Delay Constrained Offloading for Mobile Edge Computing in Cloud-enabled Vehicular Networks

Ke Zhang, Yuming Mao and Supeng Leng
School of Communication and Information Engineering,
University of Electronic Science and Technology of China
Email: {zhangke,ymmao,spleng}@uestc.edu.cn

Alexey Vinel
Halmstad University
Halmstad, Sweden
Email: alexey.vinel@hh.se

Yan Zhang
Simula Research Laboratory
Oslo, Norway
Email: yanzhang@ieee.org

***Abstract*—Cloud-based vehicular networks is a new paradigm to improve the vehicular services through distributing computation tasks between remote clouds and local vehicular terminals. To further reduce the latency and the transmission cost of the computation offloading, we propose a cloud-based Mobile Edge Computing (MEC) offloading framework in vehicular networks. In the framework, efficient computation offloading strategies are designed through a contract theoretic approach. We obtain the optimal feasible contracts that maximize the benefit of the MEC service provider while enhancing the utilities of the vehicles. Furthermore, considering the resource limitation of the MEC server and the latency tolerance of the computation tasks, we propose a contract-based computation resource allocation scheme. Numerical results show that our proposed scheme greatly enhances the utility of the MEC service provider.**

***Index Terms*—Vehicular networks; cloud; mobile edge computing; contract; offloading.**

I. INTRODUCTION

With the ever-increasing number of vehicles on the roads and the development of the Internet of Things (IoT), the vehicular terminals have been an important part of the mobile computing devices [1]. To cope with the explosive application demands of the vehicular terminals, the cloud-based vehicular networking is widely considered as a new paradigm to improve the service performance [2]-[5]. In the cloud-enabled network, by integrating both the communication and computing technologies, applications can either run locally on the vehicular terminals or offload to the remote computation cloud [6]-[8].

The mobile cloud computing greatly improves the resource utilization and computation performance. However, considering the capacity limitation and delay fluctuation of the transmission on the backhaul and backbone networks, the placement of the cloud servers far away from the mobile vehicles may cause serious degradation of the offloading efficiency. To this end, Mobile Edge Computing (MEC) is proposed as a promising solution, which pushes the cloud service to the edge of the radio network, and provides a cloud-based computation offloading in close proximity to the mobile vehicular terminals [9].

Due to the characteristic of proximity, the MEC server is able to provide fast interactive response in the computation offloading service. However, compared with the traditional cloud server locating at the backbone network, which always has enormous computation resources, the MEC server suffers from the resource limitation. Furthermore, as the MEC server operates at the edge of radio access networks, its service area may be limited by the radio coverage of the base station. Thus, an efficient computation offloading scheme for the MEC computing is needed.

Recently, there are a few studies on the cloud-enabled vehicular networks. In [10], the authors proposed a coalition game model to manage and share the resources among different cloud service providers. In [11], the authors exploited the cognitive radio and soft data fusion in vehicular access networks, and designed a distributed traffic offloading scheme for cognitive cloud vehicular networks. By combining the vehicular cloud with the fixed central cloud, the authors in [12] proposed a flexible offloading strategy to discover unutilized resources and carry out task mitigation. To investigate the MEC computing mechanisms, the authors in [13] studied a multi-user MEC computation offloading problem in a multi-channel wireless environment, and designed a distributed game theoretic offloading scheme. The authors in [14] made a case study, where 5G and MEC are collaborated to provide the real-time context-aware communication. In [15], the authors compared the performance of an infrastructure cloud with that of a mobile edge-cloud in terms of the application running time. However, few of these work has considered the impacts of the vehicle mobility on the computation offloading performance. In addition, the utility of the MEC service provider obtained from offloading various types of computation tasks has been ignored in these work.

In this paper, we propose a delay constrained MEC offloading mechanism in a cloud-based vehicular network. The main contributions of this paper are listed as below:

- We propose a cloud-based MEC offloading framework in vehicular networks, where both the MEC service limitation and the mobility of the vehicles are considered.
- We design a contract-based computation offloading scheme, which maximizes the benefit of the MEC service provider and meanwhile enhances the utility of the vehicular terminals.
- Based on the obtained optimal contracts, we propose an

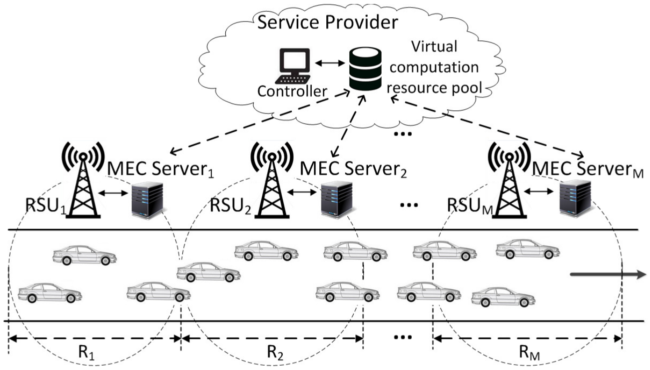

Fig. 1. The computation offloading in a cloud-based vehicular network.

efficient computation resource allocation scheme. This scheme improves the resource utilization under the delay constraints of the computation tasks.

The rest of this paper is organized as follows. The system model is presented in Section II. In Section III, we formulate the MEC computation offloading as a contract problem and obtain the optimal contracts. Based on these contracts, in Section IV, we propose a delay constrained computation resource allocation scheme. We present numerical results in Section V and conclude the paper in Section VI.

II. System Model

Figure 1 shows our proposed MEC offloading for cloud-enabled vehicular networks. We consider an unidirectional road, where M Road Side Units (RSUs) locate long the road. Each RSU is equipped with a MEC server, whose computation resources are limited. We denote the id set of these MEC servers as $\mathcal{M} = \{1, 2, ..., M\}$. The computation resources for MEC server m is b_m units, $m \in \mathcal{M}$. All the RSUs and the MEC servers along the road belong to the same service provider. The schedule of the computation task offloading for the vehicular terminals is under the management of the provider.

Due to the variety of the radio power and the wireless environment, the RSUs may have different wireless coverage areas [16]. Each vehicle access to the RSU with the strongest signal. Thus, the road can be divided into M segments, whose length are $\{R_1, R_2, ..., R_M\}$, respectively. The vehicles running within the mth segment can only access to RSU m and offload their tasks to MEC server m.

We consider that there are Q vehicles arriving at the starting end of the road, and they move at a constant speed h. Each vehicle has a computation task, which is denoted as $T = \{d, t^{max}\}$. Here d is the amount of the computation resources required to accomplish the task, and t^{max} is the maximum latency of the task. Each computation task can be accomplished either locally on the vehicle or remotely on the MEC servers through task offloading.

We assume each vehicle has a homogeneous computation resource, which is denoted as d_0. In our model, there are N types of computation tasks of these vehicles, whose resource requirements are denoted as $\{d_1, d_2, ..., d_N\}$, respectively. Thus, the tasks of the vehicular terminals can be presented as $T_i = \{d_i, t_i^{max}\}$, $i \in \mathcal{N} = \{1, 2, ..., N\}$. Without loss of generality, we assume that $d_1 < d_2 < ... < d_N$. As the time cost for performing the task T_i is proportional to the resource requirement d_i, we further assume that $t_1^{max} < t_2^{max} < ... < t_N^{max}$.

According to their computation task types, these vehicles can be correspondingly classified into N types. Let γ_i be the proportion of vehicles with task T_i in all the arriving vehicles, and $\sum_{i=1}^{N} \gamma_i = 1$. Each vehicle knows its own type. However, the computation offloading service provider may not be well aware of the types of the vehicles. Thus, we can see that an information asymmetry has occurred between the service provider and the vehicles. Although the service provider can not accurately determine the type of each vehicle, it can obtain the probability distribution of the vehicle types through some statistical information.

Both the service provider and the vehicles are rational and self-interested. Each vehicle can offload some computation tasks to the MEC servers with payment to the service provider. Thus, to maximize its revenue, the service provider derives the optimal amount of the computation resources allocated to offload the tasks and determines the corresponding payment. The information of these services is broadcasted to the vehicles running on the road in contract forms via the wireless communication. According to the contract information, each vehicle implements the task offloading in the way that maximizes its utility.

III. Contract Theoretic Computation Offloading

In this section, we first formulate the computation offloading process as a contract problem. Then we derive the optimal contract solutions, which maximize the utility of the MEC service provider while satisfying the requirement of the vehicles.

A. Contract Problem Formulation

As there are N types of vehicles according to their computation tasks, the provider needs to offer N kinds of contracts correspondingly. Let (q_i, p_i) denote the contract designed for the vehicle belonging to type i ($i \in \mathcal{N}$). Here q_i is the amount of the computation resources provided by the MEC servers for offloading the computation task from a type-i vehicle. p_i is the payment that the type-i vehicle should pay to the provider for using the offloading service.

Each vehicle can obtain the contract information through the wireless broadcast from the provider. After that, the vehicle chooses to accept the contract that brings maximal utility to it. Here we define the utility of a type-i vehicle gained from offloading its task based on the contract (q_i, p_i) as

$$\begin{aligned} &U_v^i(q_i, p_i) = \lambda \tfrac{d_i}{d_0} \ln(\alpha q_i + \beta) + (1-\lambda) e_0 q_i - p_i, \\ &q_i \le d_i, \quad i \in \mathcal{N}. \end{aligned} \tag{1}$$

In (1), the first item represents the utility of saving computation resource by the task offloading. The utility is affected by the original resource utilization d_i/d_0. α and β are coefficients. In the second item, e_0 is the utility gained from saving energy for running a task on a remote unit computation resource. λ is the weight factor, and $0<\lambda<1$. The inequality $q_i \leq d_i$ ensures that the amount of the resources defined by the contract is less or equal than the total resources required by the task. Let $G(d_i,q_i)=\lambda\frac{d_i}{d_0}\ln(\alpha q_i+\beta)+(1-\lambda)e_0 q_i$. Then, (1) can be rewritten as

$$U_v^i(q_i,p_i)=G(d_i,q_i)-p_i. \tag{2}$$

As the vehicles are rational, they will not choose the contract which brings negative utilities to them. This property is called Individual Rationality (IR), and can be mathematically expressed as $U_v^i \geq 0$ $(i\in\mathcal{N})$. Besides this property, according to the contract theory, a feasible contract should satisfy the Incentive Compatible (IC) constraint [17]. This constraint indicates that the type-i vehicles are incentivized to choose the contract specifically designed for their own type, but not the contracts of the others. The IC constraint can be represented as $U_v^i(q_i,p_i)\geq U_v^i(q_j,p_j)$, $i\neq j$, $i,j\in\mathcal{N}$.

The utility of the provider providing computation task offloading service based on these contracts is defined as

$$U_{SP}=Q\sum_{i=1}^{N}\gamma_i(p_i-cq_i), \tag{3}$$

where c is the cost for the MEC servers running a computation task on an unit resource.

Based on the utility functions of the vehicles and the service provider while considering the IR and IC constraints, the contract theoretic optimization problem of the computation offloading can be formulated as follows,

$$\begin{array}{ll} \max\limits_{\{p_n,q_n\}} U'_{SP}=\sum\limits_{n=1}^{N}\gamma_n(p_n-cq_n) & \\ s.t.\quad C1: G(d_i,q_i)-p_i\geq 0, i\in\mathcal{N} & \\ \qquad\quad C2: G(d_i,q_i)-p_i\geq G(d_i,q_j)-p_j & . \\ \qquad\qquad\quad i\neq j,\quad i,j\in\mathcal{N} & \\ \qquad\quad C3: 0\leq q_i\leq d_i,\quad i\in\mathcal{N} & \end{array} \tag{4}$$

It is noteworthy that, compared with (3), Q is omitted in (4), as this parameter does not affect the design of the contracts.

B. Optimal Contract Design

As there are N^2+N constraints, it may be complex to solve (4), especially given a large N. Thus, to solve (4) more efficiently, some constraints should be removed. In this subsection, we first present the steps to simplify the optimization problem. Then, we propose an efficient algorithm to get the optimal contract solutions.

Lemma 1. *Monotonicity: Both the computation resources q_i and the offloading payment p_i of the feasible contract $\{q_i,p_i\}$ designed for the type-i vehicles are monotonically increasing in term of i, $i\in\mathcal{N}$, i.e., for contracts $\{q_i,p_i\}$ and $\{q_j,p_j\}$, we have $q_i\geq q_j$ and $p_i\geq p_j$, if and only if $i>j$, $i,j\in\mathcal{N}$.*

Proof: According to the IC constraints of the type-i and the type-j vehicles, where $i\neq j$ and $i,j\in\mathcal{N}$, we have

$$G(d_i,q_i)-p_i\geq G(d_i,q_j)-p_j, \tag{5}$$

and

$$G(d_j,q_j)-p_j\geq G(d_j,q_i)-p_i. \tag{6}$$

Adding (5) and (6), we can get $G(d_i,q_i)+G(d_j,q_j)\geq G(d_i,q_j)+G(d_j,q_i)$. By substituting $G(\cdot,\cdot)$ with its definition, the inequality is equally changed to $(d_i-d_j)\ln(\frac{\alpha q_i+\beta}{\alpha q_j+\beta})\geq 0$. Given $i>j$, according to the vehicle type definition, we have $d_i>d_j$. Thus, we can draw the conclusion that between the feasible contracts, $q_i\geq q_j$ when $i>j$. Next we will prove that $q_i\geq q_j$ should be hold whenever $p_i\geq p_j$. Given the condition $p_i\geq p_j$, according to (5), we have $\lambda\frac{d_i}{d_0}\ln(\alpha q_i+\beta)+(1-\lambda)e_0q_i-p_i\geq\lambda\frac{d_i}{d_0}\ln(\alpha q_j+\beta)+(1-\lambda)e_0q_j-p_j$. By rearranging the inequality, we can get $\lambda\frac{d_i}{d_0}\ln(\frac{\alpha q_i+\beta}{\alpha q_j+\beta})+(1-\lambda)e_0(q_i-q_j)\geq p_i-p_j\geq 0$. Due to $0<\lambda<1$, we can get $q_i\geq q_j$. Similarly, the inequality $p_i\geq p_j$ can be proved under the condition $q_i\geq q_j$. ■

Lemma 2. *IR Constraints Reduction: If the IR constraint of the type-1 vehicles is satisfied, then other IR constraints of the type-i, $1<i\leq N$ are automatically hold [17].*

Proof: Based on the IC constraints, we can get $G(d_i,q_i)-p_i>G(d_i,q_1)-p_1$. Furthermore, according to the definition of the vehicle types, we have $d_i>d_1$, where $1<i\leq N$. As $G(d_i,q_i)$ is a monotonically increasing function in terms of d_i, we get $G(d_i,q_1)-p_1>G(d_1,q_1)-p_1$. The inequality indicates that the IR constraint of the type-i vehicles, $1<i\leq N$, is automatically satisfied whenever IR constraint of type-1 vehicles is hold. ■

Definition 1. *Local Downward Incentive Constraint (LDIC) and Downward Incentive Constraint (DIC): Considering two adjacent types, namely the type-i and the type-$(i-1)$, the IC constraint of the contracts between these two types is defined as LDIC, which can be formally presented as*

$$U_v^i(q_i,p_i)\geq U_v^i(q_{i-1},p_{i-1}),\quad i\in\{2,3,...,N\}. \tag{7}$$

By extending the definition of LDIC to the IC constraints between the the type-i and the type-j, where $j\in\{1,...,i-1\}$, we have the DICs, which are shown as

$$U_v^i(q_i,p_i)\geq U_v^i(q_j,p_j),\quad i\in\{2,3,...,N\}, j\in\{1,...,i-1\}. \tag{8}$$

Definition 2. *Upward Incentive Constraint (UIC) and Local Upward Incentive Constraint (LUIC): Similar to **Definition 1**, the definition of UIC and LUIC are formally given as $U_v^i(q_i,p_i)\geq U_v^i(q_j,p_j)$, where $i\in\{1,2,...,N-1\}, j\in\{i+1,...,N\}$, and $U_v^i(q_i,p_i)\geq U_v^i(q_{i+1},p_{i+1})$, where $i\in\{1,2,...,N-1\}$, respectively.*

Lemma 3. *Given the LDICs are hold, all the DICs are automatically hold and can be reduced. Similarly, under the condition that LUIC are hold, all the UICs can be removed.*

Proof: We give the proof of the automatically holding characteristic of the DICs as follows. According to the definition of $G(\cdot,\cdot)$, it is easy to get $G(d_{i+1},q_i)-G(d_{i+1},q_{i-1})\geq G(d_i,q_i)-G(d_i,q_{i-1})$. Given the condition that LDIC is hold, according to (7), we have $G(d_i,q_i)-G(d_i,q_{i-1}) \geq p_i - p_{i-1}$. From these two inequalities, we get

$$G(d_{i+1},q_i)-G(d_{i+1},q_{i-1}) \geq p_i - p_{i-1}. \tag{9}$$

Based on (9) and the LDIC definition between the type-$(i+1)$ and the type-i contracts, we have

$$\begin{aligned} G(d_{i+1},q_{i+1}) - p_{i+1} &\geq G(d_{i+1},q_i) - p_i \\ &\geq G(d_{i+1},q_{i-1}) - p_{i-1}. \end{aligned} \tag{10}$$

The inequality (10) could be extended to prove that all the DICs are hold. This proof can be shown as

$$\begin{aligned} G(d_{i+1},q_{i+1}) - p_{i+1} &\geq G(d_{i+1},q_{i-1}) - p_{i-1} \geq ... \\ &\geq G(d_{i+1},q_1) - p_1, \quad i \in \{1,2,...,N-1\}. \end{aligned} \tag{11}$$

Hence, we come to the conclusion that all the DICs are hold and can be reduced. The feature that all the UICs are hold can be proved in the similar way. We ignore the detailed proof here. ■

Lemma 4. *All the LDICs are binding at the optimal contracts obtained from problem* (4).

Proof: For the contract designed for the type-i vehicles, a LDIC is not binding, when $G(d_i,q_i)-p_i > G(d_i,q_{i-1}) - p_{i-1}$. Under this condition, the service provider can adapt the contract by raising all p_j $(j \geq i)$ to make the LDIC binding. This method can improve the maximum utility of the provider while not affecting the LDICs for the contracts of the other types of vehicles. ■

The fact that all the LDICs are binding for the optimal contracts, together with the monotonicity proved in Lemma 1, leads all the LUICs to be satisfied [18]. Thus, we can rewrite the optimization problem (4) as

$$\begin{aligned} &\max_{\{p_n,q_n\}} U'_{SP} = \sum_{n=1}^{N} \gamma_n (p_n - cq_n) \\ &s.t. \quad C1: G(d_1,q_1) - p_1 = 0 \\ &\qquad\ \ C2: G(d_i,q_i) - p_i = G(d_i,q_{i-1}) - p_{i-1}, 1 < i \leq N \ . \\ &\qquad\ \ C3: 0 \leq q_1 \leq q_2 \leq ... \leq q_N \\ &\qquad\ \ C4: 0 \leq q_i \leq d_i, \quad i \in \mathcal{N} \\ &\qquad\ \ C5: p_i \geq 0, \quad i \in \mathcal{N} \end{aligned} \tag{12}$$

Let $\Delta_k = G(d_k,q_k) - G(d_k,q_{k-1})$, where $1 < k \leq N$ and $\Delta_1 = 0$. According to the constraints C1 and C2 in (12), the p_i can be presented as

$$p_n = G(d_1,q_1) + \sum_{k=1}^{n} \Delta_k, \quad n \in \mathcal{N}. \tag{13}$$

By replacing p_n in (12) with the equation of (13), the objective function is given as

$$\begin{aligned} \max_{\{q_n\}} U'_{SP} &= \sum_{n=1}^{N} \gamma_n (G(d_1,q_1) + \sum_{k=1}^{n} \Delta_k - cq_n) \\ &= G(d_1,q_1) \sum_{n=1}^{N} \gamma_n - G(d_2,q_1) \sum_{n=2}^{N} \gamma_n \\ &\quad + G(d_2,q_2) \sum_{n=2}^{N} \gamma_n - G(d_3,q_2) \sum_{n=3}^{N} \gamma_n + ... \qquad . \\ &\quad + G(d_{N-1},q_{N-1}) \sum_{n=N-1}^{N} \gamma_n - G(d_N,q_{N-1}) \gamma_N \\ &\quad + G(d_N,q_N) \gamma_N + c \sum_{n=1}^{N} \gamma_n q_n \end{aligned} \tag{14}$$

Then the optimization problem of (12) is equally rewritten as follows, where the value of the objective function is only determined by the variable set $\{q_n\}$.

$$\begin{aligned} &\max_{\{q_n\}} U'_{SP} = \sum_{n=1}^{N} \{ G(d_n,q_n) \sum_{i=n}^{N} \gamma_i \\ &\qquad\qquad - G(d_{n+1},q_n) \sum_{j=n+1}^{N} \gamma_j - c\gamma_n q_n \} \ , \\ &s.t. \quad C1: 0 \leq q_1 \leq q_2 \leq ... \leq q_N \\ &\qquad\ \ C2: 0 \leq q_i \leq d_i, \quad i \in \mathcal{N} \end{aligned} \tag{15}$$

Let $S_n = G(d_n,q_n) \sum_{i=n}^{N} \gamma_i - G(d_{n+1},q_n) \sum_{j=n+1}^{N} \gamma_j - c\gamma_n q_n$. The objective function of (15) can be presented as $U'_{SP} = \sum_{n=1}^{N} S_n$. As S_n is independent from S_i, $n \neq i$, and S_n is only related to the amount of computation resources q_n, the optimal $\{q_n^*\}$ which maximizes U'_{SP} can be obtained separately by letting $q_n^* = \arg\max_{q_n} S_n$. We have $d^2 S_n / dq_n^2 = ((d_{n+1} - d_n) \sum_{i=n+1}^{N} \gamma_i - d_n \gamma_n) \alpha^2 \lambda / (\alpha q_n + \beta)^2 d_0$. Thus, given the condition that $(d_{n+1} - d_n)/d_n > \gamma_n / \sum_{i=n+1}^{N} \gamma_i$, we can get $d^2 S_n / dq_n^2 > 0$, which indicates that the S_n is a concave function. According to the Fermat's theorem, q_n^* can be derived by solving the equation $dS_n/dq_n|_{q_n = q_n^*} = 0$. Considering the constraint C2 in (15), if the obtained q_n^* is a negative number or exceeds d_n, the q_n^* should be set as 0 or d_n. Here $q_n^* = 0$ means that the contract is set to $\{0, \mathtt{Na}\}$. According to this contract, there is no computation offloading between the type-n vehicles and the MEC servers.

Besides the constraint C2, the obtained optimal q_n^* should satisfy the constraint C1 of (15). As each q_n^* is derived separately from the corresponding S_n, there may exist some sub-sequences not following the increasing order, which is described in the constraint C1. Noting that when $\{Sn\}$ are concave functions, the problem of these infeasible sub-sequences can be solved by an iteratively substitution algorithm, which is given as Algorithm 1 [19].

Based on the derived feasible optimal computation resources $\{q_n^*\}$, we can get the corresponding payments $\{p_n^*\}$ through (13). Thus, we can obtain the optimal contract set $\{q_n^*, p_n^*\}$ under the condition that $\{S_n\}$ are concave functions. If the concave condition is not satisfied, we can first solve the

Algorithm 1 The substitution algorithm for the infeasible sub-sequences

Initialization: Let $q_n^* = \arg\max_{q_n} S_n$, $n \in \mathcal{N}$.

while The set $\{q_n^*\}$ is not in the increasing order, **do**
  In the set $\{q_n^*\}$, search for the infeasible sub-sequence $\{q_i^*, q_{i+1}^*, ..., q_j^*\}$, where $q_i^* \geq q_{i+1}^* \geq ... \geq q_j^*$, $1 \leq i < j \leq N$;
  Set $q_k^* = \arg\max_{\{q\}} \sum_{x=i}^{j} S_x(q)$, $k \in \{i, i+1, ..., j\}$;
end while
return The feasible set $\{q_n^*\}$, $n \in \mathcal{N}$.

optimization problem (12) without the constraint C3 by Lagrange multiplier. Then, we can check whether the solution to this relaxed problem satisfies the constraint C3. The solving process by applying Lagrange multiplier is omitted here.

IV. Computation Resource Allocation Schemes for the Task Offloading

In this section, based on the obtained optimal contracts in Section III, we propose an efficient computation resource allocation scheme. This scheme aims to optimize the resource utilization among the MEC servers. In this scheme, both the MEC resource limitation and the latency requirements of the tasks are considered.

In a practical scenario, the sum of the task offloading demands from the arriving vehicles is a variable. The resource requirement of the demands may exceed the total computation resource of the MEC servers. Thus, a practical implementation of the contract-based offloading scheme is required.

Theorem 1. *According to the optimal contracts $\{q_n^*, p_n^*\}$, the utility gained by the service provider from offering an unit resource to serve a lower type vehicle is more than that from a higher type one.*

Proof: Let $u(n)$ denote the utility of the service provider, which is gained from offering an unit resource to offload the task of a type-n vehicle. Here, we have $u(n) = (p_n - cq_n)/q_n$. Thus, we get the difference between $u(n+1)$ and $u(n)$ as

$$D = u(n+1) - u(n) = p_{n+1}/q_{n+1} - p_n/q_n. \tag{16}$$

According to the definition of p_n in (13), we have $q_n p_{n+1} - q_{n+1} p_n = q_n(G(d_1, q_1) + \sum_{k=1}^{n+1} \Delta_k) - q_{n+1}(G(d_1, q_1) + \sum_{k=1}^{n} \Delta_k) = (q_n - q_{n+1})(G(d_1, q_1) + \sum_{k=1}^{n} \Delta_k) - q_n \Delta_{n+1}$. Due to the monotonicity proved in Lemma 1, we have $q_n < q_{n+1}$. Thus, we get $\Delta_k = G(d_k, q_k) - G(d_k, q_{k-1}) = \lambda \frac{d_k}{d_0} \ln(\frac{\alpha q_k + \beta}{\alpha q_{k-1} + \beta}) + (1 - \lambda) e_0 (q_k - q_{k-1}) > 0$, $k = \{1, 2, ..., n, n+1\}$. Then, we can find $q_n p_{n+1} < q_{n+1} p_n$, and come to the conclusion that $u(n+1) < u(n)$. ■

Theorem 1 indicates that the vehicles of lower type should be served with a higher priority, so as to make the limited resource more profitable. Besides the priority, the delay tolerance of each task should also be considered in the computation resource allocation. When a vehicle moves along the road, it may access to different RSUs at different times. For each vehicle i, there exists an RSU. This RSU is the last one that vehicle i can access to within the maximum latency of task T_i. As each RSU is equipped with a MEC server, there has a last MEC server for task T_i. We define the id of the last MEC server for task T_i as the maximum id $m_i^{\max}$. Since all the vehicles move at a constant speed h, we can get $m_i^{\max}$ as

$$m_i^{\max} = \arg\min_k \sum_{m=1}^{k} R_m \geq t_i^{\max} h, \quad i \in \mathcal{N}. \tag{17}$$

The obtained m_i^{max} can be used to determine the range of the available resources for offloading computation task T_i. Based on the $\{m_i^{max}\}$ and the priorities of each type of vehicles, the MEC computation resources allocation algorithm is described in Algorithm 2. This algorithm aims at maximizing the utility of the service provider while ensuring the delay constraints of the tasks. The computation complexity of the algorithm is given as $O(\sum_{i=1}^{N} m_i^{\max}) \leq O(NM)$.

Algorithm 2 The contract-based computation resource allocation algorithm

Initialization: The number of the arriving vehicles Q; The computation task $T_i = \{d_i, t_i^{max}\}$, $i \in \mathcal{N}$; The available computation resource $\{b_m\}$ for MEC server m, $m \in \mathcal{M}$.

Derive the optimal contract set $\{q_i^*, p_i^*\}$ ($i \in \mathcal{N}$) following the steps described in Section III;
For $i = 1 : 1 : N$ **do**
  Obtain the maximum id m_i^{max} according to (17);
  Compute the available computation resources for the type-i tasks as $B_i = \sum_{m=1}^{m_i^{\max}} b_m$;
  Get the maximum number of the served type-i vehicles as $a_i^{max} = \lfloor B_i / q_i^{\max} \rfloor$, where $\lfloor \cdot \rfloor$ is the floor function;
  Obtain the actual served type-i vehicles as $a_i = \min\{a_i^{max}, \gamma_i Q\}$;
  Allocate the computation resource from the set $\{b_m\}$ ($1 \leq m \leq m_i^{\max}$) to the a_i vehicles;
  Update the remain available resources of these MEC servers as $b_m = b_m^{remain}$ ($1 \leq m \leq m_i^{\max}$).
End For

V. Numerical Results

In this section, we evaluate the proposed contract-based computation offloading schemes. We consider a scenario where $M = 6$ RSUs along the road have the available computation resources $\{33, 18, 10, 6, 17, 25\}$ units, respectively. For each RSU, the cost for computing on an unit resource is set $c = 0.3$. The arriving vehicles are classified into $N = 5$ types in terms of their computation tasks with the probabilities $\gamma = \{0.17, 0.23, 0.28, 0.18, 0.16\}$. For each type of vehicular computation tasks, the resource requirement is $d = \{12, 13, 15, 18, 21\}$, respectively.

Figure 2 evaluates the performance of the proposed contract-based computation offloading scheme. In the resource allocation process of this scheme, both the delay tolerance and the priorities of different types of vehicles are considered.

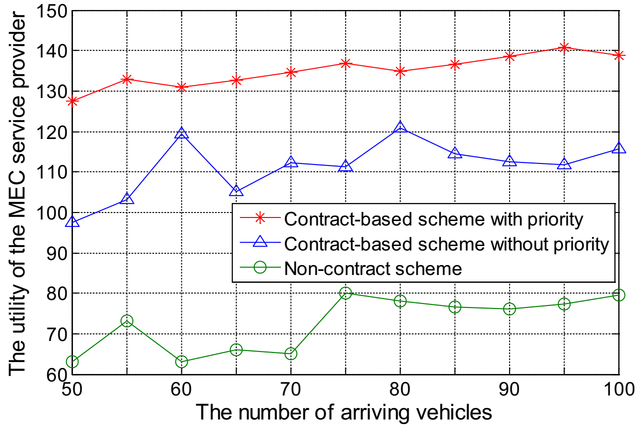

Fig. 2. The utilities of the service provider with different schemes.

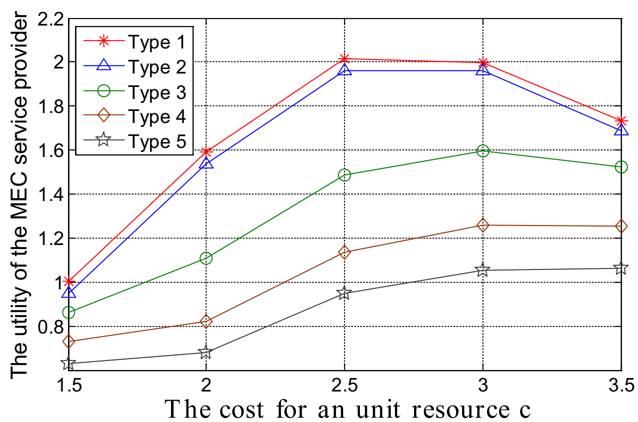

Fig. 3. The utilities of the service provider gained from an unit computation resource.

We compare the performance of our proposed scheme with these of two other schemes. One is the contract-based scheme without considering the allocation priority. The other is the non-contract scheme adopting fixed offloading payment of the vehicles. It can be seen that both the contract-based schemes make the MEC service provider gain more utility than the non-contract one. The reason is that in the contract theoretic approach, each contract is designed for the corresponding computation task type. Thus, the utility gained from providing offloading service to the vehicles can be improved through making the LDICs binding as described in Lemma 4. As our proposed scheme allocates the computation resource to the more profitable tasks with higher priorities, it brings the highest utility to the service provider.

Figure 3 shows the utility of the service provider gained from computing an offloading task on an unit resource. We can see that by utilizing an unit computation resource, the service provider gains higher profit from offloading tasks of the lower type vehicles. This result proves the Theorem 1 of this paper. In addition, the utility gained from each type of the vehicles first increases and then decreases. This can be explained as follows. According to Theorem 1, the utility gained from running a task T_i on an unit resource is defined as $u(i) = p_i/q_i - c$. In the contract-based offloading scheme, the provider raises the payment p_i to gain higher profit with the growth of the cost c. To cope with the increasing payment, vehicles will reduce the offloading resource consumption q_i. The rate of the demand reduction is higher at the lower c.

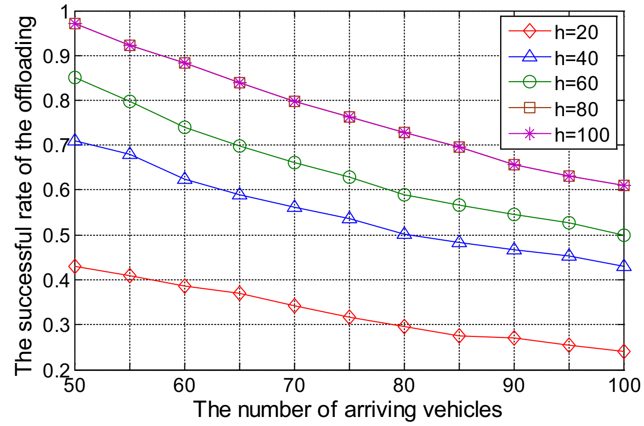

Fig. 4. The successful rate of the computation offloading service.

Figure 4 shows the successful rate of the task offloading. Due to the limitation of the MEC servers' computation resources, the rate falls down with the increasing number of the arriving vehicles. Furthermore, with the same arriving vehicles, the successful rate is higher in the case where the vehicles move faster. With a higher speed, a vehicle can access to more RSUs within its task delay constraint. Thus, it has more chance to offloading its computation task. Given the arriving number of vehicles, the successful offloading rate of the vehicles running at the speed 80 is equal to that of the vehicles running at 100. When their speed is higher than 80, the vehicles can access to all the RSUs along the road within the delay limitations of their tasks.

VI. Conclusion

In this paper, we proposed a cloud-based MEC offloading framework in vehicular networks. Based on the framework, we investigated the computation offloading mechanism. Both the resource limitation of the MEC servers and the latency tolerance of the computation tasks are considered in the design of the mechanism. Furthermore, we proposed a contract-based offloading and computation resource allocation scheme. The scheme aims to maximize the utility of the MEC service provider and satisfy the offloading requirements of the tasks. In addition, numerical results show the utility enhancement in our proposed scheme.

Acknowledgment

This work is supported by the projects 240079/F20 funded by the Research Council of Norway.

References

[1] P. Kolios, C. Panayiotou and G. Ellinas, "ExTraCT: expediting offloading transfers through intervehicle communication transmissions," *IEEE Trans. Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1238-1248, June 2015.

[2] X. Huang, R. Yu, J. Kang, N. Wang, S. Maharjan and Y. Zhang, "Software defined networking with pseudonym systems for secure vehicular clouds," *IEEE Access*, accepted.
[3] R. Yu, J. Ding, X. Huang, M. Zhou, S. Gjessing and Y. Zhang, "Optimal resource sharing in 5G-enabled vehicular networks: a matrix game approach," *IEEE Trans. Vehicular Technology*, accepted.
[4] R. Yu, Y. Zhang, S. Gjessing, W. Xia and K. Yang, "Toward cloud-based vehicular networks with efficient resource management," *IEEE Network Magazine*, vol. 27, no.5, pp. 48-55, Oct. 2013.
[5] J. Kang, R. Yu, X. Huang, H. Bogucka, S. Gjessing and Y. Zhang, "Location privacy attack and defense in cloud-enabled Internet of vehicles," *IEEE Wireless Communications*, accepted.
[6] G. Qiao, S. Leng, Y. Zhang and L. Xu, "Joint multi-RATs and cloud-service matching scheme in wireless heterogeneous networks," in *Proc. IEEE International Conference on Communications (ICC)*, pp. 1-6, May 2016.
[7] R. Yu, J. Ding, S. Maharjan, S. Gjessing, Y. Zhang and D.H.K. Tsang, "Decentralized and optimal resource cooperation in geo-distributed mobile cloud computing," *IEEE Trans. Emerging Topics in Computing*, accepted.
[8] J. Ding, R. Yu, Y. Zhang, S. Gjessing and D. Tsang, "Service providers competition and cooperation in cloud-based software defined wireless networks," *IEEE Communications Magazine*, vol.53, no.11, pp.134-140, Nov.2015.
[9] K. Zhang, Y. Mao, S. Leng, Q. Zhao, L. Li, X. Peng, L. Pan, G. Zhang, S. Maharjan and Y. Zhang, "Energy-efficient offloading for mobile edge computing in 5G heterogenous networks," *IEEE Access*, accepted.
[10] R. Yu, X. Huang, J. Kang, J. Ding, S. Maharjan, S. Gjessing and Y. Zhang, "Cooperative resource management in cloud-enabled vehicular networks," *IEEE Trans. Industrial Electronics*, vol. 62, no. 12, pp. 7938-7951, Dec. 2015.
[11] N. Cordeschi, D. Amendola and E. Baccarelli, "Reliable adaptive resource management for cognitive cloud vehicular networks," *IEEE Trans. Vehicular Technology*,vol. 64, no. 6, pp. 2528-2537, June 2015.
[12] H. Zhang, Q. Zhang and X. Du, "Toward vehicle-assisted cloud computing for smartphones," *IEEE Trans. Vehicular Technology*, vol. 64, no. 12, pp. 5610-5618, Sep. 2015.
[13] X. Chen, L. Jiao, W. Li and X. Fu, "Efficient multi-user computation offloading for mobile-edge cloud computing," *IEEE/ACM Transactions on Networking*, accepted.
[14] S. Nunna, *et al*., "Enabling real-time context-aware collaboration through 5G and mobile edge computing," in *Proc. Information Technology-New Generations (ITNG)*, pp. 601-605, Apr. 2015.
[15] U. Drolia, R. Martins, J. Tan, A. Chheda, M. Sanghavi, R. Gandhi and P. Narasimhan, "The case for mobile edge-clouds," in *Proc. Ubiquitous Intelligence and Computing*, pp. 209-215, Dec. 2013.
[16] Y. Zhang, R. Yu, W. Yao, S. Xie, Y. Xiao and M. Guizani, "Home M2M networks: architectures, standards, and QoS improvement," *IEEE Communications Magazine*, vol.49, no.4, pp.44-52, Apr. 2011.
[17] P. Bolton and M. Dewatripont, "Contract theory," *MIT press*, pp. 31-64, 2005.
[18] Y. Zhang, L. Song, W. Saad, Z. Dawy and Z. Han, "Contract-based incentive mechanisms for device-to-device communications in cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 10, pp. 2144-2155, Oct. 2015.
[19] L. Gao, X. Wang, Y. Xu and Q. Zhang, "Spectrum trading in cognitive radio networks: a contract-theoretic modeling approach," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 4, pp. 843-855, Apr. 2011.