Optimal Resource Sharing in 5G-Enabled Vehicular Networks: A Matrix Game Approach

Rong Yu, *Member, IEEE*, Jiefei Ding, Xumin Huang, Ming-Tuo Zhou, *Senior Member, IEEE*, Stein Gjessing, *Member, IEEE*, and Yan Zhang, *Senior Member, IEEE*

***Abstract*—Vehicular networks are expected to accommodate a large number of data-heavy mobile devices and multiapplication services, whereas it faces a significant challenge when we need to deal with the ever-increasing demand of mobile traffic. In this paper, we present a new paradigm of fifth-generation (5G)-enabled vehicular networks to improve network capacity and system computing capability. We extend the original cloud radio access network (C-RAN) to integrate local cloud services to provide a low-cost, scalable, self-organizing, and effective solution. The new C-RAN is named enhanced C-RAN (EC-RAN). Cloudlets in EC-RAN are geographically distributed for local services. Furthermore, device-to-device (D2D) and heterogeneous networks are essential technologies in 5G systems. They can greatly improve spectrum efficiency and support large-scale live video streaming in short-distance communications. We exploit matrix game theoretical approach to operate the cloudlet resource management and allocation. A Nash equilibrium solution can be obtained by a Karush–Kuhn–Tucker (KKT) nonlinear complementarity approach. Illustrative results indicate that the proposed resource-sharing scheme with the geodistributed cloudlets can improve resource utilization and reduce system power consumption. Moreover, with the integration of a software-defined network architecture, a vehicular network can easily reach a globally optimal solution.**

***Index Terms*—Cloud radio access network (C-RAN), fifth generation (5G), matrix game, resource management, software-defined network (SDN), vehicular network.**

I. Introduction

VEHICULAR networks are facing challenges in both communications capacity and computing capability because mobile traffic has explosively increased. As electric vehicles (EVs) become popular, an explosion of mobile traffic applications follows with the popularity of the Internet of things. Vehicular networks may have several limitations: limited spectrum resource, incapable on-board devices, and inefficient system management. To improve the performance of vehicular networks for flexible resource allocation, automated network organization, and advanced mobility management, we enhance the structure of vehicular networks by combining with fifth-generation (5G) technologies, such as the cloud radio access network (C-RAN) architecture, and cloud computing [1], which presents a novel 5G-enabled vehicular network. This new type of networks can efficiently satisfy the demand for vehicular services with heavy-data traffic or a complex computing process.

Facing the rapid increase of mobile services, the 5G wireless network has an enticing prospect in vehicular networks to enable high-data-rate applications in a vehicle environment [2]. Fifth generation employs the small-cell technology that enables high spatial and frequency reuse in limited coverage area [3]. C-RAN is defined as a centralized RAN, which separates baseband units (BBUs) from radio access units and migrates BBUs to the cloud to form a BBU pool for centralized processing [4]. This approach allows the deployment of network functions in a cloud data center to leverage traffic load through virtualization techniques. However, radio signal exchange is limited by the capability of fiber links between remote radio heads (RRHs) and the data center. Hence, decentralized processing and hierarchical management are significant to improve the performance of vehicular networks. Since data transmission is easily influenced by the surrounding environment [5], device-to-device (D2D) is an efficient approach for direct and short-distance transmission between vehicles. Moreover, D2D communications can further enhance the communications capacity by allowing nearby devices to establish local links [6]. The study in [7] discussed the technique of D2D frequency reuse to reduce communications latency. Users can simultaneously have multiple wireless signal sources and switch between them for smooth data transmission [8].

In terms of task processing, cloud computing and software-defined network (SDN) architecture are significant to improve the performance and the flexibility of vehicular networks. The main purpose of cloud computing is to provide an approach for users to run computation-intensive applications that are not easily performed on a resource-constrained mobile device [9]. Geodistributed cloud has been regarded as a promising implementation in vehicular networks for the sake of low communications delay [10], [11]. The data center, cloudlets,

Manuscript received July 19, 2015; revised December 6, 2015; accepted January 29, 2016. Date of publication March 1, 2016; date of current version October 13, 2016. This work was supported in part by programs of the National Natural Science Foundation of China under Grant 61422201, Grant 61370159, Grant U1201253, and Grant U1301255; by the Science and Technology Program of Guangdong Province under Grant 2015B010129001; by the Special-Support Project of Guangdong Province under Grant 2014TQ01X100; by the High Education Excellent Young Teacher Program of Guangdong Province under Grant YQ2013057; by the Science and Technology Program of Guangzhou under Grant 2014J2200097 (Zhujiang New Star Program); and by the projects 240079/F20 funded by the Research Council of Norway and the European Commission FP7 Project CROWN under Grant PIRSES-GA-2013-627490. The review of this paper was coordinated by Dr. T. Luan. (*Corresponding author: Yan Zhang.*)

R. Yu, J. Ding, and X. Huang are with the School of Automation, Guangdong University of Technology, Guangdong 510006, China (e-mail: yurong@ieee.org; jiefeiding@gmail.com; huangxu_min@163.com).

M.-T. Zhou is with the Shenyang Institute of Automation Guangzhou Branch, Chinese Academy of Sciences, Beijing 100864, China (e-mail: zhou.mingtuo@ieee.org).

S. Gjessing and Y. Zhang are with Simula Research Laboratory and University of Oslo, Norway (e-mail: steing@ifi.uio.no; yanzhang@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2016.2536441

and RRHs form a hierarchical network that can be managed by SDN architecture for high efficiency and better quality of service (QoS) [12]. The most significant advantage of the SDN is the three-layer vertical integration architecture, which breaks through the control logic from underlying devices [13], [14]. In case of fast-moving vehicles, cost-efficient resource allocation and joint management are important for vehicle applications to provide location-based services, e.g., navigation services, and emergency alerts [15].

Automated network organization is a new feature of 5G-enabled vehicle networks. Resource sharing provides a new aspect for allocating resource on a demand-based approach [16], [17]. It avoids some costs for resource overutilization or resource redundance to improve resource utilization. However, in a geodistributed cloudlet network, the resource allocation of a cloudlet will interact with nearby cloudlets. A matrix game is a confrontation of n players $(n \geq 2)$ whose decision incudes multiple factors in a normal and noncooperative context [18]. In [19], Sharma *et al.* adopted a matrix game to resolve the payoff matrix of strategic firms in the electricity market. Matrix game is an approach for multiple-player games and can be a potentially efficient method for resource allocation in vehicular networks.

In this paper, we propose a new paradigm for vehicular networks in a 5G communications environment, i.e., EC-RAN, which aims to support data-heavy applications, such as augmented-reality applications. Vehicle users can access services by an on-board unit (OBU) or their portable devices. With the combination of C-RAN and cloud computing, EC-RAN not only improves communications qualities by employing the mixed-network deployment of small-cell, D2D, and heterogeneous networks (HetNets) but also enhances vehicle computing capabilities by offloading mobile services to cloud. In an SDN framework, management, control, and execution are separated and supposed to be managed independently in three layers. Through intercooperation, the three layers can perform well as a whole system. To improve system performance, we adopt graph theory to demonstrate the features of geographic distribution and the relationship between cloudlets. We formulate the resource allocation problem as a noncooperation matrix game and solve it through a nonlinear concave optimization approach. In addition, illustrative results demonstrate that resource sharing among cloudlets significantly improves the performance of 5G-enabled vehicular networks and reduces system operation cost. The major contributions of this paper can be summarized as follows.

- We define a new paradigm of 5G-enabled vehicular networks. A new concept of enhanced C-RAN (EC-RAN) is proposed for communications enhancement and computing capability improvement. Geodistributed RRH and D2D approach can efficiently allocate radio resource and greatly improve the communications capacity of moving vehicles.
- We formulate the resource allocation of each cloudlet as a noncooperation matrix game. In the optimal solution, resource sharing aims to improve resource utilization. Moreover, we obtain the Nash equilibrium by employing the Karush–Kuhn–Tucker (KKT) condition into nonlinear optimization.
- We resolve the resource allocation problem in the geodistributed cloudlets to reduce system operation cost. Through exploiting the SDN framework to manage the hierarchical vehicular network, we can obtain the global optimization results with high QoS and more revenues.

The rest of this paper is organized as follows. In Section II, we describe the architecture of a 5G-enabled vehicular network and discuss the EC-RAN architecture, which includes cloudlets, D2D communications, and SDN technology. System model and problem formulation are presented in Section III. In Section IV, we formulate a noncooperation matrix game and solve it by the KKT condition approach. Illustration results are presented and discussed in Section V. Section VI concludes this paper.

II. Fifth-Generation-Enabled Vehicular Network

Fifth-generation wireless communications technologies will revolutionize the current vehicular network to satisfy the continuously increasing demand for high data rate and mobility. The proposed 5G-enabled vehicular network has three important features: 1) the EC-RAN and D2D technology for resilient communications; 2) the geodistributed cloudlets for application processing; and 3) the SDN architecture for system management and control. The integrated framework enhances the vehicular network with the advantages of a robust communications network, powerful processing capabilities, and flexible resource allocation. This system framework is also motivated by the purpose of energy saving and cost reduction. In Fig. 1, our proposed paradigm of a 5G-enabled vehicular network is presented.

A. Fifth-Generation-Enabled Communications Approaches

In Fig. 1, 5G communications enable vehicles to access base stations and communicate with cloudlets. In the traditional approach, vehicle communications and data processing are mainly handled by OBUs. The 802.11b/g protocal has limited transmission, which is often affected by vehicle density and vehicle average speed [20], [21]. Fifth-generation cellular networks can offer superior performance in terms of throughput, delay, reliability, scalability, and mobility. In 5G networks, the peak data rate for low mobility is 10 Gb/s, and the peak data rate for high mobility is 1 Gb/s. It can support transmission latency shorter than about 1 ms for moving vehicles and also serve for high-speed trains with speed from 350 to 500 km/h [22], [23]. In a vehicular network, the 5G technologies of interest are EC-RAN and D2D communications.

1) EC-RAN: The BBU constitutes a pool in each cloudlet to easily share signaling, data, and channel state information for users in a vehicular network. This approach can greatly improve the radio resource utilization by aggregating and dispatching the BBU bandwidth. EC-RAN implements a soft and virtualized BBU pool and distributed RRHs at remote sites.

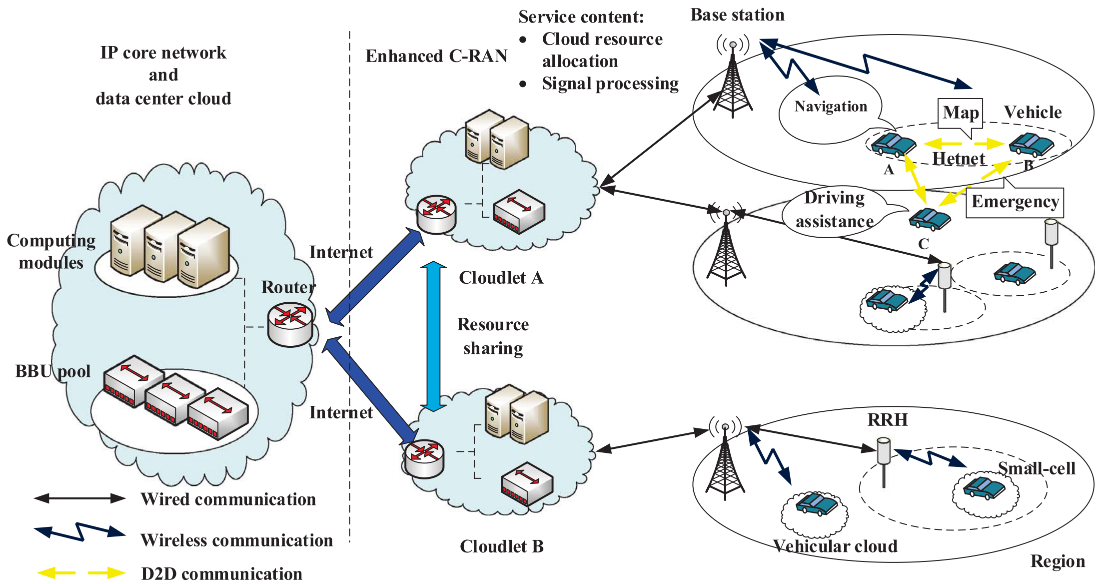


Fig. 1. Fifth-generation-enabled vehicular network architecture.

All RRHs connect to a switcher/central device that can flexibly schedule radio resource in a BBU pool for one RRH or a set of RRHs. The vehicle can directly communicate with RRHs at the road side or with nearby base stations through wireless communications. RRHs communicate with a BBU pool through wired communications for reliability and low latency, such as optical fiber links. By this approach, the BBU can assign bandwidth resource for each RRH based on user demand. Through controlling the number of RRHs, we can adjust the size of small cells and significantly reduce intercell interference. Moreover, resource allocation can be realized in a simplified procedure, e.g., programming the software.

2) D2D: D2D enables the direct transmission between different vehicle devices. It combines with device-to-base-station (D2B) to form a HetNet. Proximity-based D2D communications are emerged as a promising technology for effective sharing of resources (spectrum and computing resources, etc.) and for users who are spatially close to each other [24]. In a mixed-network deployment scenario, D2D can create opportunities for spectrum reuse and reduce communications latency for long-distance transmission or traffic congestion. On the other hand, D2D communications can potentially promote resource sharing among vehicles through device relaying, which resembles ad hoc communication. In Fig. 1, multiple D2D links are established in device pairs, i.e., A-B, B-C, and C-D. They form a D2D cell and share the same information based on geographic distribution. It is similar to a social network, which is a promising approach based on social relations [25]. The other case is that D2D can relay for traffic offloading. Vehicle A is much closer to the base station and can act as a relay for its nearby vehicle C. D2B and D2D form a combined-communications network that aims to improve data transmission performance with minimal cost. Moreover, the base station communicates with a data center through a wide area network (WAN).

D2D transmission for intervehicle communications will be affected by the number of surrounding vehicles and nearby D2D cells [26]. Fifth-generation small cells enable high spatial and frequency reuse in limited coverage area. As the density of 5G small cells grows, the transmission power of the scheduled base station is decreased, and the achieved network efficiency is increased. However, the intercell interference will be increased. A D2D cell is a tiny cell that coexists with 5G small cells. To minimize intercell interference, a 5G small cell will accept a limited number of D2D cells. We assume that the D2D transmission probability of vehicle j is denoted as

$$P_j = 1 - \nu^{n_j}. \tag{1}$$

Here, ν is the probability of D2B communication $(0 < \nu < 1)$. n_j is the number of vehicles in a D2D cell.

3) Data Transmission Through 5G Core Network: Data transmission through a 5G core network is greatly affected by the link condition. Here, we use quality of link (QoL) to evaluate the link condition between different cloudlets, because it is unsuitable to exploit a link that reaches the maximum link condition. We define QoL by the probability of a successful packet transmission in a specific small time interval. Apparently, a link with a faster transmission rate and lower delay will have a higher QoL and, therefore, obtain a larger chance at resource sharing. We employ the softmax method by using the Boltzmann distribution theory [27]. The probability that a link is selected as a backhaul during resource sharing is given by

$$\theta_{i,k} = \frac{e^{\frac{q_{i,k}}{\tau}}}{\sum_{j=1}^{n} e^{\frac{q_{i,j}}{\tau}}}. \tag{2}$$

Here, $q_{i,k}$ denotes QoL between cloudlets i and k. τ is a temperature parameter of Boltzmann distribution. A high temperature will cause the selection of a backhaul link to be all equally probable. A low temperature will generate a greater difference in the selection probability for links that differ in their QoLs.

B. Cloud Computing in 5G-Enabled Vehicular Network

Cloud computing in a 5G-enabled vehicular network includes one data center and multiple cloudlets. Long-distance transmission will increase communication latency, require more bandwidth resource, and create congestion of Internet traffic flow in a WAN [28], [29]. Compared with the centralized cloud, the distributed cloudlets are geographically close to vehicles and can serve as a subordinate unit of the data center. In Fig. 1, cloudlets are distributed in three regions to provide elastic computing services. Cloud computing in a 5G-enabled vehicular network architecture can be divided into three main scenarios.

- At the end-user side, a vehicle can access the vehicular cloud through D2B or D2D. Service can be classified into different purposes. Users will pay for this service by the amount of service time and feedback the evaluation of QoS. A vehicle can also be the source of messages. For example, a vehicle can transmit map data and share public information to nearby vehicles through the D2D approach. Application requirements will be directly offloaded to the local cloudlets through wireless communication.
- At the edge device side, a cloudlet (or called roadside cloud, RAN cloud) is a set of resource-rich computers and BBUs that are connected to the Internet and provide service to nearby users. A cloudlet has three important tasks. First, it provides direct service as a service provider and caches the required data as preparation for the next related service. Second, it manages the communication within mobile users or between the base station and mobile users, such as bridging D2D connection between two nearby EVs, allocating the transmission channel for the BBU pool, and conducting data-sharing service. Third, it serves as a middle layer by offloading some incapable applications to the superior data center cloud and then returning results to the end user.
- At the IP core side, the data center cloud is a cluster of physical devices and software services. It can possess both radio resource and computing resource and make profit by renting resource to service providers for a long-term period. Each service provider has limited cloud resource that can be based on user demand and makes profit from providing service by allocating a virtual machine to run applications. Cloud storage technology is another important approach to increase the resource utilization efficiency of cloud computing [30]. Moreover, a centralized BBU pool is located in a data center and responsible for local BBU pool management.

In this paper, resource sharing can be conducted between the nearby cloudlets. For example, when a cloudlet has many application requests to process in a busy time, it may find resource assistance from the nearby cloudlets as long as bandwidth is enough to support data transmission. Resource assistance is similar to run an application in the local cloudlet. For radio resource sharing, we consider channel coordination between vehicles [31]. For a large-area management, a unfield control platform plays an important role in system design. The SDN framework brings a promising approach for a geodistributed cloud computing network.

C. SDN Framework in EC-RAN

The basic idea of the SDN framework in EC-RAN is to separate control logic (i.e., the control plane) from many underlying cloudlets and communication equipments (i.e., the data plane), as shown in Fig. 2. The decision of resource allocation is made in the management plane (a control center) that includes service providers, a network manager, and a baseband manager. They work independently and provide service for application requests in the control center. The control plane in the second layer is responsible for data-forwarding strategies and resource-sharing strategies. The significant feature of the control plane is the virtualization of small cells in the data plane. A small cell can be regarded as a node in the EC-RAN network. Cloudlets are located in different small cells. The operation state of the cloudlet can be represented as three variables: CPU resource, memory resource, and bandwidth resource. The link state between two connected nodes can be evaluated as the QoL. Algorithm design and control decision are made based on the information from the control plane. At the bottom layer, a cloudlet and RRH in a base station provide local service to users. Cloudlets are controlled by the manager from the data center in the management plane and performs the control decision. For example, the BBU pool in a cloudlet is responsible for baseband processing to reduce the interference between different RRHs or D2D and maximize energy efficiency. The baseband manager in the management plane is responsible for the RRH on/off management to minimize the number of activated RRHs. The difference between the BBU pool and the baseband manager is that they are working in different layers. The BBU pool processes the baseband resource in the physical layer. The baseband manager controls the bandwidth resource in the application layer based on the virtualization of EC-RAN in the control plane. With the management from the SDN framework, a 5G-enabled vehicular network can make the best of the idle resource in data centers and achieve complex but efficient network management functionalities.

III. Problem Formulation

A. Network Model

We consider M geographically distributed cloudlets in different regions (e.g., one cloudlet for one region). Cloudlets can form a graph $g = \{V, E, w\}$ $(i = M)$, which is employed to describe the resource distribution and relationship of SPs. Set V represents the collection of a cloudlet L_i $(i = 1, 2, \ldots, M)$, and weight w on them denotes the link condition between two regions. The network topology depends on the geographical location of each region, e.g., the connection between L_1 and L_2 means that these two regions are adjacent. Set E is the collection of connected edges. For example, we have $e_{L_1,L_2} = 1$ $(e_{L_1,L_2} \in E)$. In this paper, the meaning of a connected edge is more than geography connection. It also represents the opportunity for cloudlet resource sharing. The cloudlet resource

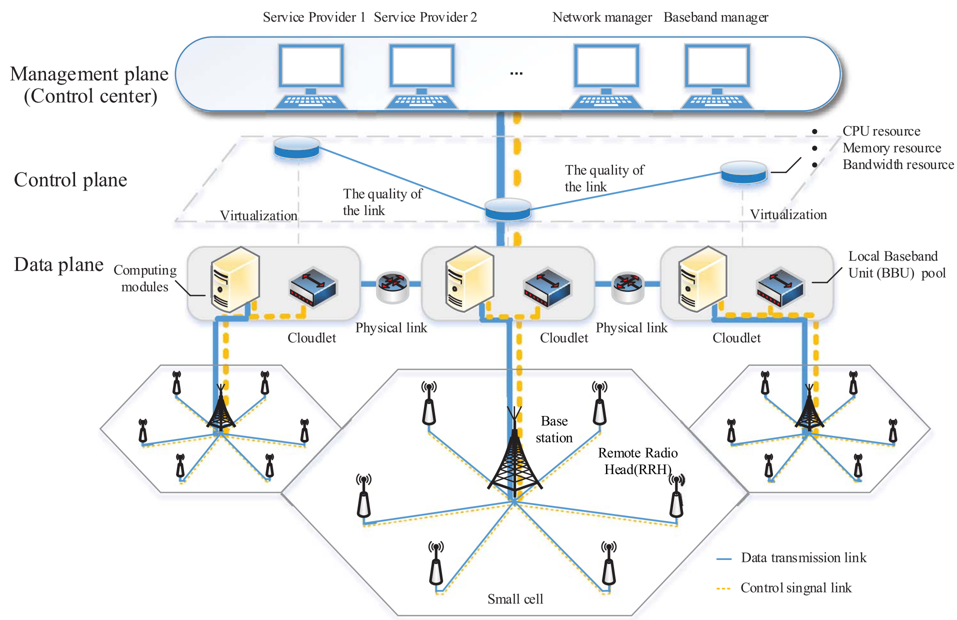


Fig. 2. SDN framework in EC-RAN.

sharing is to balance the resource utilization in the following two scenarios.

- **Resource Shortage**: If L_1 operates with a high degree of resource utilization, it will have few resource for new coming applications. This situation often happens when many vehicle users request for service simultaneously. The traditional approach is to increase the number of physical devices with the condition of raising the cost for better service. The disadvantage is that L_1 will operate on a low degree of resource utilization in most of the time. The resource-sharing approach can resolve this problem in a flexible way. For example, L_1 is allowed to rent resource from L_2, L_3, and L_4, as shown in Fig. 3. In other words, users are allowed to access the nearby cloudlets and run applications on cloudlet devices as long as the link condition can support data transmission. It is noted that L_1 can rent resource from more than one cloudlet.
- **Ultradense Deployment**: Ultradense deployment is an important feature of a cloudlet. To improve energy efficiency, a cloudlet can turn off its power automatically when it works on low utilization and can be replaced by the nearby cloudlets. For example, at midnight, most people are in deep sleep. The cloud manager will evaluate the utilization of each cloudlet and turn off some cloudlets, such as L_1. The user in L_1 will access the nearby cloudlets for service, such as that in L_2.

Intercloud resource sharing is very important for cloudlets to improve profit and reduce the system power consumption. The control result includes the amount of allocating resource and the on/off state of cloudlets. Moreover, the resource-sharing strategy should satisfy the capability of each cloudlet. For a cloudlet L_i, the maximum capacity can be evaluated by the CPU resource $\max C_i$, the memory resource $\max M_i$, and the bandwidth resource $\max B_i$.

The resource requests from user j can be denoted by $R_j = (C_j, M_j, B_j, T_j)$. Requests include the information of the amount of CPU resource C_j, the amount of memory resource M_j, the required bandwidth B_j, and the maximum latency time T_j. Every application has a specific maximum latency T_j, which should not be exceeded to provide a satisfactory QoS to a user. Therefore, the link condition will be an important factor to determine the success of resource sharing. To evaluate the resource utilization of a cloudlet, we normalize the three types of resources in a uniformed value and make the value of resource utilization no more than 1, i.e.,

$$y_j = c\frac{C_j}{\max C_i} + a\frac{M_j}{\max M_i} + b\frac{B_j}{\max B_i} \tag{3}$$

where y_j is the weighted sum of resource that a cloudlet L_i allocates to user j. c, a, and b are the fixed coefficients of three kinds of resources (CPU resource, memory resource, and bandwidth resource) and satisfy the relationship of $c+a+b=1$.

B. Utility Function

The stimulating approach in a 5G-enabled vehicular network includes revenue and operation utility. A user will pay for service by the amount of service time. If a cloud operates more applications, it can receive more revenues. However, unoccupied resource will also consume energy. Thus, the purpose of resource management of the cloud is trying to improve the resource utilization of each cloudlet and decrease the number of idle or low-utilization cloudlets. The utility function of a 5G-enabled vehicular network is given by

$$U = R\sum_{i=1}^{M} y_i + u\sum_{i=1}^{M}\log(y_i + 1) \tag{4}$$

where y_i is the utilization of cloudlet i, and $0 \le y_i \le 1$. R is the unit of revenue that a user pays for service. u is the fixed

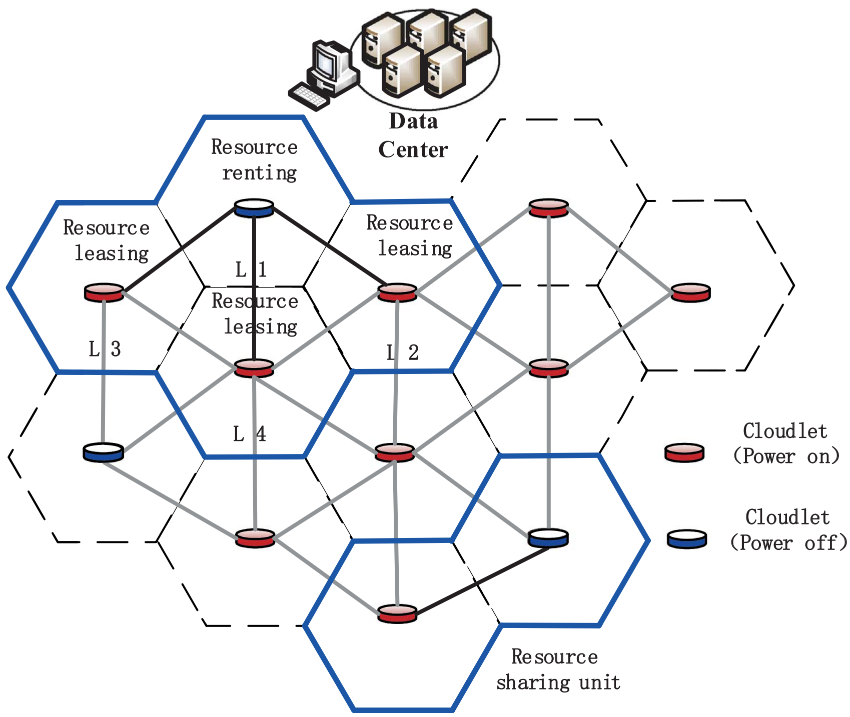


Fig. 3. Cloudlet resource sharing in a 5G-enabled vehicular network.

coefficient of operation utility. The operation utility is a concave and increases monotonically. This indicates that operating in a low resource utilization is bad for utility. From the utility function, we observe that the maximum value of utility will appear when resource utilization reaches a good balance between cloudlets. The balance will be obtained through the resource-sharing approach.

C. Cloudlet Resource Sharing

Cloudlet resource sharing is based on the conditions of satisfying link quality and cloud resource redundance. For example, a cloudlet will prefer to cooperate with the cloudlets with more unoccupied resource than itself. The unoccupied resource here is the resource available for new coming applications. We employ a penalty factor $(y_i)^m$ $(m \geq 2)$, which is an upward curve. The penalty is intended to avoid cloudlet resource overutilization. This means that the penalty will be increased as y_i is increasing. As a result, L_i will try to decrease utilization through resource sharing with the nearby cloudlets. Thus, $1-(y_i)^m$ denotes the willingness of a cloudlet to accept new coming applications. Resource sharing aims to balance the utilization of cloudlets. Thus, the allocation function of each cloudlet is related to the willingness of all related cloudlets, i.e.,

$$F(x_i) = (1-(x_{i,i}+\rho_i)^m)+\chi_{i,k}\sum_{k=1,k\neq i}^{M}(1-(x_{i,k}+\rho_k)^m)$$
$$\chi_{i,k} = \tau_k \frac{\delta_2}{(\theta_{i,k}+\delta_1)}$$
$$y_i = x_{i,i}+\rho_i, \quad y_k = x_{i,k}+\rho_k$$
$$\tau_k = \begin{cases}1, & \text{if turn on}\\ 0, & \text{if turn off}\end{cases} \tag{5}$$

where $x_{i,i}$ is the amount of resource to be allocated on L_i for new coming applications. $x_{i,k}$ is the amount of resource renting from L_k. ρ_i is the amount of resource that has been occupied on a cloudlet L_i. The occupied resource includes the shared resource with other cloudlets in the former optimization result (e.g., $x_{k,i}$). Therefore, the resource utilization of L_i, y_i, is the sum of $x_{i,i}$ and ρ_i. τ_k is the decision variable (either 0 or 1) of a cloudlet state. $\theta_{i,k}$ is the link choice probability. δ_1 and δ_2 are the parameters that determine the influence from $\theta_{i,k}$. $\chi_{i,k}$ denotes link condition. The value of $x_{i,k}$ should satisfy the constraint of the total required resource π_i. Then, we have

$$\text{s. t.} \quad \sum_{k=1}^{M} x_{i,k} = \pi_i$$
$$x_{i,k} \geq 0 (k=1,2,\ldots,M). \tag{6}$$

To a single cloudlet, the solution of (5) is a nonlinear optimization problem. However, the result of all cloudlets is a matrix $X = (x_1^T, x_2^T, \ldots, x_M^T)^T \in R^M$, in which $x_1 = (x_{1,1}, x_{1,2}, \ldots, x_{1,M})$ is the optimal solution of a cloudlet L_1. The result of one cloudlet may have an effect on other connected cloudlets for resource sharing. For example, if $x_{1,2} > 0$, this means that a cloudlet L_1 rents resource from L_2. Therefore, $x_{1,2}$ is the amount of trading resource between two cloudlets. Moreover, before allocating resource, we first evaluate the occupied resource of L_i by

$$\rho_i = \varphi_i + \sum_{k=1,k\neq 1}^{M} x_{i,k}. \tag{7}$$

Here, φ_i is the resource utilization of L_i in the last time slot. We try to maximize the utility of each cloudlet based on the allocation function, since the allocation function has the same tendency as the utility of each cloudlet. This means that if we maximize the value of the allocation function, we obtain the best utility of the cloudlet. The strategies of other cloudlets are denoted by $\overline{x_i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_M)$. The optimal function of a cloudlet i can be denoted by

$$\begin{aligned} \max_{x_i} F(x_i, \overline{x_i}) &= (1 - (x_{i,i} + \rho_i)^m) \\ &\quad + \chi_{i,k} \sum_{k=1, k\neq i}^{M} (1 - (x_{i,k} + \rho_k)^m) \\ \text{s. t.} \quad \sum_{k=1}^{M} x_{i,k} &= \pi_i \\ x_{i,k} &\geq 0 (k = 1, 2, \ldots, M). \end{aligned} \tag{8}$$

IV. Solution

The optimal function of all cloudlets is too complicated to be resolved as they are tightly coupled by the network topology. Equation (8) leads us to a local optimal solution of one cloudlet, in which we find that the utilization of cloudlets is closely related with each other and can be balanced by resource sharing. Therefore, the globally optimal solution is based on the SDN architecture to maximize the allocation function of the entire cloud system. The main work of this section is to linearize the optimal function and get the explicit form of the solution.

There are two main difficulties for this problem. First, the solution of (8) is a set of value, $x_1 = (x_{1,1}, x_{1,2}, \ldots, x_{1,M})$. It includes the resource sharing from other cloudlets, such as $(x_{1,2}, \ldots, x_{1,M})$, and the resource allocation on its own devices, $x_{1,1}$. The traditional concave optimization approach is to construct the entropy function here and search for the optimal solution in a feasible zone. However, this approach is time consuming when there are many cloudlets in a network. We employ a simplified linear approach, namely, KKT condition, to find the optimal solution. Second, This problem has special properties. All cloudlets in EC-RAN try to balance the utilization of a network. Thus, global optimization can be obtained when cloudlets achieve a local optimal solution. This means that the globally optimal function is equal to the sum of all local optimal functions. The study in [32] proposes a tree topology approach for the load balance problem in a homogeneous network. However, the operation time of a tree formation will grow as the number of nodes increases, whereas the accuracy of a global solution will decrease. To address the problem, we propose a noncooperation matrix game to obtain global optimization, which is to motivate each cloudlet to optimize resource management.

A. Matrix Game

We consider a noncooperative matrix game with M players. For each player i, it obtains the optimal solution x_i based on other player's strategies $\overline{x_i}$. Thus, we regard this approach as the same as solving a generalized Nash equilibrium problem. To find the Nash equilibrium solution, we denote X^* as the generalized Nash equilibrium, $X^* = \{x_1^*, x_2^*, \ldots, x_n^*\}$, and $x_1^* = \{x_{1,1}^*, x_{1,2}^*, \ldots, x_{1,n}^*\}$. According to KKT conditions, the Nash equilibrium solution is

$$\begin{aligned} \max_X F(X^*) &= \sum_{i=1}^{M} \Bigg(\left(1 - \left(x_{i,i}^* + \rho_i\right)^m\right) \\ &\qquad + \sum_{k=1}^{M} \chi_{i,k} \left(1 - \left(x_{i,k}^* + \rho_k\right)^m\right) \Bigg) \\ \text{s. t.} \quad g(x_i) &= \sum_{k=1}^{M} x_{i,k}^* - \pi_k = 0 \\ G(x_i) &= x_{i,k}^* - \upsilon_k \geq 0. \end{aligned} \tag{9}$$

In [33], Li and Lin employed the KKT condition to transform the problem of generalized Nash equilibrium into nonlinear complementarity. To get the Nash equilibrium solution, the resource-sharing function needs to satisfy one condition: To all cloudlets, solutions $x_i \ (i = 1, \ldots, M)$, $F : R^M \mapsto R$ are concave functions. Thus

$$\begin{aligned} \frac{\partial F(X)}{\partial x_{i,k}} &= -\chi_{i,k}(x_{i,k} + \rho_k)^{m-1} \\ \frac{\partial^2 F(X)}{\partial {x_{i,k}}^2} &= -2\chi_{i,k} \ \ (\text{if } \ m = 2). \end{aligned} \tag{10}$$

Therefore, $F(X)$ is concave in each $x_{i,k}$. If x_i is a normalized stationary point of the optimal problem, there exist multipliers λ_i, β_i such that for each $i = 1, 2, \ldots, M$, we have

$$\begin{aligned} &-\nabla_{x_i} F(x_i, \overline{x_i}) + \nabla_{x_i} g(x_i)\lambda_i + \nabla_{x_i} G(x_i)\beta_i = 0 \\ &\lambda_i > 0 \perp g(x_i)\lambda_i = 0 \\ &\beta_i \geq 0 \perp G(x_i)\beta_i = 0. \end{aligned} \tag{11}$$

However, when the number of cloudlets is larger than two, (11) is very complicated to solve and get an explicit expression. We notice that a Nash equilibrium solution will be obtained when we solve (8) for each cloudlet. Therefore, we first simplify the optimal function (8). Then, we calculate the local optimal solution of each cloudlet independently. After iteration for several rounds, the local optimal solution will gradually converge to the globally optimal solution, which is the Nash equilibrium. We refer to the following steps to obtain the global solution.

- **Step 1:** Simplify the optimal function (8) of each cloudlet. Then, calculate the related parameter ρ_i^n for each cloudlet according to (7), where n is the index of rounds. Go to step 2.
- **Step 2:** Calculate the local optimal solution of each cloudlet. Go to step 3.
- **Step 3:** Compare solution X^{n-1} with current solution X^n. If $|X^n - X^{n-1}| < \varphi$, X^n is the final solution. Otherwise, go back to step 1. φ is a positive value.

B. Simplification

The principle of simplification is based on the idea of excluding the cloudlets that are inappropriate to rent resource. First, cloudlets without edge connection are not capable of resource sharing. Take L_1 for example. L_1 and L_6 have no connected edge, $e(L_1, L_6) = 0$. Thus, $x_{1,6} = 0$. Then, we notice that a cloudlet L_i only sends a resource sharing request to the connected cloudlets whose unoccupied resource is richer than itself. If a cloudlet L_1 rents resource from L_2, the resource allocation result includes occupied resource $x_{1,1}$ and renting resource $x_{1,2}$, i.e.,

$$x_{1,1} = \frac{\chi_{1,2}(\rho_2 + \pi_1) - \rho_1}{1 + \chi_{1,2}}$$
$$x_{1,2} = \begin{cases} 0, & \text{if } \chi_{1,2}\rho_2 - \rho_1 \geq \pi_1 \\ 1 - \frac{\chi_{i,k}(\rho_2+\pi_1)-\rho_1}{1+\chi_{1,2}}, & \text{if } \chi_{1,2}\rho_2 - \rho_1 < \pi_1. \end{cases} \quad (12)$$

The result shows that L_1 may rent resource from L_2 if the condition of $\chi_{1,2}\rho_2 - \rho_1 < \pi_1$ is satisfied. This condition is also correct in multiple-cloudlet resource sharing. Therefore, the resource sharing between two cloudlets is based on two conditions: They are directly connected $e(L_i, L_k) = 1$, and the amount of unoccupied resource is enough for resource sharing $\chi_{i,k}\rho_k - \rho_i < \pi_i$. Otherwise, the resource sharing request will be refused $x_{i,j} = 0$.

C. Concave Optimization in Each Cloudlet

We employ a nonlinear concave optimization approach based on KKT conditions to solve the local optimal solution of each cloudlet. From [34], we can work out an explicit solution for any matrix dimension M. Equation (5) can be rewritten as

$$L(x_i, \lambda_i, \beta_i, v_i) = -F(x_i) + \lambda_i \left(\sum_{k=1}^{n} x_{i,k} - \pi_i\right) + \sum_{k=1}^{n} \beta_k \left(x_{i,k} - v_k^2\right). \quad (13)$$

Here, λ_i, β_k, and v_k are related parameters in Lagrange decoupling. n is the number of cloudlets, satisfying condition $\chi_{i,k}\rho_k - \rho_i < \pi_i$. Then, we have a set of equations as follows:

$$\frac{\partial L}{\partial x_i} = -\frac{\partial F_{(x_i)}}{\partial x^i} + \frac{\partial \lambda_i \sum_{k=1}^{n} x_{i,k}}{\partial x^i} + \frac{\partial \sum_{k=1}^{n} \beta_k \left(x_{i,k} - v_k^2\right)}{\partial x^i} = 0 \quad (14)$$

$$\sum_{k=1}^{n} x_{i,k} - \pi_i = 0 \quad (15)$$

$$x_{i,k} \geq 0 \quad (16)$$

$$\beta_k x_{i,k} = 0. \quad (17)$$

The reformulation of this problem is a standard mathematical procedure and relatively easy to solve. The solution of cloudlet L_i will fulfill limitations from (14)–(17), no matter how many cooperating cloudlets there are. In this paper, we only elaborate the progress of resolving the case that L_i has two cooperating cloudlets. Take $n = 2$ for example. We combine (14) and (15) and then get the results as follows:

$$\lambda_i = -\frac{2\chi_{i,1}\chi_{i,2}\Delta + \chi_{i,1}\chi_{i,2}\beta_1 + \chi_{i,2}\beta_2 + \chi_{i,1}\beta_3}{\chi_{i,1}\chi_{i,2} + \chi_{i,1} + \chi_{i,2}} \quad (18)$$

$$x_{i,1} = \frac{2\chi_{i,1}\chi_{i,2}\Delta + \chi_{i,2}(\beta_2 - \beta_1) + \chi_{i,1}(\beta_3 - \beta_1)}{2(\chi_{i,1}\chi_{i,2} + \chi_{i,1} + \chi_{i,2})} - \rho_1 \quad (19)$$

$$x_{i,2} = \frac{2\chi_{i,2}\Delta + \chi_{i,2}(\beta_1 - \beta_2) + (\beta_3 - \beta_2)}{2(\chi_{i,1}\chi_{i,2} + \chi_{i,1} + \chi_{i,2})} - \rho_2 \quad (20)$$

$$x_{i,3} = \frac{2\chi_{i,1}\Delta + \chi_{i,1}(\beta_1 - \beta_3) + (\beta_2 - \beta_3)}{2(\chi_{i,1}\chi_{i,2} + \chi_{i,1} + \chi_{i,2})} - \rho_3 \quad (21)$$

$$\Delta = \rho_1 + \rho_2 + \rho_3 + \pi_i. \quad (22)$$

Here, Δ is the sum of utilization. Then, we combine (18) with (19)–(21) and (17) and get three equations about β_i, i.e.,

$$\left(\frac{2\chi_{i,1}\chi_{i,2}\Delta + \chi_{i,2}(\beta_2 - \beta_1) + \chi_{i,1}(\beta_3 - \beta_1)}{2(\chi_{i,1}\chi_{i,2} + \chi_{i,1} + \chi_{i,2})} - \rho_1\right)\beta_1 = 0 \quad (23)$$

$$\left(\frac{2\chi_{i,2}\Delta + \chi_{i,2}(\beta_1 - \beta_2) + (\beta_3 - \beta_2)}{2(\chi_{i,1}\chi_{i,2} + \chi_{i,1} + \chi_{i,2})} - \rho_2\right)\beta_2 = 0 \quad (24)$$

$$\left(\frac{2\chi_{i,1}\Delta + \chi_{i,1}(\beta_1 - \beta_3) + (\beta_2 - \beta_3)}{2(\chi_{i,1}\chi_{i,2} + \chi_{i,1} + \chi_{i,2})} - \rho_3\right)\beta_3 = 0. \quad (25)$$

Therefore, through calculating β_i, we can get a solution of x_i. According to (16), there is no less than one β_i that is equal to zero. Therefore, conditions can be classified into seven cases as $2^3 - 1 = 7$. If all β_i''s are equal to zeros, the result of x_i can be presented as

$$\begin{cases} x_{i,1} = \frac{2\chi_{i,1}\chi_{i,2}\Delta}{2(\chi_{i,1}\chi_{i,2}+\chi_{i,1}+\chi_{i,2})} - \rho_1 \\ x_{i,2} = \frac{2\chi_{i,2}\Delta}{2(\chi_{i,1}\chi_{i,2}+\chi_{i,1}+\chi_{i,2})} - \rho_2 \\ x_{i,3} = \frac{2\chi_{i,1}\Delta}{2(\chi_{i,1}\chi_{i,2}+\chi_{i,1}+\chi_{i,2})} - \rho_3. \end{cases} \quad (26)$$

If only β_1 is equal to zero $\beta_1 = 0$, this means that only x_1 is larger than zero. Therefore, we have the set of results $x_i = [\pi_i, 0, 0]$. If only β_1 is larger than zero $\beta_1 > 0$, this means that only x_1 is equal to zero, and we have

$$-(\chi_{i,1}+\chi_{i,2})\beta_1 = 2\rho_1(\chi_{i,1}\chi_{i,2} + \chi_{i,1} + \chi_{i,2}) - 2\chi_{i,1}\chi_{i,2}\Delta. \quad (27)$$

Moreover, the result of x_i can be presented by

$$\begin{cases} x_{i,1} = 0 \\ x_{i,2} = \frac{\chi_{i,2}\Delta(2\chi_{i,1}+\chi_{i,2})}{(\chi_{i,1}\chi_{i,2}+\chi_{i,1}+\chi_{i,2})(\chi_{i,1}+\chi_{i,2})} - \frac{\chi_{i,2}}{\chi_{i,1}+\chi_{i,2}}\rho_1 - \rho_2 \\ x_{i,3} = \frac{\chi_{i,1}\Delta(1+\chi_{i,1}\chi_{i,2})}{(\chi_{i,1}\chi_{i,2}+\chi_{i,1}+\chi_{i,2})(\chi_{i,1}+\chi_{i,2})} - \frac{\chi_{i,1}}{\chi_{i,1}+\chi_{i,2}}\rho_1 - \rho_3. \end{cases} \quad (28)$$

The rest of the four cases can be resolved similarly. Thus, we narrow down the searching area for solution x_i to eight points. According to (16), we can further remove some points out of the feasible zone of the solution. Then, we go back to (8) and find the final solution at the point with the maximum value of $F(x_i, \overline{x_i})$. The Kuhn–Tucker conditions are able to find the final solution of more than two cooperated cloudlets.

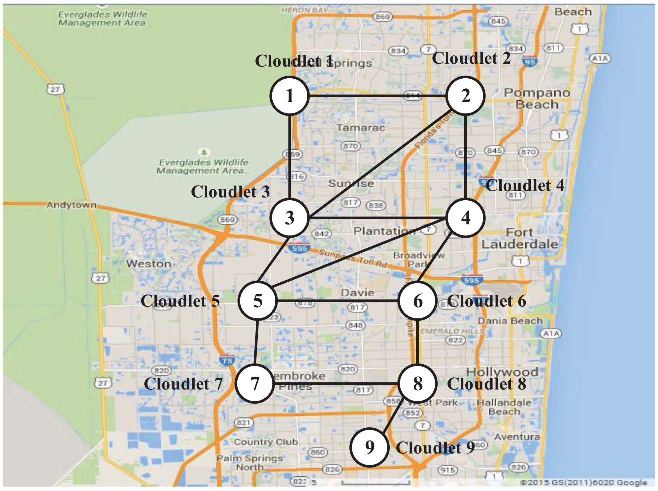


Fig. 4. Example of distributed cloudlets in southeast Florida.

Theorem 1: When we maximize the resource-sharing function, we will obtain the best network utility.

Proof: According to (4), the utility function is concave in $(\partial U/\partial y_i) = R + (u/(y_i+1))$ and $(\partial U^2/\partial^2 y_i) = -(u/(y_i+1))^2 < 0$. According to (10), the resource-sharing function is also concave. The resource sharing result is to balance the utilization between the different cloudlets. We assume a two-cloudlet network, whose utilization is denoted by y_1 and y_2 before resource sharing $(y_1 > y_2)$. By maximizing (9), the optimal solution $y_1' = y_2' = ((y_1+y_2)/2)$ is to balance the utility between cloudlets. Furthermore, the utility is compared between before and after resource sharing, i.e.,

$$\begin{aligned} U - U' &= u\left(\log(y_1+1) + \log(y_1+1) - 2\log(y_1'+1)\right) \\ &= u\left(\log\left(\frac{(y_1+1)(y_2+1)}{(y_1'+1)^2}\right)\right) \\ &< u\left(\log\left(\frac{2}{(y_1'+1)}\right)\right) < 0. \end{aligned} \tag{29}$$

The utility before resource sharing is less than that after resource sharing. Therefore, when we maximize the resource-sharing function, the utility of cloudlet will be maximized. ■

V. Numerical Results

A. Simulation Setup

In the simulation, the system model is supported by the travel data from the household travel data dictionary of the southeast Florida region [35].There are three countries and 18 regions in the southeast Florida, Broward, Dade, and Palm Beach. We select nine adjacent regions with a high traffic rate. The nine adjacent regions are shown in Fig. 4. We assume that each cloudlet is located in one region and responsible for local vehicular service. The connected lines between each region are the edges of a network and indicate geographical proximity. Cloudlets with connection can form a resource sharing group. The remote data center is responsible for managing the operation of all cloudlets. The simulation is conducted to last for 24 h.

TABLE I
Resource Requirement of AR Applications

Applications	CPU resource	Memory resource	Bandwidth resource
For navigation without D2D assist	5	5	3
For online video without D2D assist	4	4	4
For video and navigation without D2D assist	7	7	5
For navigation without D2D assist	5	4	2
For video and navigation with D2D assist	7	6	3

TABLE II
Parameter Setting

$\max C_i$	800	a	0.5	τ	1.5
$\max M_i$	800	b	0.4	m	2
$\max B_i$	480	c	0.1	M	9
R	5	δ_1	1	ν	0.7
u	3	δ_2	0.5	φ	0.0001

We assume that the capability of all cloudlets is the same, and the cloud resource is shared by service providers equally. We consider two kinds of vehicular applications: service for navigation and service for online video. Since a vehicle may have many passengers, the two kinds of vehicular applications can be conducted at the same time. The required resources are in Table I, in which resource values are transformed into the units of each kind of resource (e.g., one unit of CPU means 4000 MIPS, one unit of memory means 4000 MB). We adopt this approach to simplify the calculation as that in [36] and [37]. The ratios of three kinds of resource are evaluated by the cost, i.e., $c = 0.5$, $a = 0.4$, and $b = 0.1$. Other parameter settings can be found in Table II.

B. Convergence of the Proposed Approach

The optimal solution of each cloudlet will be obtained by maximizing (9). In the optimal progress, each cloudlet will interact with each other according to the network topology. Following the proposed approach, the global solution will be obtained through multiple iterations. Fig. 5 shows us the process in obtaining the global solution. Error E means the absolute difference between the current solution and the previous solution, $E = \sum_{k=1}^{M} |x_{i,k}^n - x_{i,k}^{n-1}|$. After several iterations, such as point $n = 15$ in Fig. 5, the local solution will converge to a stable state of the global optimal solution. This means that no cloudlet will adjust its resource allocation strategy to obtain a higher utility.

C. Performance Evaluation of D2D Communications

The D2D approach is an attractive characteristic of 5G networks. Vehicles can employ D2D to transmit data rather than D2B. Therefore, frequency reuse will enable a cloudlet to allocate less D2B bandwidth resource by using D2D bandwidth.

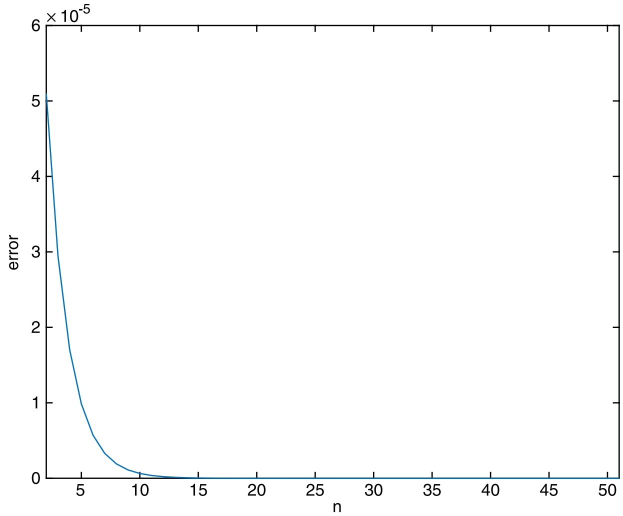


Fig. 5. Convergence of the KKT approach.

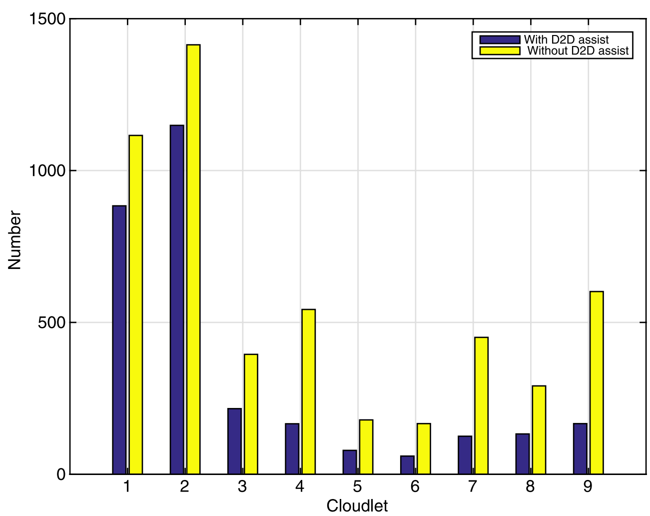


Fig. 6. Total number of dropped applications in 24 h.

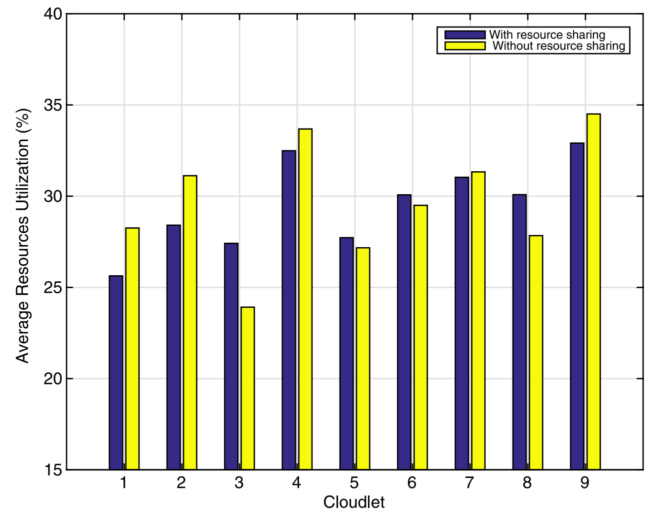


Fig. 7. Resource utilization of cloudlets at 8:00.

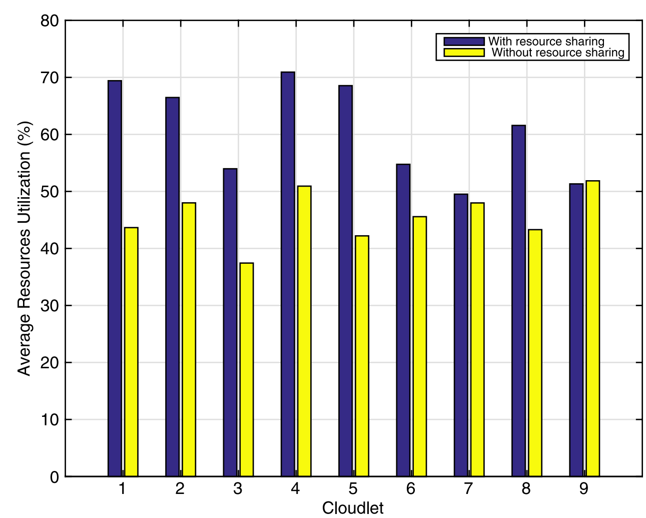


Fig. 8. Average resource utilization in a whole day.

Moreover, with less data transmission to mobile users, cloudlets can save memory resource, such as Read-Only Memory for task running. Fig. 6 compares the total number of dropped applications on each cloudlet. In this model, we assume that a user will continue to send an application request if they are refused. Therefore, one application may be dropped several times if a cloudlet has no enough resource to provide service. In Fig. 6, we notice that the number of dropped applications in a D2D-assisted network is less than that without D2D. In the case without D2D, each application will consume more resource. A cloudlet will refuse and drop user applications if it has not enough resource to provide service.

D. Performance Evaluation of Cloudlet Resource Sharing

Cloudlet resource sharing is an approach to improve the unbalance in resources utilization. It performs well in two situations: resource shortage and cloudlet hibernation. In the case of resource shortage, a cloudlet can enhance its capability by renting resource from nearby cloudlets. Through offloading application on the cooperated cloudlet, the number of mobile users in service is increasing. A cloudlet can be turned off in a state of hibernation for energy saving, if its amount of occupied resource is very low and it can be replaced by the nearby cloudlets. However, in the case without resource sharing, a cloudlet will keep operating, even if there is no user. Fig. 7 compares the resource utilization of each cloudlet at 8:00. The blue bars of cloudlets 1, 2, 4, 7, and 9 are lower than yellow bars. The rest of the blue bars are higher than the yellow bars. The total resource utilization of all cloudlets is the same. The results indicate that a cloudlet with a large amount of occupied resource can reduce its resource utilization through renting from the cloudlets with low occupation. Therefore, the resource utilization between different cloudlets will become more balanced. The utility of the two approaches (with and without resource sharing) is 5.65866 and 5.65882, respectively. This further illustrates that the proposed allocation strategy is optimal to network utility.

The average resource utilization of each cloudlet in a whole day is shown in Fig. 8. The utilization of blue bars is higher than that of the yellow bars. This is because the cloudlet that has less application requirements can improve its resource utilization

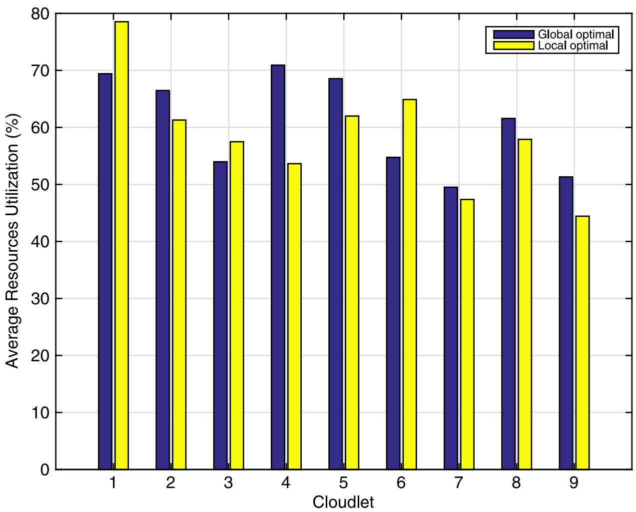


Fig. 9. Result of average resource utilization at 9:00 in a global optimal approach and a local optimal approach.

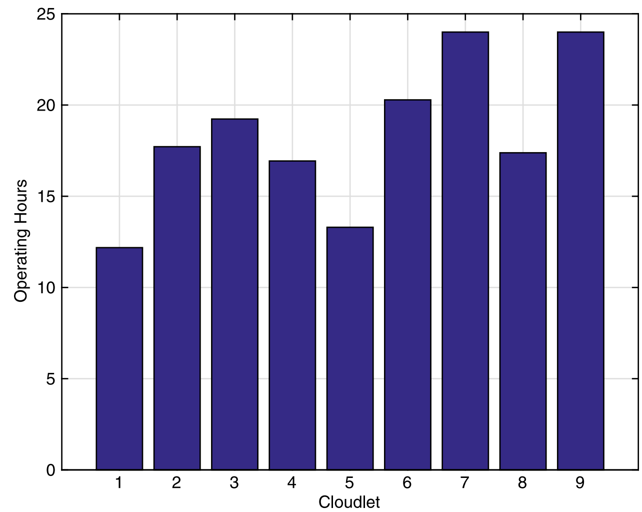


Fig. 10. Total operation time of nine cloudlets.

through the resource-sharing approach. The cloudlet that has abundant resource can process more application requests by resource sharing. The number of application requirements is affected by the number of mobile users. The result turns out that the total utilization is significantly enhanced.

E. Performance Evaluation of the SDN Architecture

The 5G-enabled vehicular network with SDN architecture has many benefits. One of them is that with an appropriate separation between control and management, cloudlets are controlled and managed by the control center in the management plane. This allows the cloud manager to control the network conveniently. Without SDN architecture, a cloudlet only has knowledge of nearby cloudlets, and the manager has to make the decision based on local information. Fig. 9 shows the result of average resource utilization at 9:00 and compares the results of two approaches: with and without SDN architecture (global and local optimality).

The balance of resource utilization is important to show the advantage of SDN architecture and can be evaluated according to the heterogeneity index. The heterogeneity index in this paper is based on the Lorenz curve and a Gini coefficient that describes income inequality in microeconomics [38]. Here, we employ the heterogeneity index H to evaluate the unbalance degree of resource utilization of cloudlets. Thus

$$H = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}|y_i - y_j|}{2N^2\bar{y}} \tag{30}$$

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i. \tag{31}$$

Here, $\bar{y}$ is the average value of utilization of all cloudlets. N is the number of cloudlets. The heterogeneity index of cloudlet utilization with and without SDN architecture is 0.1328 and 0.1765 respectively. This shows that the globally optimal solution is slightly better than the local solution of resource balance. Since the balanced result is better for utility, we come to the conclusion that with SDN architecture, the 5G-enabled vehicular network can reach the global optimality.

F. Performance Evaluation of Energy Cost

Fig. 10 shows the total operation time of nine cloudlets in the resource-sharing approach. Let us take the cloudlet L_1 for example. It has 12 h to turn on for work and 12 h to turn off for hibernation. The power-off duration is good for the lifetime of the cloudlet and also reduces the operation energy. If without resource sharing, the cloudlets need to keep on working. By combining with the result in Fig. 7, we conclude that the 5G-enabled vehicular network increases the utilization of resources with lower power consumption.

VI. Conclusion

In this paper, we have proposed a new paradigm of a 5G-enabled vehicular network. The paradigm integrates the concept of EC-RAN, D2D communications, and SDN technologies. The main purpose is to provide efficient and elastic services for mobile applications, which requires large bandwidth resource and high computing capability. We have also discussed the cloudlet resource management approach, which includes resource allocation and sharing. The D2D approach was investigated to increase the bandwidth resource reuse and relieve the communications congestion. We employed a matrix game to solve the problem and used KKT conditions to work out the explicit solutions of global optimization. Through cloudlet resource sharing, the management approach of noncooperation game can significantly improve the resource utilization and reduce the operation cost for the cloudlet devices. In future work, communications security is also important in cloud computing [39], [40]. Particularly for fast-moving vehicles, the cooperation between small cells is an effective approach to deal with the network topology change problem, such as rapid authentication of mobile terminals or fast switching between different wireless networks [41].

References

[1] P. Demestichas *et al.*, "5G on the horizon: Key challenges for the radio-access network," *IEEE Veh. Technol. Mag.*, vol. 8, no. 3, pp. 47–53, Sep. 2013.

[2] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[3] N. Bhushan *et al.*, "Network densification: The dominant theme for wireless evolution into 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 82–89, Feb. 2014.

[4] R. Wang, H. Hu, and X. Yang, "Potentials and challenges of C-RAN supporting multi-RATs toward 5G mobile networks," *IEEE Access*, vol. 2, pp. 1187–1195, Oct. 2014.

[5] J. Shen, H. Tan, J. Wang, J. Wang, and S. Lee, "A novel routing protocol providing good transmission reliability in underwater sensor networks," *J. Internet Technol.*, vol. 16, no. 1, pp. 171–178, 2015.

[6] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu, "Device-to-device communication in 5G cellular networks: Challenges, solutions, and future directions," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 86–92, May 2014.

[7] A.-H. Tsai, L.-C. Wang, J.-H. Huang, and T.-M. Lin, "Intelligent resource management for device-to-device (D2D) communications in heterogeneous networks," in *Proc. WPMC*, 2012, pp. 75–79.

[8] S. Patel, M. Chauhan, and K. Kapadiya, "5G: Future mobile technology-vision 2020," *Int. J. Comput. Appl.*, vol. 54, no. 17, pp. 6–10, 2012.

[9] K. Kumar and Y.-H. Lu, "Cloud computing for mobile users: Can offloading computation save energy?" *Computer*, vol. 43, no. 4, pp. 51–56, Apr. 2010.

[10] J. Li, D. Li, J. Zheng, and Y. Quan, "Location-aware multi-user resource allocation in distributed clouds," *Adv. Comput. Archit.*, vol. 451, pp. 152–162, 2014.

[11] M. Quwaider and Y. Jararweh, "Cloudlet-based efficient data collection in wireless body area networks," *Simul. Model. Pract. Theory*, vol. 50, pp. 57–71, 2015.

[12] R. Kaewpuang, D. Niyato, P. Wang, and E. Hossain, "A framework for cooperative resource management in mobile cloud computing," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, pp. 2685–2700, Dec. 2013.

[13] H. Kim and N. Feamster, "Improving network management with software defined networking," *IEEE Commun. Mag.*, vol. 51, no. 2, pp. 114–119, Feb. 2013.

[14] D. Kreutz *et al.*, "Software-defined networking: A comprehensive survey," *Proc. IEEE*, vol. 103, no. 1, pp. 14–76, Jan. 2015.

[15] A. Beloglazov and R. Buyya, "Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers," *Concurrency Comput., Pract. Exp.*, vol. 24, no. 13, pp. 1397–1420, Sep. 2012.

[16] S. Barbarossa, S. Sardellitti, and P. Di Lorenzo, "Joint allocation of computation and communication resources in multiuser mobile cloud computing," in *Proc. IEEE 14th SPAWC*, 2013, pp. 26–30.

[17] D. Kusic, J. O. Kephart, J. E. Hanson, N. Kandasamy, and G. Jiang, "Power and performance management of virtualized computing environments via lookahead control," *Cluster Comput.*, vol. 12, no. 1, pp. 1–15, Mar. 2009.

[18] S. Belhaiza, "On perfect Nash equilibria of polymatrix games," *Game Theory*, vol. 2014, pp. 1–11, 2014.

[19] K. C. Sharma, R. Bhakar, and H. Tiwari, "Extreme Nash equilibrium of polymatrix games in electricity market," in *Proc. Recent Adv. Innovations Eng.*, 2014, pp. 1–5.

[20] Z. H. Mir and F. Filali, "LTE and IEEE 802.11 p for vehicular networking: A performance evaluation," *EURASIP J. Wireless Commun. Netw.*, vol. 2014, no. 1, pp. 1–15, 2014.

[21] X. Wu *et al.*, "Vehicular communications using DSRC: Challenges, enhancements, and evolution," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 399–408, Sep. 2013.

[22] C.-X. Wang *et al.*, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 122–130, Feb. 2014.

[23] M. Eiza and Q. Ni, "An evolving graph-based reliable routing scheme for VANETs," *IEEE Trans. Veh. Technol.*, vol. 62, no. 4, pp. 1493–1504, May 2013.

[24] L. Q. Clara, N. Huaning, P. A. Tolis, and W. Geng, "5G network capacity: Key elements and technologies," *IEEE Veh. Technol. Mag.*, vol. 9, no. 1, pp. 71–78, Mar. 2014.

[25] M. Tinghuai *et al.*, "Social network and tag sources based augmenting collaborative recommender system," *IEICE Trans. Inf. Syst.*, vol. 98, no. 4, pp. 902–910, 2015.

[26] S. Caixing, L. Supeng, Z. Yan, A. Vinel, and M. Jonsson, "Performance analysis of connectivity probability and connectivity-aware MAC protocol design for platoon-based VANETs," *IEEE Trans. Veh. Technol.*, vol. 64, no. 12, pp. 5596–5609, Dec. 2015.

[27] H. A. Tran, A. Mellouk, and S. Hoceini, "QoE content distribution network for cloud architecture," in *Proc. 1st Int. Symp. NCCA*, 2011, pp. 14–19.

[28] D. T. Hoang, D. Niyato, and P. Wang, "Optimal admission control policy for mobile cloud computing hotspot with cloudlet," in *Proc. WCNC*, 2012, pp. 3145–3149.

[29] T. Soyata, R. Muraleedharan, C. Funai, M. Kwon, and W. Heinzelman, "Cloud-vision: Real-time face recognition using a mobile-cloudlet-cloud acceleration architecture," in *Proc. IEEE ISCC*, 2012, pp. 59–66.

[30] Y. Ren, S. Jun, W. Jian, H. Jin, and S.-Y. Lee, "Mutual verifiable provable data auditing in public cloud storage," *J. Internet Technol.*, vol. 16, no. 2, pp. 317–323, 2015.

[31] Q. Wang, S. Leng, H. Fu, and Y. Zhang, "An IEEE 802.11p-based multichannel MAC scheme with channel coordination for vehicular ad hoc networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 449–458, Jun. 2012.

[32] S. Xie and Y. Wang, "Construction of tree network with limited delivery latency in homogeneous wireless sensor networks," *Wireless Pers. Commun.*, vol. 78, no. 1, pp. 231–246, 2014.

[33] P. Li and G. Lin, "Solving a class of generalized Nash equilibrium problems," *J. Math. Res. Appl.*, vol. 33, no. 3, pp. 372–378, 2013.

[34] M. A. Hanson, "On sufficiency of the Kuhn–Tucker conditions," *J. Math. Anal. Appl.*, vol. 80, no. 2, pp. 545–550, Apr. 1981.

[35] Florida transportation modeling, 2014. [Online]. Available: http://www.fsutmsonline.net/Floridatravelsurveys/FL_travel_surveys.htm

[36] T.-H. Tram and C. Tham, "A game-theoretic model for dynamic pricing and competition among cloud providers," in *Proc. IEEE/ACM 6th Int. Conf. UCC*, 2013, pp. 235–238.

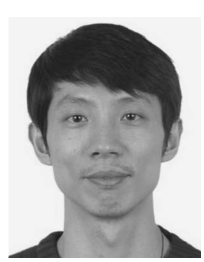
[37] H. Liang, L. X. Cai, D. Huang, X. Shen, and D. Peng, "An SMDP-based service model for interdomain resource allocation in mobile cloud networks," *IEEE Trans. Veh. Technol.*, vol. 61, no. 5, pp. 2222–2232, Jun. 2012.

[38] H.-B. Hu and X.-F. Wang, "Unified index to quantifying heterogeneity of complex networks," *Phys. A, Statist. Mech. Appl.*, vol. 387, no. 14, pp. 3769–3780, Jun. 2008.

[39] Z. Xia, X. Wang, X. Sun, and Q. Wang, "A secure and dynamic multi-keyword ranked search scheme over encrypted cloud data," *IEEE Trans. Parallel Distrib. Syst.*, vol. 27, no. 2, pp. 340–352, Feb. 2016.

[40] Z. Fu, X. Sun, L. Qi, and Z. Lu, "Achieving efficient cloud search services: Multi-keyword ranked search over encrypted cloud data supporting parallel computing," *IEICE Trans. Commun.*, vol. E98-B, no. 1, pp. 190–200, 2015.

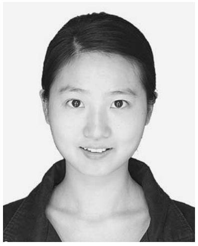
[41] P. Guo, J. Wang, X. H. Geng, C. S. Kim, and J.-U. Kim, "A variable threshold-value authentication architecture for wireless mesh networks," *J. Internet Technol.*, vol. 15, no. 6, pp. 929–935, 2014.



Rong Yu (S'05–M'08) received the Ph.D. degree from Tsinghua University, Beijing, China, in 2007.

After that, he was with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China. In 2010, he joined the Institute of Intelligent Information Processing, Guangdong University of Technology, Guangzhou, where he is currently a Full Professor. He is currently the Deputy Secretary General of the Internet of Things (IoT) Industry Alliance, Guangdong, China, and the Deputy Head of the IoT Engineering Center, Guangdong. He is a coinventor of over ten patents and has authored or coauthored over 100 international journal and conference papers. His research interest mainly focuses on wireless communications and mobile computing, including mobile cloud computing, vehicular networks, smart grid, cogtinive radio, etc.

Dr. Yu is a member of the Home Networking Standard Committee in China, where he leads the standardization work of three standards.



Jiefei Ding received the M.S. degree from Guangdong University of Technology, Guangzhou, China in 2015. She is currently working toward the Ph.D. degree with the University of Manitoba, Winnipeg, MB, Canada.

In 2015, she was a postgraduate visiting internship student for five months with the Hong Kong University of Science and Technology, Hong Kong. Her research interests include cloud computing resource management, software-defined network, and cloud radio access networks in fifth-generation networks.

Xumin Huang is currently working toward the Master's degree in networked control systems with Guangdong University of Technology, Guangzhou, China.

His research interests mainly focus on network performance analysis and simulation and enhancement in wireless communications and networking.

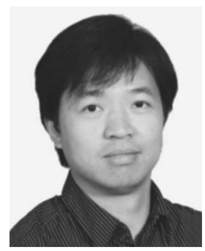
Ming-Tuo Zhou (SM'11) received the Ph.D. degree in telecommunications from the Asian Institute of Technology, Khlong Luang, Thailand, in 2003.

In July 2015, he joined Shenyang Institute of Automation Guangzhou Branch, Chinese Academy of Sciences. From July 2004 to June 2015, he was a Senior Research Scientist with the Smart Wireless Laboratory, National Institute of Information and Communications Technology Singapore Representative Office. His current research interests include industrial wireless sensor networks and wireless smart utility networks.

Mr. Zhou was the Technical Coeditor of IEEE 802.16n and IEEE 802.16.1a and a voting member and a technical contributor of the IEEE 802.11, 802.15, and 802.16 Working Groups. He received a Prize from the Finnish Universities of Technology and Helsinki Consulting Group and the IEEE-SA Standards Board Acknowledges with Appreciation Award for contributions to IEEE Std 802.16n, IEEE Std 802.16.1a, and IEEE Std 802.15.4m.

Stein Gjessing (M'91) received the Dr. Philos. degree from the University of Oslo, Oslo, Norway, in 1985.

He is an Adjunct Researcher with the Simula Research Laboratory and a Professor of computer science with the Department of Informatics, University of Oslo, where, from 1987 to 1990, he was the Head of the Department of Informatics. From February 1996 to October 2001, he was the Chairman of the national research program "Distributed IT-System" founded by the Research Council of Norway. He also participated in three European funded projects: Macrame, Arches, and Ascissa. His current research interests include routing, transport protocols, and wireless networks, including cognitive radio and smart-grid applications.



Yan Zhang (SM'10) received the Ph.D. degree from the School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore in 2015.

He is currently the Head of Department with the Department of Networks, Simula Research Laboratory, Norway, and an Associate Professor (part-time) with the Department of Informatics, University of Oslo, Oslo, Norway. His current research interests include wireless networks and reliable and secure cyberphysical systems (e.g., healthcare, transport, and smart grids).

Dr. Zhang is an Associate Editor or on the Editorial Board of a number of well-established scientific international journals, e.g., *Wiley Wireless Communications and Mobile Computing*. He also serves as the Guest Editor for the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, the IEEE COMMUNICATIONS MAGAZINE, the IEEE WIRELESS COMMUNICATIONS, and the IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING. He serves as the Chair in a number of conferences, including the 2016 IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications; the 2016 IEEE Consumer Communications and Networking Conference; the 2016 International ICST Wireless Internet Conference; the 2015 IEEE International Conference on Smart Grid Communications; and the 2015 IEEE International Conference on Cloud Computing Technology and Science. He serves as a Technical Program Committee Member for numerous international conferences, including the IEEE Conference on Computer Communications, the IEEE International Conference on Communications, the IEEE Global Telecommunications Conference, and the IEEE Wireless Communications and Networking Conference. He has received seven Best Paper Awards. He is a Senior Member of the IEEE Communications Society and the IEEE Vehicular Technology Society.