Latency minimization for the MEC-aided vehicular network under Doppler Effect

A Y Du¹, P Wang² and L Song²
¹ The High School Affiliated to Renmin University of China, Beijing, 100080 P.R. China
² School of Electronics Engineering and Computer Science, Peking University, Beijing, 100871 P.R. China
E-mail: Allydu@qq.com

Abstract. In this paper, we consider a heterogeneous Mobile Edge Computing (MEC)-aided multi-layer vehicular network consisting of multiple vehicles as Edge Device (EDs), a roadside unit (RSU), and a cloud center (CC). The raw data produced by the vehicles can be processed at the three nodes separately. EDs transmit the data to the RSU via wireless communication link, and a wired link is considered between the RSU and the CC. Taking the mobility of vehicles and the limited amount of resources into account, we jointly discuss the task offloading and resource allocation in this three-layer MEC-aided vehicular network. Specifically, the interactions of Doppler Effect and the bandwidth provided by RSU to different vehicles are investigated. Aiming at minimizing the latency while increasing the utilization of resources, we design an iterative task assignment and resource allocation (ITR) algorithm. The simulation results indicate that: 1) The average latency of all tasks is significantly decreased, and the optimal bandwidth allocation scheme to the EDs is achieved; 2) Due to the Doppler Effect of vehicles, it is important to consider both the mobility of vehicles and the length of communication links when allocating resource.

1. Introduction

Vehicular network has gained an increasing popularity in recent years due to the pursuit of quick-response and high-reliability vehicular communication environment [1]. To improve the driving safety of vehicles, a larger amount of data should be calculated timely. To compensate for the vehicles’ limited calculation capacity, cloud computing is considered [2], but the excessive transmission latency for cloud computing has greatly restricted the performance of vehicular networks. Mobile Edge Computing (MEC) is then proposed to break through this bottleneck. By enabling every device from the data source to the cloud center, with computing capacity, large volume of data can be processed near the data source, and thus, a relatively shorter latency can be achieved [3][4].

There are two main challenges in such heterogeneous MEC-aided vehicular network, that is, 1) task offloading and resource allocation bring cross-layer interaction, and 2) The Doppler Effect leads to a distinct spectrum offset in signal transmission.

To address these challenges, we developed an iterative task assignment and resource allocation (ITR) algorithm that jointly optimized the task offloading and resources allocation to reach the minimum average latency, in which Doppler Effect is also considered.
Several researches have previously been conducted related to MEC and vehicular network, but most of them focus on the two-layer network [3], [5]-[6]. In [5], MEC-assisted network architecture is proposed and proven to be effective in reducing the latency. Beside these researches, only few initial works considered three-layer structure, such as [7]. Unfortunately, the influence of Doppler Effect and mobility of vehicles are not fully considered in the aforementioned works.

To ensure the best performance for vehicular networks, we analyzed the best way of task offloading and resources allocation and evaluated the influences of Doppler Effect.

The main contributions and results of this paper are summarized as below:
1) Consider MEC-aided multi-layer vehicular network.
2) To jointly reduce the task processing latency and increase the utilization of resource, a two-step iterative algorithm is proposed and simulated.
3) Numerical results show that the average latency is significantly reduced by performing our proposed ITR algorithm.

2. System model
We consider a communication structure in the Internet of Vehicle environment, including one CC, one RSU, and N Vehicles loaded with edge devices (EDs). The EDs are connected with the RSU through wireless links, while the RSU is connected with the CC through the wired links. The wireless links between the ED and RSU experience the Doppler effect because the vehicles are moving along the road.

We assume that all the N EDs connect only to the given RSU, and the RSU can connect only to the one CC. Each node in the network possesses a certain amount of computing capacity. The raw data is generated at the EDs. The results of the data from one ED must be aggregated at the same ED.

![Figure 1. The vehicle network scenario.](image)

2.1. Doppler Effect
Because the vehicle and the RSU are not relatively static, Doppler Effect appears when transmitting raw data through wireless links between the vehicle and RSU. The frequency of signal on each link is then influenced, and the signals will interfere each other because their actual bandwidth is changed. This interference will then decrease the quality of signal transmission or cause error in the result.

To eliminate these negative effects of Doppler Effect, we can adjust the communication bandwidth of vehicle \(i\) (for all \(1 \leq i \leq N\)) to \(B^i_s\) so that the receiving bandwidth does not exceed the bandwidth provided by the RSU for vehicle \(i\), \(B_i\). Assume the bandwidth allocated to vehicle \(i\) is from \(f^i_1\) to \(f^i_2\), so

\[
f^i_2 - f^i_1 = B_i. \quad (1)
\]
Let $v_c^i$ denote the velocity of vehicle $i$, and $\alpha_i$ denote the angle between $v_c^i$ and the line connecting vehicle $ED_i$ and the RSU. Let $\Delta v_i$ denote the relative velocity between vehicle $i$ and the RSU, which can be calculated as

$$\Delta v_i = v_c^i \cos \alpha_i.$$  \hspace{1cm} (2)

According to Doppler Effect, the difference between the actual frequency of the signal and the original frequency can be expressed differently for vehicles driving toward and away from RSU. For vehicles driving towards RSU:

$$\Delta f_i = \frac{\Delta v_i}{c+\Delta v_i} f_s^i$$ \hspace{1cm} (3)

For vehicles driving away from RSU:

$$\Delta f_i = \frac{\Delta v_i}{c-\Delta v_i} f_s^i$$ \hspace{1cm} (4)

where $\Delta f_i$ is the spectrum offset, and $c$ is light speed.

The relationship between $B_i$ and $B_s^i$ can then be expressed as

$$B_s^i = B_i - \Delta f_i.$$ \hspace{1cm} (5)

2.2. Edge device

The EDs loaded on the vehicles should generate the raw data, process part of raw data, and transmit the remaining part of the raw data to the next layer, the RSU.

Let $\lambda_{ED}$ denote the data generation speed of $ED_i$, and $\rho_i$ represent the compression ratio after data processing. The percentage of raw data processed by $ED_i$ is represented by $S_{ED}^i$. The computing capacity and transmitting capacity of each ED per unit time is denoted by $\theta_{ED}$ and $\phi_{ED}(B_s^i)$, where $\phi_{ED}(B_s^i)$ is a function of the bandwidth of $ED_i$, $B_s^i$, calculated by the Shannon Formula. Since the total bandwidth provided by the RSU, $B_{RSU}$, is limited, the sum of all $B_s^i$ is limited.

2.3. Road side unit

The RSU should receive the remaining raw data from the vehicles and process part of the received raw data. Then, the RSU must transmit the remaining raw data to the Cloud Center through a wired link and transmit the processed data back to the vehicle through wireless channels.

Let $S_{RSU}^i$ denotes the division percentage of RSU for the data from $ED_i$. The computing capacity of RSU per unit time is denoted by $\theta_{RSU}$. The total amount of raw data processing at the RSU should be no greater than $\theta_{RSU}$. The transmitting capacity between the RSU and the CC is $\phi_{RSU}$. The amount of raw data transmitted to the CC is limited by $\phi_{RSU}$.

2.4. Cloud center

The CC is responsible for receiving and processing all the remaining raw data and transmitting the processed data all the way back to the vehicle.

The raw data arriving speed at the CC is $\lambda_{CC}$. $S_{CC}^i$ denotes the division percentage of CC for the data from $ED_i$. The computing capacity of the CC can be denoted by $\theta_{CC}$, which limits the data arriving speed. The processed data is transmitted to the RSU through a wired link first, and then transmitted to $ED_i$ through wireless channels. The transmitting capacity of each link is $\phi_{RSU}$ and $\phi_{ED}(B_s^i)$.

3. Problem formulation

In this scenario, we aim at minimizing the overall system latency. The latency of processing a task generated from one vehicle is defined as the sum of the total time used to process the raw data at the nodes and total transmitting time between the nodes.

When the data is processed at $ED_i$, the latency for the data with the unit size can be calculated as
because the processed data do not need to transmit to the RSU or CC, the transmitting time is 0.

When the data is processed at the RSU, the latency for the data with the unit size can be calculated as:

$$L_{RSU}^i = \frac{1}{\theta_{RSU}} + \frac{1}{\phi_{RSU}} + \frac{\rho_i}{\phi_{ED}(B_i^j)}$$

(7)

The first term is the transmitting time of the raw data from the $ED_i$ to RSU, the second term is the data processing time at the RSU, and the third term is the transmitting time of the processed data from RSU back to $ED_i$.

When the data is processed at the CC, the latency for the data with the unit size can be calculated as:

$$L_{CC}^i = \frac{1}{\phi_{ED}(B_i^j)} + \frac{1}{\theta_{CC}} + \frac{1}{\phi_{RSU}} + \frac{\rho_i}{\phi_{ED}(B_i^j)}$$

(8)

The first two terms are the transmitting time of the raw data from the $ED_i$ to CC, the third term is the data processing time at CC, the last two terms are the transmitting time of the processed data from the CC back to $ED_i$.

Considering the task division percentage at ED, RSU, and CC as $S_{ED}^i$, $S_{AP}^i$, and $S_{CC}^i = 1 - S_{AP}^i - S_{ED}^i$.

Since the data processing at each node can be done simultaneously, the latency of $ED_i$ can be formulated as:

$$L_i = \lambda_{ED} \max(S_{ED}^i L_{ED}^i, S_{RSU}^i L_{RSU}^i, S_{CC}^i L_{CC}^i)$$

(9)

We consider the overall minimal latency for all cars in the scenario, so we need to reach the minimum average latency for N cars, which can be calculated as:

$$L = \frac{1}{N} \sum_{i=1}^{N} L_i$$

(10)

Hence, the latency minimization problem in our scenario can be formulated as

$$\min L$$

s.t. (1) - (9).

4. Algorithm design

In this section, we design the iterative task assignment and resource allocation (ITR) algorithm for the average latency minimizing problem.

In our proposed ITR algorithm, the task assignment of the three layers (i.e., the vehicle, RSU and CC) and the resource allocation is performed iteratively. In each iteration, two steps are included:

- Step1: The task assignment among the vehicle, RSU and CC for each transmission link is optimized given the previous resource allocation scheme, so that the maximum latency among these nodes are minimized and a maximum resource utilization is achieved.

- Step2: According to the updated task assignment, the computing and transmission resource of each node is reallocated, aiming to minimize the overall latency of all links, and thus, the average processing latency of each vehicle can be reduced.

5. Simulation and Results

In this section, we evaluate how Doppler Effect influence the average latency and resource allocation to the nodes. We consider a typical MEC-aided three-layer vehicular network that includes a Cloud Center, a Roadside Unit, and 20 vehicles on the same road.

5.1. Experiment scenario and setup

As shown in Figure 1, to better analyze the effect of Doppler Effect from the aspect of vehicular locations and movement directions, all the considered vehicles are assumed to move at the same speed, and the
distance between two neighboring vehicles is constant. Each vehicle produces the same amount of data files that will be computed in the three nodes: vehicles, RSU, and CC.

We used MATLAB to simulate the computation and transmission procedures. The parameter settings are listed in Table 1. The influence of Doppler Effect is evaluated by the following indicators:

- Latency: The latency represents the minimum time it takes for the data file to be completely computed. We evaluate both the latency for each vehicle and the average latency of the 20 vehicles.
- Resources Allocation: The resources allocation represents the way to allocate data files to the three nodes (i.e., Vehicles, RSU and CC) that ensures the minimum latency.

| Table 1. Parameter settings. |
| Computing capacity of each vehicle | $30 \times 10^3$ CPU Cycles/s |
| Computing capacity of RSU          | $3 \times 10^6$ CPU Cycles/s  |
| Computing capacity of CC          | $10^{10}$ CPU Cycles/s        |
| Wireless transmission resources of RSU | 80 kHz                       |
| Wired transmission resource      | 10 Mbps                       |
| Data generation speed             | 1 kbps                        |
| Compression ratio                 | 10%                           |

5.2. Simulation results

In this subsection, we consider the influences of Doppler Effect in two variables: the location of vehicles and the velocity of the vehicles.

Figure 2 describes how latency changes with the locations of vehicles. When the distance between vehicle and RSU (at location 10.5) reduces, the latency decreases, because the relative speed between the vehicle and RSU, $\Delta v_i$, becomes smaller with the decreasing vehicle-RSU distance. The farther the vehicle is from the RSU, the larger the $\Delta v_i$, and thus the smaller the spectrum efficiency and higher latency.

Figure 3 shows the raw data allocation proportion to the vehicle, RSU, and CC change as the location of vehicles change. In location 1 to 8 and 14 to 20, more of the tasks are allocated to the vehicle than to RSU and CC. This is because the distances between the vehicles and RSU are so large that the low spectrum efficiency caused $L_{RSU}^i$ and $L_{CC}^i$ to be bigger than $L_{ED}^i$. For location 9 to 13, as the distance from vehicles to RSU decreases, the influence of Doppler Effect decreases and that of computing capacity became visible. Since $\theta_{ED}$ is much smaller than $\theta_{RSU}$, and $\theta_{RSU}$ is smaller than $\theta_{CC}$, more tasks are assigned to RSU and CC than to the Vehicles.
Figure 4. The latency vs. velocity of vehicles.

Figure 4 shows how the average latency changes as the velocity of vehicles change. As the velocity increases, the average latency increases in an accelerating rate. This is because an increasing vehicular speed causes a larger relative speed $\Delta v_i$, leading to a smaller spectrum efficiency. Therefore, we are more likely to experience a high latency when driving in higher speed, such as on highways.

6. Conclusion

In this paper, we analyzed the MEC-based vehicular network performance under Doppler Effect, proposed a task assignment and resource allocation (ITR) algorithm, and showed that the average latency can be minimized while the utilization of bandwidth resource is maximized. The simulation results indicate that:

- As vehicles drive along the road passing by a RSU, the task processing latency first decreases when the vehicle getting close to the RSU, and then increases when it leaving away.
- As the velocity of vehicles increases, the latency increases in an accelerating rate, which indicates the necessity of providing more bandwidth to vehicles with high speed.
- Vehicles farther from the RSU require fewer computational resources from RSU and CC, rendering the importance of considering both the mobility of vehicles and the path loss of communication links when allocating resources.

References

[1] H. Peng, Le Liang, X. Shen and G. Y. Li, "Vehicular Communications: A Network Layer Perspective," in IEEE Transactions on Vehicular Technology, vol. 68, no. 2, pp. 1064-1078, Feb. 2019, doi: 10.1109/TVT.2018.2833427.

[2] M. Gerla, "Vehicular Cloud Computing," 2012 The 11th Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net), Ayia Napa, 2012, pp. 152-155, doi: 10.1109/MedHocNet.2012.6257116.

[3] K. Zhang, Y. Mao, S. Leng, Y. He and Y. Zhang, "Mobile-Edge Computing for Vehicular Networks: A Promising Network Paradigm with Predictive Off-Loading," in IEEE Vehicular Technology Magazine, vol. 12, no. 2, pp. 36-44, June 2017, doi: 10.1109/MVT.2017.2668838.

[4] P. Wang, C. Yao, Z. Zheng, G. Sun and L. Song, "Joint Task Assignment, Transmission, and Computing Resource Allocation in Multilayer Mobile Edge Computing Systems," in IEEE Internet of Things Journal, vol. 6, no. 2, pp. 2872-2884, April 2019, doi: 10.1109/JIOT.2018.2876198.

[5] M. Emara, M. C. Filippou and D. Sabella, "MEC-Assisted End-to-End Latency Evaluations for C-V2X Communications," 2018 European Conference on Networks and Communications (EuCNC), Ljubljana, Slovenia, 2018, pp. 1-9, doi: 10.1109/EuCNC.2018.8442825.

[6] K. Zhang, Y. Mao, S. Leng, Y. He and Y. Zhang, "Mobile-Edge Computing for Vehicular Networks: A Promising Network Paradigm with Predictive Off-Loading," in IEEE Vehicular Technology Magazine, vol. 12, no. 2, pp. 36-44, June 2017, doi: 10.1109/MVT.2017.2668838.
[7] K. Zhang, Y. Mao, S. Leng, A. Vinel and Y. Zhang, "Delay constrained offloading for Mobile Edge Computing in cloud-enabled vehicular networks," 2016 8th International Workshop on Resilient Networks Design and Modeling (RNDM), Halmstad, 2016