1, t_2) \), we estimate the location \((x_t, y_t)\) by the following interpolation:

\[
l_t = l_1 + \frac{t - t_1}{t_2 - t_1}(l_2 - l_1), \quad \text{where} \ l = x \ or \ y.
\]

After obtaining \((x_t, y_t)\), we again need to adjust it to be in a city’s road using the city map.

Using the preprocessing method, we obtain an instantaneous two-dimensional distribution map of the taxis’ positions for each minute. Because the sampling intervals are equal, we can map the time to the set of natural number \( \mathbb{N} \) as mentioned above. One of the maps is shown in Fig. 3, which is a view of the downtown region of Beijing. As can be seen although we consider the entire area which is more than 10 thousand square kilometers, more than 70% taxis are inside an area of 900 square kilometers, which is the downtown region of Beijing. We also get a view of the topology of main roads in Beijing through the distribution of taxis, which implicitly confirms that our preprocessing method is reasonably accurate.

After preprocessing the data, we obtain the serviceability of one or several MVCs during different times and under different link conditions using Algorithm 1. Similarly, the connectivity and mobility of MVCs can also be computed from the preprocessed data.

V. MACROSCOPIC ANALYSIS

The computing speeds of vehicles are assumed to be uniformly distributed in [0.2, 1.8] units/min. After a temporal evolution of the serviceability, we mathematically analyze the relationship between serviceability and delay tolerance and investigate the influences of node density and network congestion.

A. Temporal Evolution of Serviceability

The serviceability of MVC indicates how much computation can the initiator acquire from other vehicles in the MVC. First, we provide an intuitive study of the temporal dynamics of the serviceability of MVCs. For the sake of clarity, we specially consider the multi-hop enabled scenario with V2V range \( R_V = 100 \) m. We observe similar evolution trends among different days and thus we take the average of the evolutions over 31 days of the month. The average serviceability of MVCs as the function of the initiating time \( t_0 \) and delay tolerance \( t_d \) is depicted in Fig. 4. In order to focus on the temporal dynamics of serviceability, we further plot the average serviceability of MVCs over the duration of 24 hours, given four different delay tolerance requirements, in Fig. 5. With respect to the initiating time, we observe that from 0:00 to 7:00 in the morning, the serviceability is very small and is slightly increasing. However, the serviceability increases significantly to reach a peak value at around 9:00. The second peak of the serviceability shows up between 15:00 and 16:00 in the afternoon. After 20:00 in the evening, the serviceability decreases slowly back to the lowest level of midnight again.

More specifically, the peak value of serviceability is about twice of the value at 12:00 between the two peaks, and the lowest serviceability at midnight is only about one-tenth of the peak value in the daytime. This indicates that there exists a wide range of serviceability levels during the 24 hours of a day. In particular, there are two time periods of huge serviceability. With the delay tolerance of two hours, for example, the peak value of serviceability is about 18,000 min, equivalent to the average computation power provided by more than 150 vehicles. These temporal dynamics of the serviceability of MVCs are of course highly consistent with the behaviors of taxis. For example, the two peaks in the evolution of serviceability are reached at rush hours, and because most of the drivers are not working at night, we observe a low level of serviceability.
Fig. 6. Serviceability of MVCs averaging over different initiating times as the function of delay tolerance in six different network scenarios.

### Table I
VALUES OF $k$ AND $r$ IN SIX DIFFERENT NETWORK SCENARIOS

<table>
<thead>
<tr>
<th>Communication Protocol</th>
<th>$R_V$</th>
<th>$k$</th>
<th>$r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-hop</td>
<td>50 m</td>
<td>2.863</td>
<td>1.056</td>
</tr>
<tr>
<td></td>
<td>100 m</td>
<td>3.699</td>
<td>1.097</td>
</tr>
<tr>
<td></td>
<td>200 m</td>
<td>4.889</td>
<td>1.194</td>
</tr>
<tr>
<td>Multi-hop</td>
<td>50 m</td>
<td>4.246</td>
<td>1.108</td>
</tr>
<tr>
<td></td>
<td>100 m</td>
<td>11.291</td>
<td>1.350</td>
</tr>
<tr>
<td></td>
<td>200 m</td>
<td>2838.4</td>
<td>1.164</td>
</tr>
</tbody>
</table>

**B. Relationship Between Serviceability and Delay Tolerance**

From Fig. 4 or Fig. 5, we also notice that the serviceability is highly influenced by the delay tolerance. Hence we analyze the relationship between the serviceability $S$ and the delay tolerance $t_d$ in more details. We consider the six scenarios, where the V2V communication range is $R_V = 50$ m, 100 m or 200 m and the communication protocol is either single-hop or multi-hop enabled, respectively. By averaging the initiating time $t_0$ over the duration of 24 hours, we obtain the average serviceability of MVCs as the function of the delay tolerance $t_d$, shown in Fig. 6. It can be observed clearly that the serviceability increases as the delay tolerance increases. Furthermore, using curve fitting, we can obtain a strong linear relationship between $\ln S$ and $t_d$ with the average of the correlation coefficients greater than 0.99. Thus we obtain the fitted mathematical relationship between the serviceability $S$ and the delay tolerance $t_d$

$$S = k t_d^r,$$  \hspace{1cm} (1)

where $k$ and $r$ are the two parameters depending on the properties of the network. These two parameters enable us to estimate the average serviceability for a given delay tolerance in a specific network scenario using the model (1). Table I lists the values of $k$ and $r$ obtained for the six network scenarios.

1) **Analysis of $k$:** The most important factor that influences the value of $k$ is the connectivity of the network, which is also the underlying reason for the evolution of $S$ during a day. We obtain $c(t_0, t_d = 30$ min, $q$) and $k(t_0, t_d = 30$ min, $q$) for $t_0 \in \{0:00, 2:00, \ldots, 22:00\}$ from the Beijing trace and plot the pairs in Fig. 7 for the five different scenarios. The case of multi-hop enabled with $R_V = 200$ m is not included as the values of $k$ are very large in this case. As can be seen from Fig. 7, the relationship between $k$ and the connectivity $c$ in each scenario is closely to linear and $k$ increases as $c$ increases.

Additionally, when the connectivity $c$ is given, $k$ is different in different scenarios. Thus the V2V communication range $R_V$ and whether multi-hop communication is enabled are the secondary factors that influences $k$.

More specifically, except for the multi-hop enabled scenario with $R_V = 200$ m that is not included in Fig. 7, with the fixed connectivity value $c$, $k$ is smaller when $R_V$ is larger. This can be explained by the decrease of stability of connectivity. As the communication range $R_V$ becomes larger, the average distance of connections is larger and the probability of reconnection is smaller. This reduces the growth rate of serviceability, i.e., a smaller $k$. We can also observe that $k$ is larger when multi-hop communication is enabled, which is probably because there exist more than one possible paths between two vehicles and thus the reconnection is easier.

As for the multi-hop enabled scenario with $R_V = 200$ m, $k$ (and $c$) is more than 100 times larger than in the five other scenarios shown in Fig. 7. Because of the strong connectivity in this case, $k$ shows some different properties. In particular, if $c$ is given, the value of $k$ is larger than in the case of multi-hop enabled and $R_V = 100$ m. This can be explained reasonably. When the connectivity of the network is so strong, the decrease in the probability of reconnection caused by the increase of distance of connects becomes negligible. On the other hand, because the connectivity becomes stronger, fewer connections have no contribution to the serviceability. Therefore, $k$ increases in spite of the increase of $R_V$.

2) **Analysis of $r$:** Although $r$ slightly fluctuates during a day, we do not find obvious relationship between $r$ and connectivity or mobility of the network. However, $r$ is definitely influenced by communication protocols. Except for the multi-hop enabled scenario with $R_V = 200$ m, $r$ increases as the communication range increases. When the communication ranges are the same, $r$ is larger when multi-hop communication is enabled. The reason is that with larger communication range, a connection between two nodes has a larger probability to keep longer. As for the multi-hop enabled scenario with $R_V = 200$ m, the reason is similar to the case of $k$. Because the network is more like a wired one, new connections are hard to establish and existing connections are hard to break. Thus, $r$ is closer to 1 in this case than in the case of $R_V = 100$ m.

In summary, the average serviceability of an MVC has a relationship with the delay tolerance expressed by $S = k t_d^r$, with the two characteristic parameters $k$ and $r$. Except for the exceptional multi-hop enabled scenario with $R_V = 200$ m,
Fig. 8. Average serviceability of MVCs as the function of the node density of the network given different delay tolerance requirements in the multihop-enabled scenario with the V2V range $R_V = 100$ m.

Fig. 9. Average serviceability of MVCs as the function of delay tolerance given different node densities of the network in the multihop-enabled scenario with the V2V range $R_V = 100$ m.

$k$ increases with the increase of the network connectivity. Furthermore, the value of $k$ decreases with the increase of the V2V communication range $R_V$, while $r$ increases with the increase of $R_V$. Both $k$ and $r$ are larger when multi-hop communication is enabled.

C. Influence of Node Density

In order to investigate the influence of the vehicle or node density in a vehicular network to the achievable serviceability of the network, we randomly select a proportion of the nodes in the graph from Beijing trace and we measure the serviceability of the resulting sub-network. The results obtained for the multi-hop enabled scenario with $R_V = 100$ m are shown in Fig. 8. It can be seen from Fig. 8 that for different given delay tolerances $t_d$, the relationship between the density of nodes and the serviceability is always close to linear. Given different node densities in the network, Fig. 9 shows that the relationship between $\ln S$ and $\ln t_d$ is a linear one.

D. Influence of Network Congestion

We also investigate the coefficient of congestion and average serviceability of MVC as the functions of the number of initiators. The results are shown in Fig. 11. When multi-hop communication is enabled, we observe that both the coefficient of congestion and average serviceability decrease as the

Then we investigate the influence of node density on the two characteristic parameters $r$ and $k$. Again we consider the multi-hop enabled scenario with the V2V range $R_V = 100$ m. As the proportion of selected nodes changes from 10% to 100% under the condition of the initiating time $t_0 = 10:00$ and the delay tolerance $t_d = 120$ min, we find that the values of $r$ are restricted within a very small interval with the variance nearly $1 \times 10^{-5}$. This indicates that $r$ is independent of the node density. But we find that $k$ increases as the density of nodes increases. Fig. 10 shows that this trend is also caused by the increase of the connectivity $c$. In summary, the density of nodes does not influence $r$ but it impacts $k$. These observations are useful as they may simplify the evaluation of serviceability of a network.

Our results provide useful information to evaluate and to predict the serviceability of vehicular networks in large urban environments. We will provide a more detailed discussion in Section VII.
number of initiators increases, while the average serviceability increases and the coefficient of congestion decreases as the communication range becomes larger, especially as $R_V$ increases to the maximum value of 200 m. This suggests that when multi-hop communication is enabled, enlarging the communication range properly is beneficial both to improve serviceability and to avoid network congestion. As for the one-hop scenario, the analysis of average serviceability is similar to the multi-hop scenario. However, we find that as the communication range increases from 50 m to 200 m, the coefficient of congestion becomes larger. This means that in the one-hop scenario, enlarging the communication range improves serviceability but also increases network congestion.

VI. MICROSCOPIC ANALYSIS

We concentrate on the properties of the serviceability of an MVC by studying the spatial distribution of the serviceability of an MVC and analyzing the influencing factors. In these analyses, the computing speeds of vehicles are still set to be uniformly distributed in [0.2, 1.8] units/min.

A. Spatial Distribution

The spatial distribution of serviceability critically affects the performance of computation offloading in a vehicular network. To give a rigorous analysis, we consider six scenarios as shown in Fig. 12, where the spatial distribution of serviceability is plotted for May 6th at the initiating time $t_0 = 8:00$ with the delay tolerance $t_d = 30$ min. The area covered is the entire Beijing city, and each dot in Fig. 12 represents the serviceability obtained by a taxi. From Fig. 12, we can observe that the average serviceability of MVCs is greater when the V2V range is larger and multi-hop communication is enabled. Apart from the different levels of the average serviceability of MVCs, we observe similar spatial distributions of the serviceability of MVCs in all the scenarios except for the multi-hop enabled scenario with $R_V = 200$ m. For example, the serviceability is greater where the density of vehicles is greater, and the heterogeneity is evident even in the center of the city. However, the situation in the multi-hop enabled scenario with $R_V = 200$ m is different. In a large area around the city center, an obvious homogeneity of the serviceability of MVCs can be seen. In this area, the serviceability levels $S_i$ of different MVCs initiated by different vehicles $v_i$ are very similar. This can be explained by the strong connectivity between those nodes, which makes the network behaves more like a wired one.

The complementary cumulative distribution functions (CCDFs) of the serviceability of MVCs are portrayed in Fig. 13 for these six different scenarios. In the log-linear scale, the
distribution of serviceability has a piecewise linear approximation. More specifically, except for the multi-hop enabled scenario with $R_V = 200$ m, all the nodes can be divided into four parts in each scenario. The first part contains the nodes of which the serviceability is zero, and the serviceability in each of the other three parts has an approximately linear distribution in the log-linear scale, as can be seen from Fig. 13. Obviously, as the V2V communication range increases, the proportion of the MVCs with zero serviceability decreases and the serviceability of each MVC increases. For the multi-hop enabled scenario with $R_V = 200$ m, the distribution offers us a more intuitive feeling on the homogeneity in the central region of the network, corresponding to the dark red region of the spatial distribution plot for the multi-hop enabled scenario with $R_V = 200$ m in Fig. 12.

B. How Serviceability Related to Connectivity and Mobility

Because the serviceability is related to the connectivity, the connectivity of each MVC must be a principal factor that influences the distribution of serviceability of MVCs. Besides, there exists significant differences in the mobility between nodes, and the mobility of MVCs is also an important factor. Therefore we further analyze how these two factors influence the serviceability of MVCs.

We focus on the case of $R_V = 100$ m, and Fig. 14 shows how the serviceability related to the connectivity on May 6th at the initiating time $t_0 = 8:00$ with the delay tolerance $t_d = 30$ min, for both single-hop and multi-hop scenarios. Clearly, the serviceability of MVCs has a linear relationship with the connectivity of MVCs in the both cases but the linear relationship is much more strict in the first scenario, which is reasonable because the multi-hop scenario is much more complex. The growth rate of the serviceability approximately equals to the delay tolerance $t_d = 30$ min, and only about 5% of the MVCs’ serviceability are more than 5% greater than $t_d \cdot c_i$, where $c_i$ denotes the connectivity of the MVC initiated by $v_i$. Moreover, as shown in Fig. 15, if we only consider the distribution of the serviceability of the MVCs for which $c_i \neq 0$ and $c_i$ are restricted in a narrow range by eliminating other MVCs whose levels of connectivity are outside the range, the whole distribution is almost linear in the log-linear scale.

Fig. 16 shows the influence of mobility on serviceability, where the MVCs are classified by mobility. The most important observation is that as the mobility $m_i$ of the MVC initiated by $v_i$ increases from 0–100 m to 1000+ m, the probability of $S_i = 0$ falls from 0.29 to 0.06 when multi-hop communication is not enabled, and it falls from 0.27 to 0.04 when multi-hop communication is enabled. We also observe that the variation of the serviceability of MVCs is smaller when the mobility is larger, which means that higher mobility causes lower predictability. Although the influence of mobility on serviceability is smaller than the influence of connectivity, it should still be taken into consideration when we design architectures and applications or when vehicles make decisions on computation offloading.

In summary, the increases of connectivity and mobility of an MVC both strengthen its serviceability, although when the mobility increases, the serviceability becomes more unstable. The analysis of this subsection provides useful information to help a vehicle when deciding whether it should offload its computation in a specific state restricted by time, location, communication condition, vehicular mobility and density.

VII. DISCUSSIONS

A. Versatility

Due to the rareness of available large-scale urban vehicular mobility traces, our analysis is limited to Beijing trace only,
which is extracted from GPS devices on taxis. Although taxis have broad coverage in space and time and they are sensitive to urban environment in terms of underlying road topology, they are just part of the whole traffic in the city. Thus the influence of vehicle density on the serviceability of a generic vehicular network must be studied. In Section V, we have specifically considered the influence of taxi density on the serviceability of the network. The results show that one of the characteristic parameters $k$ is related to the density of taxis. We also observe a linear relationship between the density of taxis and the serviceability when the density is not too small. Intuitively, the density of taxis in Beijing city is correlated with the density of vehicles in Beijing. This linear relationship between the serviceability and the density of taxis extracted from Beijing trace can be applied to the generic vehicular network of Beijing involving all types of vehicles. Furthermore, if a city has similar size and road topology to Beijing, our results can also be applied.

For cities which do not have similar size to Beijing or where the road topology is very different, the specific statistical values obtained in this study may not be valid. However, the statistical model proposed in this study, which defines a relationship with the delay tolerance of a task, which can be described by the two characteristic parameters. The serviceability of MVCs is also investigated by considering the influence of network congestion. In our microscopic analysis, the spatial distribution of the serviceability has been analyzed. Moreover, we have analyzed the influences of connectivity and mobility on the serviceability in both macroscopic and microscopic views. The connectivity is shown to have a linear relationship with the serviceability. The mobility is seen to have a smaller influence on the serviceability than the connectivity but it also provides useful information. According to our analysis, the serviceability of MVCs can be estimated and predicted.

Future work is warranted to investigate the serviceability of MVCs in other big cities around the world in order to further verify the findings and conclusions of this study. Another research problem worthy of investigating is the influence of the speed of vehicles on the serviceability, which is not directly studied in our work. Besides, the influence of network congestion should be thoroughly investigated. Ultimately, a real system to offload computation between vehicles should be demonstrated and tested in the future to allow researchers to collect real-life data for studying the properties of MVCs.

B. Significance

We can utilize the insights and conclusions obtained in this study in the design of vehicular cloud computing architectures and applications. Specifically, consider the system design of vehicular cloud computing. According to our analysis, the serviceability of MVCs is highly influenced by the V2V communication range. For example, in Beijing, the range of 100 m performs well in most areas. However, the performance in some areas is limited due to the sparse connectivity. Then, we can arrange infrastructures such as RSUs in these areas to improve the connectivity and thus to enhance the serviceability by providing V2I communication. When designing protocols, the multi-hop communication should be considered, according our findings. If multi-hop communication is not allowed in MVCs, the serviceability may be severely limited, as clearly demonstrated in this study.

Whether an application should provide an offloading option and how to offload the tasks to the MVC always require careful considerations by the designer, and our obtained insights in the serviceability of MVCs can help to design better applications. Drivers can also benefit from our work. Because the serviceability depends on specific environment, the choices of offloading a task can be varying in different scenarios. With the knowledge of the level of serviceability in a city, the drivers can make better decisions. We also suggest that an automatic analyzing software should be designed based on the model we proposed to assist drivers making choices.

VIII. CONCLUSION AND FUTURE WORK

We have assessed the serviceability of MVCs by studying the taxi mobility trace of Beijing and have built a model for the vehicular network under different scenarios in such a large-scale urban environment. In our macroscopic analysis, the evolution of serviceability is observed, and the serviceability is shown to have a relationship with the delay tolerance of a task, which can be described by the two characteristic parameters. The serviceability of MVCs is also investigated by considering the influence of network congestion. In our microscopic analysis, the spatial distribution of the serviceability has been analyzed. Moreover, we have analyzed the influences of connectivity and mobility on the serviceability in both macroscopic and microscopic views. The connectivity is shown to have a linear relationship with the serviceability. The mobility is seen to have a smaller influence on the serviceability than the connectivity but it also provides useful information. According to our analysis, the serviceability of MVCs can be estimated and predicted.

Future work is warranted to investigate the serviceability of MVCs in other big cities around the world in order to further verify the findings and conclusions of this study. Another research problem worthy of investigating is the influence of the speed of vehicles on the serviceability, which is not directly studied in our work. Besides, the influence of network congestion should be thoroughly investigated. Ultimately, a real system to offload computation between vehicles should be demonstrated and tested in the future to allow researchers to collect real-life data for studying the properties of MVCs.

REFERENCES

Chuanmeizhi Wang received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2015. His research interests include vehicular communications and networking.

Yong Li (M’09) received the B.S. degree in electronics and information engineering from Huazhong University of Science and Technology, Wuhan, China, in 2007 and the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2012. From July to August 2012 and 2013, he was a Visiting Research Associate with Telekom Innovation Laboratories and The Hong Kong University of Science and Technology, Kowloon, Hong Kong, respectively. During December 2013 to March 2014, he was a Visiting Scientist with University of Miami, Coral Gables, FL, USA. He is currently a Faculty Member with the Department of Electronic Engineering, Tsinghua University. His research interests are in the areas of networking and communications, including mobile opportunistic networks, device-to-device communication, software-defined networks, network virtualization, and future Internet.

Dr. Li has served as a Technical Program Committee (TPC) Chair for the WWoW workshop of Simplex 2013 and the TPC member for several international workshops and conferences. He is a Guest Editor for ACM/Springer Mobile Networks and Applications, Special Issue on Software-Defined and Virtualized Future Wireless Networks. He is currently the Associate Editor of EURASIP Journal on Wireless Communications and Networking. He was a recipient of the Outstanding Postdoctoral Researcher, Outstanding Ph.D. Graduates, and Outstanding Doctoral Thesis awards from Tsinghua University.

Depeng Jin (M’08) received the B.S. and Ph.D. degrees in electronics engineering from Tsinghua University, Beijing, China, in 1995 and 1999, respectively. He is an Associate Professor with Tsinghua University and Vice Chair with the Department of Electronic Engineering. His research fields include telecommunications, high-speed networks, ASIC design, and future Internet architecture.

Dr. Jin was the recipient of the National Scientific and Technological Innovation Prize (Second Class) in 2002.