Contract Design for Traffic Offloading and Resource Allocation in Heterogeneous Ultra-Dense Networks

Jun Du, *Student Member, IEEE*, Erol Gelenbe, *Life Fellow, IEEE*, Chunxiao Jiang, *Senior Member, IEEE*, Haijun Zhang, *Senior Member, IEEE*, and Yong Ren, *Senior Member, IEEE*

***Abstract*— In heterogeneous ultra-dense networks (HetUDNs), the software-defined wireless network (SDWN) separates resource management from geo-distributed resources belonging to different service providers. A centralized SDWN controller can manage the entire network globally. In this paper, we focus on mobile traffic offloading and resource allocation in SDWN-based HetUDNs, constituted of different macro base stations and small-cell base stations (SBSs). We explore a scenario where SBSs' capacities are available, but their offloading performance is unknown to the SDWN controller: this is the *information asymmetric* case. To address this asymmetry, incentivized traffic offloading contracts are designed to encourage each SBS to select the contract that achieves its own maximum utility. The characteristics of large numbers of SBSs in HetUDNs are aggregated in an analytical model, allowing us to select the *SBS types* that provide the off-loading, based on different contracts which offer rationality and incentive compatibility to different SBS types. This leads to a closed-form expression for selecting the SBS types involved, and we prove the monotonicity and incentive compatibility of the resulting contracts. The effectiveness and efficiency of the proposed contract-based traffic offloading mechanism, and its overall system performance, are validated using simulations.**

***Index Terms*— Traffic offloading, software defined wireless networks (SDWNs), contract theory, heterogeneous ultra-dense networks (HetUDNs), resource sharing.**

I. Introduction

FIFTH Generation (5G) cellular networks were first proposed to meet the increasing mobile data traffic, which will expand one thousand times from 2010 to 2020 [1]–[3]. To meet this increasing data challenge, ultra-densification, i.e., overlaying macro base stations (MBSs) with a large number of small-cell base stations (SBSs) such as pico base stations (BSs), femto BSs and WiFi hotspots, etc., which constitute heterogeneous networks (HetNets), is one of the "big three" 5G technologies [4]. With assistance of these SBSs, *mobile traffic offloading* technology provides a solution to address the enormous expansion of mobile data, by moving traffic load from cellular networks to alternative wireless networks consisting of densely distributed SBSs. To operate such heterogeneous ultra-dense networks (HetUDNs), an effective and efficient network architecture and resource management mechanisms are indispensable. In recent years, cloud-based Software-defined Wireless Networks (SDWNs) are proposed to control and manage HetUDNs in a central manner efficiently. SDWNs can potentially revolutionize network design and resource management, and enable the applications to manipulate various services by separating the control plane from the data plane [5], [6]. In SDWNs, mobile traffic offloading can be enabled by the SDWN at the edge [7], which can exploit knowledge of the data requests and the network resource status of MBSs and SBSs. With a centralized controller, resources in HetUDNs can be managed efficiently to meet data requests from mobile users, and optimize system performance including data rate, load balancing and energy consumption [8], [9].

Recently, mobile traffic offloading in HetNets received significant attention for its effectiveness on rescuing the heavy traffic load in cellular networks by switching and exchanging traffic, and using access control and compatibility protocols [10]. Focusing on energy consumption optimization [11], [12], security guarantees [13] and performance analysis [14], much work has paid attention to mechanism design for mobile traffic offloading. As in [5], resource management in SDWNs is a form of competitive market, where resource requesters and providers compete and cooperate to maximize their own utilities. For traffic offloading in SDWN-based HetUDNs, competition and cooperation among resource providing and utilizing entities can be modeled and analyzed through economics theory [15]. Game theory is used to model the supply and demand relationship of resources for traffic offloading in HetNets, and many different game theory based offloading approaches have been applied, including Nash bargaining game [16], [17], coalition game [18], Stackelberg game [19], etc. In HetNets with densely distributed BSs, the computational complexity of such approaches grows exponentially, so that mean-field games can provide low-complexity tractable partial differential equation based solutions for traffic offloading [20]–[22].

Manuscript received April 1, 2017; revised September 12, 2017; accepted September 25, 2017. Date of publication October 6, 2017; date of current version December 1, 2017. This work was supported in part by the National Natural Science Foundation China projects under Grant 61371079, Grant 61471025, and Grant 91338203, in part by the Imperial College by Engineering and Physical Sciences Research Council under Grant EP/C548051/1, in part by the Young Elite Scientist Sponsorship Program by CAST under Grant 2016QNRC001, and in part by the Fundamental Research Funds for the Central Universities. *(Corresponding author: Chunxiao Jiang.)*

J. Du and Y. Ren are with the Department of Electronics Engineering, Tsinghua University, Beijing 100084, China (e-mail: blgdujun@gmail.com; reny@tsinghua.edu.cn).

E. Gelenbe is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: e.gelenbe@imperial.ac.uk).

C. Jiang is with the Tsinghua Space Center, Tsinghua University, Beijing 100084, China (e-mail: jchx@tsinghua.edu.cn).

H. Zhang is with the Beijing Engineering and Technology Research Center for Convergence Networks and Ubiquitous Services, University of Science and Technology Beijing, Beijing 100083, China (e-mail: haijunzhang@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/JSAC.2017.2760459

Auction theory is another important tool in network economics to model resource supply and demand, especially for networks with heterogeneous transmission resource [23]. Analytical work in this area includes the analysis of price and income in single [24] and networked multiple auctions [25], and for sealed bids can be found in [26]. In [27]–[29], the authors focus on the effect of an auction on the success of bidders who wish to make optimal choices. Auctions can also be used in a network to automatically carry out objectives which may satisfy either the sellers or the buyers [30], and in resource allocation for admission control based on resource availability [31]. In SDWN-based HetNets, the central controller can act as an auction broker, and traffic offloading can be operated efficiently with appropriate auction mechanisms, such as "double" [32] and reverse auctions [33].

In SDWN-based HetUDNs, despite the existence of a central controller, SBSs are selfish and may hide or fabricate the status of their resource. However, the SBSs' resource availability can be recognized by the central controller, although their offloading performance may be unknown. To offer an incentive compatibility, contract theory can be introduced to the traffic offloading mechanism designs. Contract theory, as a powerful microeconomics framework, is proposed to essentially deal with the information asymmetry in the market, regarding the service capability of "employees" which cannot be observed by "employers" before they are employed. According to contract theory, an incentive mechanism that encourages every employee to consciously choose the contract designed for its service capability will be realized. This approach has already been employed in resource allocation problems for device-to-device communications [34], heterogeneous Long-term Evlution-Advanced (LTE-A) networks [35] and heterogeneous cloud-based radio access networks [36]. Classic contract theory is based on the definition of different *employee types*, which is only considered as an abstract index without any specific definition in the aforementioned studies. Especially in HetUDNs with a large number of SBSs, this broad definition of "SBS type" can cause difficulties when contract models are applied to the real network environment. Thus in this work, we pay special attention to the SBS types and investigate how they can affect the performance of a contract-based traffic offloading mechanism.

The rest of this paper is organized as follows. Section II describes the SDWN framework for resource sharing in HetUDNs. The contract formulation and three contract-based traffic offloading mechanisms are designed in Section III and Section IV, respectively. Conditions for contract feasibility are analyzed and derived in Section V. Simulations are presented in Section VI, and Section VII summarizes our conclusions.

II. Architecture of SDWN

SDWN is an emerging network framework which separates the control plane from the data plane. The architecture of the SDWN-based resource sharing system of HetUDNs is shown as Fig. 1. In the resource and application level, the network provides data services with distributed MBSs and SBSs. These heterogeneous BSs are operated by the same or different operators (service providers) and deployed with a high density, which means that their coverage areas are overlapped seriously. The MBSs' mobile user equipments (MUEs) and small cell user equipments (SUEs) are randomly distributed in the coverage of the BSs. Through traffic offloading, the throughput and other performance of the system can be improved.

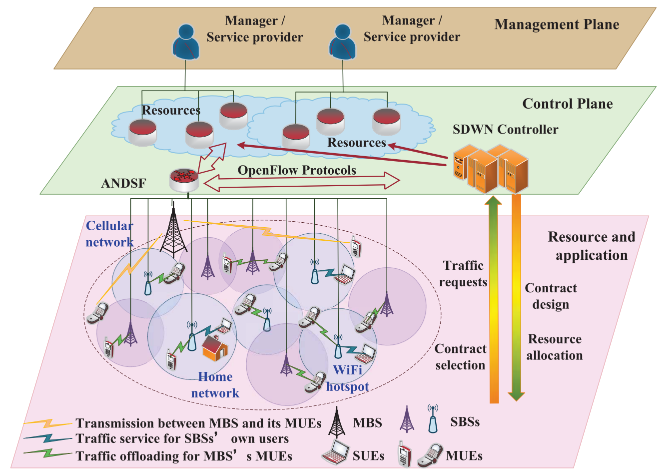


Fig. 1. Traffic offloading and resource allocation for SDWN-based HetUDNs.

As shown in Fig. 1, the SDWN separates resource management from the geo-distributed resource cloud, which forms a virtual network topology in the control plane. In the control plane, the centralized *SDWN Controller* discovers the traffic demands of MUEs, available transmission resource and the channel status in of the HetUDN through the *Access Network Discovery and Selection Function (ANDSF)*. The ANDSF fulfils this mission above by requiring to the MBS and SBSs, the current LTE/cellular network operators of which are more than willing to share the status information above to maximize their service capability and resource utilization. After receiving the supply and demand status of network resource, the SDWN controller designs a bundle of contracts for different types of SBSs, and then broadcasts the contracts to the MBS and all SBSs through the ANDSF. Every SBS distributed within the coverage of the MBS selects one contract to maximize its own payoff, and reports its selection to the SDWN controller that it will provide the certain amount of traffic offloading and get the certain payments from the MBS specified by the selected contract. According to the SBS's contract selection, Then the SDWN controller allocates this SBS's bandwidth resource for MUEs covered by this SBS, and requests the MBS to pay the SBS for its offloading service. During the process above, the ANDSF performs as a medium for the information interaction between the HetUDN and the SDWN Controller, and an executive of resource allocation. The ANDSF can interact with the SDWN controller for traffic offloading and resource allocation by standardized interfaces such as OpenFlow-enable switches [7], [15], which need some corresponding modifications for the requested and released information mentioned above. Moreover, in this work, we focus on the bandwidth resource allocation of the heterogeneous SBSs in the system.

III. Contract Formulation for Traffic Offloading

Consider an SDWN-based HetUDN with one MBS and a number of SBSs randomly distributed in the coverage of the MBS. These SBSs are not owned by the MBS operator, which means that the MBS cannot obtain the local information, such as transmission capacity, load status, operation and offloading cost, etc., of these SBSs. This model is flexible to be applied into a system with multiple MBSs, in which all SBSs can be associated to their respective MBS according to a certain association strategy, and then each subsystem consisting of an MBS and its associating SBSs can be analyzed by the system model of this work. We denote with $\mathcal{N} = \{1, 2, \cdots, N\}$ as the set of SBSs. Consider a set of $\mathcal{M} = \{1, 2, \cdots, M\}$ MUEs who are randomly distributed within the coverage of the MBS. In addition, let $\mathcal{M}_n$ be the set of MUEs in the coverage of SBS $n \in \mathcal{N}$, then we have $\bigcap_{n\in\mathcal{N}} \mathcal{M}_n = \mathcal{M}$. Let $\mathcal{N}_i$ denote the set of SBSs who can cover MUE $i \in \mathcal{M}$, which means that SBS $n \in \mathcal{N}_i$ can provide the traffic offloading service for MUE i.

Assume that the time is slotted. During the duration of a time slot, the location of MUEs, the offloading decision of SBSs and resource allocation are considered to be fixed. Denote $\mathbf{s}_n = \{s_{ni}\}_{i\in\mathcal{M}_n}$ as the scheduling vector, where $s_{ni} = 1$ indicates that the traffic of MUE i is allocated to be offloaded by SBS n, and $s_{ni} = 0$, otherwise. Assume that each MUE can be associated with at most one SBS, i.e., $\forall i \in \mathcal{M}$, $\sum_{n\in\mathcal{N}} s_{ni} \leq 1$. The case of $\sum_{n\in\mathcal{N}} s_{ni} = 0$ indicates that the traffic requested by MUE i is not offloaded by any SBS and is delivered by the MBS directly. Let $s_{0i} = 1$ denote that MUE i is served by the MBS without any SBS offloading for it.

A. Transmission Model Formulation

The transmission data rate can be used to evaluate the performance of the HetUDN, and is related to the signal-to-interference-plus-noise ratio (SINR). In this work, we model the channel between MUEs and BSs as a Rayleigh fading channel. Then $\forall n \in \mathcal{N}$ and $i \in \mathcal{M}$, the SINR from SBS n and the MBS to MUE i is defined as

$$\gamma_{ni} = \frac{p_n |h_{ni}|^2}{\sum_{n'\in\mathcal{N}_i} \kappa_{n'i}\, p_{n'} |h_{n'i}|^2 + \kappa'_{0i}\, p_0 |h_{0i}|^2 + \sigma^2}, \quad \text{(1a)}$$

$$\gamma_{0i} = \frac{p_0 |h_{0i}|^2}{\sum_{n\in\mathcal{N}} \kappa'_{ni}\, p_n |h_{ni}|^2 + \sigma^2}, \quad \text{(1b)}$$

respectively. In (1a) and (1b), σ^2 is the constant addictive noise power, while p_n and $p_{n'}$ are the transmission power consumption of SBS n and n', respectively. $\kappa_{n'i} \in [0, 1]$ is the interference parameter among SBSs, and κ'_{0i}, $\kappa'_{ni} \in [0, 1]$ are the interference parameters between the MBS and SBSs. Considering different licensed spectrum applied for direct transmission by the MBS and traffic offloading by SBSs, the interference between the MBS and SBSs can be ignored, i.e., $\kappa'_{0i} = \kappa'_{ni} = 0$. Then considering the channel allocation, the achievable service rate for MUE i can be presented by

$$r_i = \omega_0 s_{0i} \log(1 + \gamma_{0i}) + \sum_{n\in\mathcal{N}} \omega_n s_{ni} \log(1 + \gamma_{ni}), \quad (2)$$

where ω_n and ω_0 are the bandwidths of spectrum used by SBS n and the MBS. Take LTE-A for instance, the bandwidth for one resource block is $\omega = 180$ kHz. Considering all SBSs in the network utilize the common spectrum for traffic offloading, then let $\omega_n = \omega$, $\forall n \in \mathcal{N}$. Let y_n denote the traffic offloading accepted by or allocated to SBS n, i.e., $y_n = \sum_{i\in\mathcal{M}_n} r_{ni} s_{ni}$.

B. Economic Models Formulation

The offloading quality provided by heterogeneous SBSs is different. On the other hand, the benefit for the MBS from different SBSs is different as well. For instance, the MBS tends to get much more benefit from SBSs those located closed to the edge of MBS's coverage. Therefore, it is better for the SDWN controller to design diverse contracts for the heterogeneous BSs, to improve the performance of the HetUDN.

1) Utility of MBS: Let $T_n(y_n)$ be the payment for SBS n when it helps to offload the amount of y_n traffic. Assume $T_n(0) = 0$, and in addition, $T_n(y_n)$ is a strictly increasing function of y_n, $\forall n \in \mathcal{N}$. Then we define the utility of MBS as

$$U(\mathbf{s}, \mathbf{y}, \mathbf{T}(y)) = \delta \sum_{i\in\mathcal{M}} r_i - \sum_{n\in\mathcal{N}} T_n(y_n), \quad (3)$$

where $\mathbf{s} = \{s_{0i}, s_{ni}\}_{(N+1)\times M}$ denotes the association matrix, $\mathbf{y} = \{y_n\}_{n\in\mathcal{N}}$ is the traffic offloading vector, $\mathbf{T}(y) = \{T_n(y)\}_{n\in\mathcal{N}}$ denotes the vector of payment bundles for different types of SBSs, and δ is the MBS's unit monetary gain through the traffic rate.

2) Utility of SBSs: Let x_n $(n \in \mathcal{N})$ denote the SBS n's own traffic demands. Assume that traffic requests arrival for different SBSs are independent and identically distributed (i.i.d.), and follows a probability distribution function $f_n(x)$. In this work, we consider the traffic requests are sequences of Poisson arrivals with arrival rate λ_n, $\forall n \in \mathcal{N}$.

We define ψ_n, transmission efficiency of SBS n, as the average amount of data traffic (bits) can be delivered by one unit of bandwidth resource per time unit, which is given by

$$\psi_n = \frac{\sum_{i\in\mathcal{M}_n} r_{ni}}{\omega_n \left\| \{r_{ni}\}_{i\in\mathcal{M}_n} \right\|_0}, \quad (4)$$

where $r_{ni} = \omega_n \log(1 + \gamma_{ni})$ denotes the achievable data rate of MUE i receiving from SBS n, and $\|\cdot\|_0$ calculates the number of non-zero elements. Then the average bandwidth resource consumption for SBS n on delivering one unit traffic is $1/\psi_n$.

Let $\mathbf{\Omega}_n$ denote the resource capacity of SBS n. Then we have $y_n \leq \mathbf{\Omega}_n \psi_n$, $\forall n \in \mathcal{N}$. In addition, denote $w_n > 0$ as SBS n's average revenue achieved from one unit of its bandwidth resource consumption caused by its own traffic demands. Let c_n $(0 < c_n < w_n)$ represent SBS n's average cost on one unit of bandwidth utilization. Then the expected revenue of SBS n resulting from serving its own traffic demands is given by

$$\begin{aligned} &P_n(\mathbf{\Omega}_n) \\ &= (w_n - c_n)\, E(x_n/\psi_n) \\ &= (w_n - c_n) \left[\int_0^{\mathbf{\Omega}_n \psi_n} \frac{x}{\psi_n} f_n(x)\, dx + \int_{\mathbf{\Omega}_n \psi_n}^{\infty} \mathbf{\Omega}_n f_n(x)\, dx \right] \\ &= a_n \left(1 - e^{-\frac{\mathbf{\Omega}_n \psi_n}{\lambda_n}}\right), \end{aligned} \quad (5)$$

where $a_n = \lambda_n (w_n - c_n)/\psi_n$, and $f_n(x) = \lambda_n^{-1} e^{-\lambda_n^{-1} x}$.

Furthermore, given feasible amount of traffic offloading by SBS n, the expected revenue from the rest bandwidth resource for serving this SBS's own traffic demands can be obtained as

$$P_n\left(\Omega_n - y_n/\psi_n\right) = a_n\left(1 - e^{-\frac{\Omega_n\psi_n - y_n}{\lambda_n}}\right). \tag{6}$$

Then the utility of SBS n from traffic offloading is given by

$$V_n = P_n\left(\Omega_n - y_n/\psi_n\right) + T_n\left(y_n\right) - c_n y_n/\psi_n, \quad \forall n \in \mathcal{N}. \tag{7}$$

In addition, we define the net utility of SBS n as the SBS utility improvement when offloading traffic for the MBS:

$$V'_n = V_n - P_n\left(\Omega_n\right), \quad \forall n \in \mathcal{N}. \tag{8}$$

In (7), we assume that the total revenue of SBS n:

$$\upsilon_n\left(y_n\right) = P_n\left(\Omega_n - y_n/\psi_n\right) + T_n\left(y_n\right) \tag{9}$$

is a strictly increasing concave function of y_n, i.e., $\upsilon'\left(y_n\right) > 0$, and $\upsilon''\left(y_n\right) < 0$. This setting is reasonable, due to that as the amount of offloading traffic increasing, payment $T\left(y\right)$ from the MBS increases slowly, and meanwhile the income brought to the SBSs grows slowly, which also results from less service for SBSs' own traffic requests. This property of revenue function will be further analyzed in Section V.

3) Social Welfare: The social welfare of HetUDN is defined as the aggregate utility of the MBS and SBSs, denoted by

$$\begin{aligned} W &= \sum\nolimits_{n\in\mathcal{N}} U_n + \sum\nolimits_{n\in\mathcal{N}} V_n \\ &= \sum\nolimits_{n\in\mathcal{N}}\left[\delta\sum\nolimits_{i\in\mathcal{M}_n} r_i - T_n\left(y_n\right)\right] \\ &\quad + \sum\nolimits_{n\in\mathcal{N}}\left[P_n\left(\Omega_n - \frac{y_n}{\psi_n}\right) + T_n\left(y_n\right) - \frac{c_n y_n}{\psi_n}\right] \\ &= \underbrace{\delta\sum\nolimits_{i\in\mathcal{M}} r_i}_{\substack{\text{MBS: Profit from}\\ \text{MUEs' throughput}}} + \underbrace{\sum\nolimits_{n\in\mathcal{N}} P_n\left(\Omega_n - \frac{y_n}{\psi_n}\right)}_{\substack{\text{SBS: Profit from serving}\\ \text{its own traffic demands}}} - \underbrace{\sum\nolimits_{n\in\mathcal{N}} \frac{c_n y_n}{\psi_n}}_{\substack{\text{SBS: Cost}\\ \text{of offloading}}}. \end{aligned} \tag{10}$$

IV. Contract Design for Traffic Offloading

According to contract theory, a reasonable definition of SBS's type is very important to realize the contract-based traffic offloading. So first of all, we propose a new definition of the *SBS type* for traffic offloading in the HetUDN as Definition 1, based on the models established previously.

Definition 1 (SBS Type): In the HetUDN with multiple SBSs, the definition of SBS n's type, which is determined by SBS's transmission efficiency ψ_n, resource capability Ω_n, average revenue achieved from per unit of its bandwidth resource consumption caused by SUEs' traffic demands ϖ_n, average cost on one unit of bandwidth utilization c_n and the arrival rate of SUEs' traffic requests $\lambda_n = \lambda$, is given by

$$\theta_n = \psi_n/\left[c_n + \left(\varpi_n - c_n\right)e^{-\frac{\Omega_n\psi_n}{\lambda}}\right]. \tag{11}$$

Remarks: Notice that the definition of the SBS type is reasonable since that (11) gives an index which can reflect the SBS's capability of providing traffic offloading service for the MBS. To be specific, Definition 1 indicates that a larger value of SBS type θ_n, which means a smaller c_n/ψ_n (SBS n's cost by one unit of traffic transmission), smaller $\left(\varpi_n - c_n\right)/\psi_n$ (SBS n's net benefits from one unit of traffic transmission for SUEs), larger $\Omega_n\psi_n$ (SBS n's maximum resource can be provided for transmission) and lower λ (SBS n's traffic load from its own users), indicates a stronger capability of providing the traffic offloading service for the MBS.

According to Definition 1, each of the N SBSs in the HetUDN belongs to one of the N types. In the SDWN-based HetUDN, the SDWN controller needs to design a bundle of contracts $\{\mathbf{T}\left(\mathbf{y}\right), \mathbf{y}\}$ for these N types of SBSs. Consequently, based on the definitions above, the traffic offloading contract for SBS n with type θ_n can be expressed by $\{T_n\left(y_n\right), y_n\}$. Next, we will introduce necessary principles that ensure a contract to be valid and feasible.

A. Contract Design With Information Asymmetry

1) Individual Rationality (IR): No matter whether the MBS and the SDWN controller can identify the types of SBSs, the designed traffic offloading contract must ensure that every SBS has an incentive to provide the traffic offloading service for MUEs. Therefore, the following *Individual Rationality (IR)* constraint must be satisfied when designing the contracts.

Definition 2 (Individual Rationality (IR)): Any type of SBSs in the HetUDN will only select the traffic offloading contract that can guarantee that the utility received is not less than its utility can be received when it does not provide the traffic offloading service, i.e., $\forall n = 1, 2, \cdots, N$,

$$V_n = P_n\left(\Omega_n - \frac{y_n}{\psi_n}\right) + T_n\left(y_n\right) - \frac{c_n y_n}{\psi_n} > P_n\left(\Omega_n\right). \tag{12}$$

2) Incentive Compatibility (IC): Under the situation with information asymmetry, SBSs tends to request high payment and provide the traffic offloading service as little as possible, according to (7). To ensure that every SBS will select the right contract designed for its type specially, the designed bundle of contracts must make sure that the maximum utility can be achieved if and only if the SBS selects the contract for its type specially. This principle of contract designing is called *Incentive Compatibility (IC)*, which is defined as Definition 3.

Definition 3 (Incentive Compatibility (IC)): Any type of SBSs in the HetUDN will obtain the maximum utility if and only if it selects the contract for its own type specially. In other words, selecting the traffic offloading contract designed for its type will bring to this SBS more utility than any other contracts in the contract bundle, i.e., $\forall n, m = 1, 2, \cdots, N$,

$$\begin{aligned} V_n\left(T_n, y_n\right) &= P_n\left(\Omega_n - \frac{y_n}{\psi_n}\right) + T_n\left(y_n\right) - \frac{c_n y_n}{\psi_n} \\ &\geq V_n\left(T_m, y_m\right) \\ &= P_n\left(\Omega_n - \frac{y_m}{\psi_n}\right) + T_m\left(y_m\right) - \frac{c_n y_m}{\psi_n}. \end{aligned} \tag{13}$$

Due to the case of information asymmetry, the types of SBSs cannot be accessed by the ANDSF. However, the knowledge of the probability π_n, with which an SBS might belong to type θ_n, is available for the SDWN controller, and $\sum_{n\in\mathcal{N}}\pi_n = 1$. Therefore, with the IR and IC constraints, the SDWN controller will formulate the bundle of traffic offloading contracts which will maximize the MBS's utility. Then the contract-based traffic offloading optimization problem in the scenario with information asymmetry is formulated as

$$\max\ U^*(\mathbf{s},\mathbf{y},\mathbf{T}(y))=\sum_{n\in\mathcal{N}}\pi_n\left[\delta\sum_{i\in\mathcal{M}_n}r_i-T_n(y_n)\right], \quad (14a)$$

$$\text{s.t. } s_{0i}+\sum\nolimits_{n\in\mathcal{N}}s_{ni}=1,\quad \forall i=1,2,\cdots,M, \quad (14b)$$

$$y_n=\sum\nolimits_{i\in\mathcal{M}_n}r_{ni}s_{ni}\geq 0,\quad \forall n=1,2,\cdots,N, \quad (14c)$$

$$\Omega_n-\frac{y_n}{\psi_n},\quad \forall n=1,2,\cdots,N, \quad (14d)$$

$$T_n(y_n)-\frac{c_ny_n}{\psi_n}-a_ne^{-\frac{\Omega_n\psi_n}{\lambda_n}}\left(e^{\frac{y_n}{\lambda_n}}-1\right)\geq 0,\quad \forall n=1,2,\cdots,N,\quad \text{(IR)} \quad (14e)$$

$$P_n\left(\Omega_n-\frac{y_n}{\psi_n}\right)+T_n(y_n)-\frac{c_ny_n}{\psi_n}\geq P_n\left(\Omega_n-\frac{y_m}{\psi_n}\right)+T_m(y_m)-\frac{c_ny_m}{\psi_n},\quad \forall n,m=1,2,\cdots,N,\quad \text{(IC)} \quad (14f)$$

$$y_n\geq 0,\quad \forall n=1,2,\cdots,N. \quad (14g)$$

The feasibility conditions of the traffic offloading contacts formulated in (14) will be analyzed and derived in Section V.

B. Contract Design Without Information Asymmetry

Without information asymmetry, the IC constraint is unnecessary because any SBS cannot be disguised as other types of SBSs. Then the optimization problem of traffic offloading processed by the SDWN controller can be formulated as (14), with the IC constraint being removed. We provide the optimal traffic offloading solution for the HetUDN as Lemma 1.

Lemma 1 Without information asymmetry, the optimal traffic offloading contract for type θ_n ($\forall n=1,2,\cdots,N$), which is defined by (11), is given by

$$y_n^{upper}=\lambda_n\left[\ln(\delta\psi_n-c_n)-\ln(w_n-c_n)\right]+\Omega_n\psi_n, \quad (15a)$$

$$T_n^{upper}(y_n)=\frac{c_ny_n^{upper}}{\psi_n}+a_ne^{-\frac{\Omega_n\psi_n}{\lambda_n}}\left(e^{\frac{y_n^{upper}}{\lambda_n}}-1\right). \quad (15b)$$

Proof: We provide the detailed proof in [37].

Remarks: Notice that the value of social welfare in (10) is equal to the MBS's utility. In addition, all SBSs receive zero net utility due to the selfish property of the MBS, who tries to extract as much profit from SBSs' offloading as possible when satisfying the IR constraint shown as (14e). Solutions given in Lemma 1 provide the first best contract solution for the traffic offloading problem, since both the social welfare and MBS utility are maximized and achieve the Pareto efficiency.

C. Contract Design by Linear Pricing

Linear pricing based contracts are designed for the scenario with information asymmetry. the SDWN controller designs a optimal payment β^* to optimize the MBS utility without the IC constraint, and then requests the MBS to pay β^* for every SBS equally for one unit of offloaded traffic. In other words, the SBS requesting more offloading traffic will get more payment linearly. To maximize the SBS utility, every SBS tends to request an appropriate amount of offloading traffic y_n^{lower}. We provide the optimal traffic offloading contract selected by the SBS and the optimal unit-price β^* in Lemma 2.

Lemma 2 With information asymmetry, the optimal traffic offloading contract for type θ_n ($\forall n=1,2,\cdots,N$) under the linear pricing rule is given by

$$y_n^{lower}=\lambda_n\left[\ln\left(\beta^*\psi_n-c_n\right)-\ln(w_n-c_n)\right]+\Omega_n\psi_n, \quad (16a)$$

$$T_n^{lower}=\beta^*y_n^{lower}, \quad (16b)$$

where β^*, designed by the SDWN controller to maximize the MBS utility, is the solution of the following equation:

$$(\delta-\beta)\left(\beta-\frac{c_n}{\psi_n}\right)^{-1}=\ln\left(\frac{\psi_n\beta-c_n}{w_n-c_n}\right)+\frac{\Omega_n\psi_n}{\lambda_n}. \quad (17)$$

Proof: We provide the detailed proof in [37].

The contract-based traffic offloading designed above is feasible and can be realized under the SDWN framework. The required status information in Definition 1, 2 and 3 is obtained through the ANDSF, and contracts satisfying IR and IC are designed by the SDWN controller. However, enough computing capacity of the SDWN controller and corresponding modifications of the interface and interaction protocols are still necessary to realize the contract-based traffic offloading.

V. CONDITIONS FOR CONTRACT FEASIBILITY

First, we propose the following Lemma 3 which provides the condition that ensures the increasing concave property of revenue function $\upsilon_n(y_n)$ defined as (9).

Lemma 3 In a traffic offloading system with a set $\mathcal{N}$ of SBSs indicated by $n=1,2,\cdots,N$. The arrival rate of SBS's own traffic requests is $\lambda_n=\lambda$, $\forall n\in\mathcal{N}$. Define

$$\varphi_n=\frac{w_n-c_n}{\psi_n}e^{-\frac{\Omega_n\psi_n}{\lambda}}. \quad (18)$$

With a bundle of traffic offloading contracts satisfying IR an IC conditions, the traffic offloading allocated to SBS n is y_n, $\forall n\in\mathcal{N}$. Given $y_n\leq y_m$, $(n,m\in\mathcal{N})$, if $\varphi_n\leq\varphi_m$, then the revenue function shown in (9) is a strictly increasing concave function of the amount of traffic offloading allocated.

Proof: Take the first derivative of υ_n in (9) and we get

$$\frac{\partial\upsilon_n(y_n)}{\partial y_n}=\frac{w_n-c_n}{\psi_n}e^{-\frac{\Omega_n\psi_n-y_n}{\lambda}}>0. \quad (19)$$

Therefore, revenue function $\upsilon_n(y_n)$ is a strictly increasing function of y_n, $\forall n\in\mathcal{N}$.

Given $y_n\leq y_m$, and according to IC conditions, the revenue margin between SBS n and SBS m can be calculated by

$$\begin{aligned}&\upsilon_m(y_m)-\upsilon_n(y_n)\\&=a_m\left(1-e^{-\frac{\Omega_m\psi_m-y_m}{\lambda}}\right)+T_m(y_m)-c_my_m/\psi_m\\&\quad+c_my_m/\psi_m-\left[a_n\left(1-e^{-\frac{\Omega_n\psi_n-y_n}{\lambda}}\right)+T_n(y_n)\right]\end{aligned}$$

$$\begin{aligned}
&\geq a_m\left(1-e^{-\frac{\Omega_m\psi_m-y_n}{\lambda}}\right)+T_n\left(y_n\right)-c_m y_n/\psi_m\\
&\quad+c_m y_m/\psi_m-\left[a_n\left(1-e^{-\frac{\Omega_n\psi_n-y_n}{\lambda}}\right)+T_n\left(y_n\right)\right]\\
&=a_m\left(1-e^{-\frac{\Omega_m\psi_m-y_n}{\lambda}}\right)-a_n\left(1-e^{-\frac{\Omega_n\psi_n-y_n}{\lambda}}\right)\\
&\quad+\frac{c_m}{\psi_m}\left(y_m-y_n\right)\\
&\geq a_m\left(1-e^{-\frac{\Omega_m\psi_m-y_n}{\lambda}}\right)-a_n\left(1-e^{-\frac{\Omega_n\psi_n-y_n}{\lambda}}\right)\triangleq F_1\left(y_n\right).
\end{aligned}$$

Take the first derivative of $F_1\left(y_n\right)$, and then we get

$$\frac{\partial F_1\left(y_n\right)}{\partial y_n}=\left(\frac{\varpi_n-c_n}{\psi_n}e^{-\frac{\Omega_n\psi_n}{\lambda}}-\frac{\varpi_m-c_m}{\psi_m}e^{-\frac{\Omega_m\psi_m}{\lambda}}\right)e^{\frac{y_n}{\lambda}}.$$

When $\varphi_n\leq\varphi_m$, then we have $\partial F_1\left(y_n\right)/\partial y_n\leq 0$, which reflects that with y_n and y_m increasing, the revenue margin between y_n and y_m tends to be smaller. Consequently, the revenue function shown in (9) is a strictly increasing concave function of the amount of traffic offloading provided by SBSs. This completes the proof of Lemma 3.

A feasible traffic offloading contract for the information-asymmetry situation must ensure that without the knowledge of SBS types, all SBSs can receive maximum net utility only if they select the right contracts designed for their types. Based on Lemma 3, the following Theorem 1 proposes the monotonic property of SBS's offload amount, payment, and net utility. Theorem 1 demonstrates the feasibility of the proposed contract based traffic offloading and resource allocation method in Section IV-A for the HetUDN with different types of SBSs.

Theorem 1 (Monotonicity): In an SDWN-based HetUDN with N heterogeneous SBSs, the type of each SBS θ_n ($n\in\mathcal{N}$) is defined by Definition 1. Without the information of SBS types, the SDWN controller designs a bundle of traffic offloading contracts $\{\mathbf{T}(\mathbf{y}),\mathbf{y}\}$ for these N types of SBSs and the MBS, according to the optimization problem formulated as (14). Consider that the arrival rates of traffic requests from SUEs are equal for every SBS, i.e., $\lambda_n=\lambda$, $\forall n$. Then for each contract $\{T_n(y_n),y_n\}$, the amount of traffic offload y allocated to each SBS and payment $T(y)$ obtained by (14) have the monotonicity. Specifically, if and only if $\theta_1<\theta_2<\cdots<\theta_N$,

$$y_1<y_2<\cdots<y_N, \tag{20a}$$
$$T_1\left(y_1\right)<T_2\left(y_2\right)<\cdots<T_N\left(y_N\right), \tag{20b}$$
$$V'_1<V'_2<\cdots<V'_N. \tag{20c}$$

Proof: We first prove that $y_1<y_2<\cdots<y_N$ if and only if $\theta_1<\theta_2<\cdots<\theta_N$. According to the IC constraints in (14f), we have $\forall n,m=1,2,\cdots,N$,

$$a_n\left(1-e^{-\frac{\Omega_n\psi_n-y_n}{\lambda}}\right)+T_n\left(y_n\right)-c_n y_n/\psi_n \geq a_n\left(1-e^{-\frac{\Omega_n\psi_n-y_m}{\lambda}}\right)+T_m\left(y_m\right)-c_n y_m/\psi_n, \tag{21a}$$

$$a_m\left(1-e^{-\frac{\Omega_m\psi_m-y_m}{\lambda}}\right)+T_m\left(y_m\right)-c_m y_m/\psi_m \geq a_m\left(1-e^{-\frac{\Omega_m\psi_m-y_n}{\lambda}}\right)+T_n\left(y_n\right)-c_m y_n/\psi_m. \tag{21b}$$

Necessity: Consider that $0\leq y_n\leq y_m$ ($\forall n,m\in\mathcal{N}$, $n\neq m$). For the concave property of the revenue function, the condition of $\varphi_n\leq\varphi_m$ is satisfied according to Lemma 3, and $\varphi_n=\varphi_m$ if and only if $y_n=y_m$. Then add the two inequalities above in (21) together and then we get the following inequality

$$\begin{aligned}
0&\leq\left(a_m e^{-\frac{\Omega_m\psi_m}{\lambda}}-a_n e^{-\frac{\Omega_n\psi_n}{\lambda}}\right)\left(e^{\frac{y_m}{\lambda}}-e^{\frac{y_n}{\lambda}}\right)\\
&=\lambda\left(\varphi_m-\varphi_n\right)\left(e^{\frac{y_m}{\lambda}}-e^{\frac{y_n}{\lambda}}\right)\leq\left(\frac{c_n}{\psi_n}-\frac{c_m}{\psi_m}\right)\left(y_m-y_n\right).
\end{aligned} \tag{22}$$

For $0\leq y_n\leq y_m$, the following inequality is always satisfied:

$$e^{\frac{y_m}{\lambda}}-e^{\frac{y_n}{\lambda}}\geq\left(y_m-y_n\right)/\lambda\geq 0. \tag{23}$$

According to (23), (22) can be transformed to

$$\frac{1}{\lambda}\left(a_m e^{-\frac{\Omega_m\psi_m}{\lambda}}-a_n e^{-\frac{\Omega_n\psi_n}{\lambda}}\right)\leq\frac{c_n}{\psi_n}-\frac{c_m}{\psi_m}, \tag{24}$$

which can be further derived as

$$\begin{aligned}
&\frac{a_m}{\lambda}e^{-\frac{\Omega_m\psi_m}{\lambda}}+\frac{c_m}{\psi_m}\leq\frac{a_n}{\lambda}e^{-\frac{\Omega_n\psi_n}{\lambda}}+\frac{c_n}{\psi_n}\\
&\quad\Rightarrow\frac{\varpi_m-c_m}{\psi_m}e^{-\frac{\Omega_m\psi_m}{\lambda}}+\frac{c_m}{\psi_m}\leq\frac{\varpi_n-c_n}{\psi_n}e^{-\frac{\Omega_n\psi_n}{\lambda}}+\frac{c_n}{\psi_n},
\end{aligned} \tag{25}$$

which is equal to

$$\frac{\psi_n}{c_n+\left(\varpi_n-c_n\right)e^{-\frac{\Omega_n\psi_n}{\lambda}}}\leq\frac{\psi_m}{c_m+\left(\varpi_m-c_m\right)e^{-\frac{\Omega_m\psi_m}{\lambda}}}. \tag{26}$$

According to the definition of θ, we can get $\theta_n\leq\theta_m$, and $\theta_n=\theta_m$ if and only if $y_n=y_m$.

Sufficiency: Consider Definition 1, $0<\theta_n\leq\theta_m$ is equal to

$$\frac{c_n}{\psi_n}+\frac{\varpi_n-c_n}{\psi_n}e^{-\frac{\Omega_n\psi_n}{\lambda}}\geq\frac{c_m}{\psi_m}+\frac{\varpi_m-c_m}{\psi_m}e^{-\frac{\Omega_m\psi_m}{\lambda}}, \tag{27}$$

which can be written as

$$\left(\frac{\varpi_m-c_m}{\psi_m}e^{-\frac{\Omega_m\psi_m}{\lambda}}-\frac{\varpi_n-c_n}{\psi_n}e^{-\frac{\Omega_n\psi_n}{\lambda}}\right)+\left(\frac{c_m}{\psi_m}-\frac{c_n}{\psi_n}\right)\leq 0. \tag{28}$$

Hypothesise $y_n>y_m>0$, then $\varphi_n>\varphi_m$, and inequality

$$\lambda\left(e^{\frac{y_n}{\lambda}}-e^{\frac{y_m}{\lambda}}\right)>y_n-y_m>0 \tag{29}$$

is always satisfied. Then (28) can be further derived as

$$\begin{aligned}
0&\leq\left(\frac{\varpi_n-c_n}{\psi_n}e^{-\frac{\Omega_n\psi_n}{\lambda}}-\frac{\varpi_m-c_m}{\psi_m}e^{-\frac{\Omega_m\psi_m}{\lambda}}\right)\left(y_n-y_m\right)\\
&\quad+\left(\frac{c_n}{\psi_n}-\frac{c_m}{\psi_m}\right)\left(y_n-y_m\right)
\end{aligned} \tag{30}$$

$$\begin{aligned}
&<\lambda\left(\frac{\varpi_n-c_n}{\psi_n}e^{-\frac{\Omega_n\psi_n}{\lambda}}-\frac{\varpi_m-c_m}{\psi_m}e^{-\frac{\Omega_m\psi_m}{\lambda}}\right)\left(e^{\frac{y_n}{\lambda}}-e^{\frac{y_m}{\lambda}}\right)\\
&\quad+\left(\frac{c_n}{\psi_n}-\frac{c_m}{\psi_m}\right)\left(y_n-y_m\right).
\end{aligned} \tag{31}$$

However, according to IC constraints and adding (21a) and (21b) together, then we have

$$\left(a_m e^{-\frac{\Omega_m\psi_m}{\lambda}}-a_n e^{-\frac{\Omega_n\psi_n}{\lambda}}\right)\left(e^{\frac{y_m}{\lambda}}-e^{\frac{y_n}{\lambda}}\right)\leq\left(\frac{c_n}{\psi_n}-\frac{c_m}{\psi_m}\right)\left(y_m-y_n\right). \tag{32}$$

Considering $y_n > y_m > 0$, (32) can be transformed as

$$\left(a_n e^{-\frac{\Omega_n \psi_n}{\lambda}} - a_m e^{-\frac{\Omega_m \psi_m}{\lambda}}\right)\left(e^{\frac{y_n}{\lambda}} - e^{\frac{y_m}{\lambda}}\right) + \left(\frac{c_n}{\psi_n} - \frac{c_m}{\psi_m}\right)(y_n - y_m) \le 0.$$

As $a_n = \lambda (w_n - c_n)/\psi_n$, we can get

$$\lambda\left(\frac{w_n - c_n}{\psi_n} e^{-\frac{\Omega_n \psi_n}{\lambda}} - \frac{w_m - c_m}{\psi_m} e^{-\frac{\Omega_m \psi_m}{\lambda}}\right)\left(e^{\frac{y_n}{\lambda}} - e^{\frac{y_m}{\lambda}}\right) + \left(\frac{c_n}{\psi_n} - \frac{c_m}{\psi_m}\right)(y_n - y_m) \le 0, \quad (33)$$

which is a contradiction with (30). Therefore, the hypothesis $y_n > y_m$ is invalid, which means that $y_n \le y_m$ if $\theta_n \le \theta_m$.

Then we have demonstrated the proposition that $y_n < y_m$ if and only if $\theta_n < \theta_m$, and $y_n = y_m$ if and only if $\theta_n = \theta_m$. Next, we will prove that $T_1(y_1) < T_2(y_2) < \cdots < T_N(y_N)$ if and only if $y_1 < y_2 < \cdots < y_N$.

Sufficiency: $\forall n, m = 1, 2, \cdots, N, n \neq m$, we have (21b) according to IC constraints, which can be transformed to

$$T_n(y_n) - T_m(y_m) \le a_m e^{-\frac{\Omega_m \psi_m}{\lambda}}\left(e^{\frac{y_n}{\lambda}} - e^{\frac{y_m}{\lambda}}\right) + \frac{c_m}{\psi_m}(y_n - y_m),$$

and then we get $T_n(y_n) \le T_m(y_m)$ if $y_n \le y_m$.

Necessity: Inequality (21a) obtained by the IC constraints can be transformed to

$$a_n e^{-\frac{\Omega_n \psi_n}{\lambda}}\left(e^{\frac{y_n}{\lambda}} - e^{\frac{y_m}{\lambda}}\right) + \frac{c_n}{\psi_n}(y_n - y_m) \le T_n(y_n) - T_m(y_m).$$

Given $T_n(y_n) \le T_m(y_m)$, the left part of the inequality above can be written by

$$a_n e^{-\frac{\Omega_n \psi_n}{\lambda}}\left(e^{\frac{y_n}{\lambda}} - e^{\frac{y_m}{\lambda}}\right) + \frac{c_n}{\psi_n}(y_n - y_m) \le 0, \quad (34)$$

which can be satisfied only by $0 \le y_n \le y_m$.

Then we have demonstrated the proposition that $T_n(y_n) < T_m(y_m)$ if and only if $y_n < y_m$, and $T_n(y_n) = T_m(y_m)$ if and only if $y_n = y_m$. Due to the transferability of the necessary and sufficient conditions, $T_n(y_n) < T_m(y_m)$ if and only if $\theta_n < \theta_m$, and $T_n(y_n) = T_m(y_m)$ if and only if $\theta_n = \theta_m$.

Last, we will prove the monotonicity of SBS's net utility. According to (7), (8) and the IC constraints in Definition 3, the net utility difference between SBS n and SBS m ($\forall n, m = 1, 2, \cdots, N$, $n \neq m$) can be calculated as

$$\begin{aligned}
&V'_m - V'_n \\
&= V_m - P_m(\Omega_m) - (V_n - P_n(\Omega_n)) \\
&= a_m\left(1 - e^{-\frac{\Omega_m \psi_m - y_m}{\lambda}}\right) + T_m(y_m) - \frac{c_m y_m}{\psi_m} - P_m(\Omega_m) \\
&\quad - \left[a_n\left(1 - e^{-\frac{\Omega_n \psi_n - y_n}{\lambda}}\right) + T_n(y_n) - \frac{c_n y_n}{\psi_n}\right] + P_n(\Omega_n) \\
&\ge a_m\left(1 - e^{-\frac{\Omega_m \psi_m - y_n}{\lambda}}\right) + T_n(y_n) - \frac{c_m y_n}{\psi_m} - P_m(\Omega_m) \\
&\quad - \left[a_n\left(1 - e^{-\frac{\Omega_n \psi_n - y_n}{\lambda}}\right) + T_n(y_n) - \frac{c_n y_n}{\psi_n}\right] + P_n(\Omega_n)
\end{aligned} \quad (35)$$

$$\begin{aligned}
&= a_m\left(1 - e^{-\frac{\Omega_m \psi_m - y_n}{\lambda}}\right) - a_n\left(1 - e^{-\frac{\Omega_n \psi_n - y_n}{\lambda}}\right) \\
&\quad + \left(\frac{c_n}{\psi_n} - \frac{c_m}{\psi_m}\right) y_n - P_m(\Omega_m) + P_n(\Omega_n) \triangleq F_2(y_n)
\end{aligned} \quad (36)$$

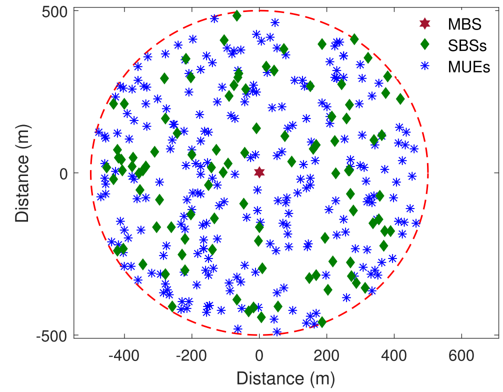


Fig. 2. Distributions of MUEs, MBS and SBSs in the simulation scene. (The red dotted circle is the coverage of the MBS.)

Consider that $\theta_n < \theta_m$, then $y_n < y_m$ and $\varphi_n < \varphi_m$ according to (20a) in Theorem 1 proved previously and Lemma 3, respectively. Let $y_n = 0$ in $F_2(y_n)$, and according to the results of $P_m(\Omega_m)$ and $P_m(\Omega_m)$ calculated by (5), we have

$$F_2(0) = a_m\left(1 - e^{-\frac{\Omega_m \psi_m}{\lambda}}\right) - a_n\left(1 - e^{-\frac{\Omega_n \psi_n}{\lambda}}\right) - P_m(\Omega_m) + P_n(\Omega_n) = 0. \quad (37)$$

According to the expression of θ_n defined in Definition 1 and considering that $y_n > 0$, the first derivative of $F_2(y_n)$ with respect to y_n can be written as

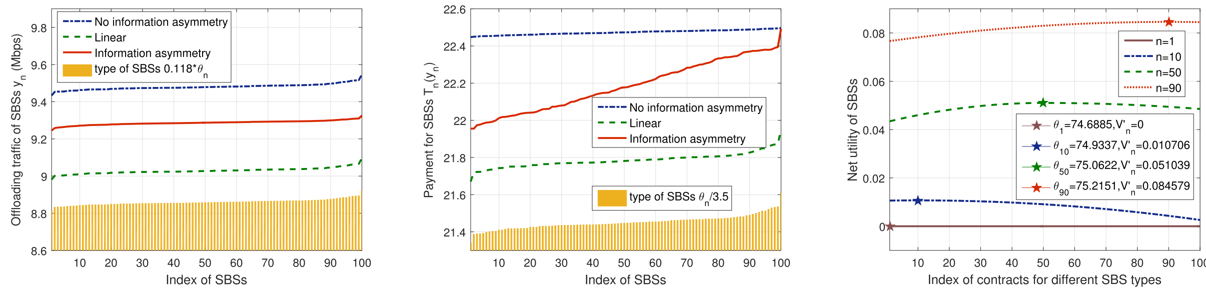
$$\begin{aligned}
\frac{\partial F_2(y_n)}{\partial y_n} &= \left(\frac{c_n}{\psi_n} - \frac{c_m}{\psi_m}\right) + (\varphi_n - \varphi_m) e^{\frac{y_n}{\lambda}} \\
&\ge \left(\frac{c_n}{\psi_n} - \frac{c_m}{\psi_m}\right) + (\varphi_n - \varphi_m) \\
&= \left(\frac{c_n}{\psi_n} + \varphi_n\right) - \left(\frac{c_m}{\psi_m} + \varphi_m\right) = \frac{1}{\theta_n} - \frac{1}{\theta_m} > 0.
\end{aligned} \quad (38)$$

The necessity of (20c) can be proved by applying the reduction to absurdity. Since the proving idea is similar to (35) - (38), we omit the proof of necessity for (20c). Therefore, we have $V'_n < V'_m$, if and only if $\theta_n < \theta_m$, and $V'_n = V'_m$, if and only if $\theta_n = \theta_m, \forall n, m \in \mathcal{N}$, $n \neq m$. This completes the proof of Theorem 1.

Remarks:

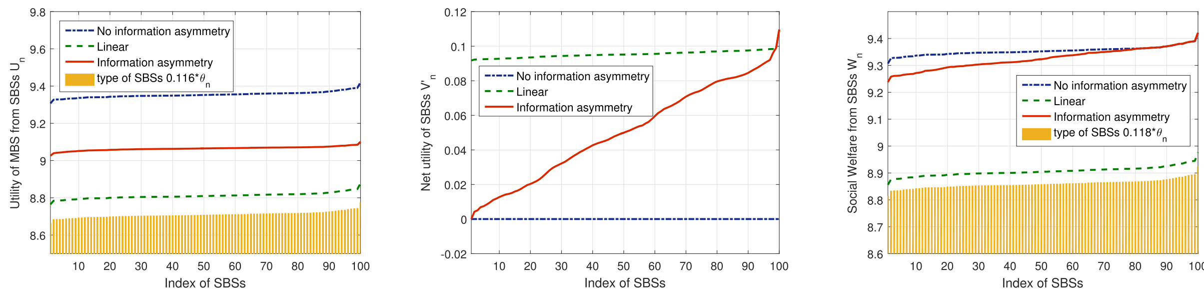
1) Valid of SBS Type Definition: Theorem 1 demonstrates that SBS type θ_n proposed and defined in Definition 1 is reasonable, since it can effectively reflect the influence of heterogeneous SBSs' performance and capacity on the contract designed by a competitive market based economics theory.

2) Fairness and Monotonicity: Theorem 1 demonstrate that, for both of the service requester and service providers, i.e., the MBS and SBSs, respectively, the proposed contract-based traffic offloading and resource allocation mechanism as (14) guarantees the fairness and incentive property of the transmission resource market, in the scenario of information asymmetry and that service providers are heterogeneous. On the one hand, monotonicity of (20a) and (20b) implies that for the SBSs with higher θ, they are more suitable for offloading traffic, and their best choice to achieve highest payoff is offloading larger amount of traffic. Meanwhile, they will receive more payment. This contract principle can ensure the fairness among the heterogeneous SBSs. On the other hand, monotonicity also



(a) Traffic offloading y_n versus SBS types. (b) Payments $T_n(y_n)$ versus SBS types. (c) Net utility received by different contracts.

Fig. 3. The contract monotonicity and incentive compatibility versus different SBS types.



(a) Utility of MBS U_n (b) Net utility of SBSs V_n' (c) Social welfare W_n

Fig. 4. System performance of different types of SBSs when applying different traffic offloading and resource allocation mechanisms.

TABLE I

SIMULATION PARAMETERS

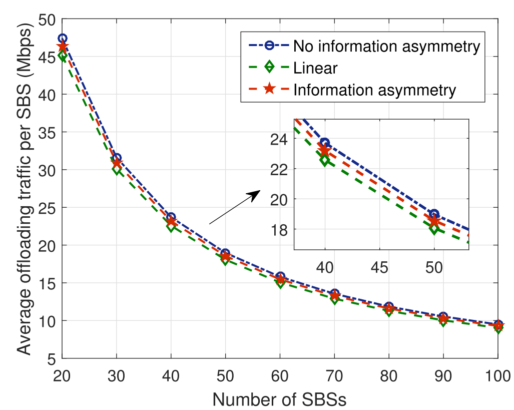
System parameters	Value settting
Transmission power of MBS	46 dBm
Transmission power of SBSs	$\sim U\,[15, 35]$ dBm
Path loss of MBS	$28.3 + 22.0\log_{10}l$, l (km)
Path loss of SBSs	$30.5 + 36.7\log_{10}l$, l (km)
MBS / SBS bandwidth	20 MHz
MBS / SBS operating frequency	2.6 GHz / 2.4 GHz
SBSs' own traffic requests arrival rate	$\lambda = 10$ Mbps
Power spectral density of thermal noise	-174 dBm/Hz

provides an incentive for SBSs. Specifically, if a high type of SBS selects the contract designed for low types of SBSs, even though a small amount of traffic offloading will be requested by the SDWN controller, the corresponding low payment will deteriorate this high-type SBS's payment.

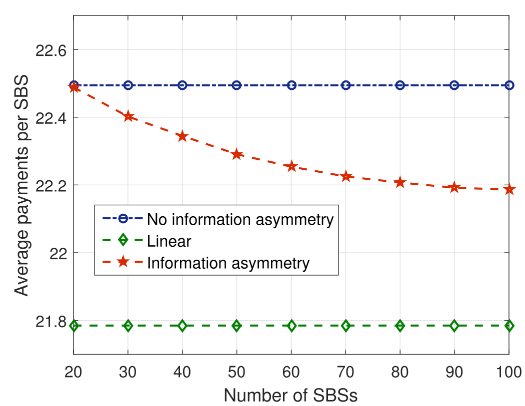
3) Incentive Compatibility: Monotonicity of (20c) also implies that the incentive for SBSs is compatible, which means that SBSs with high capability will receive more net utility than low ones. For those SBSs whose types cannot be aware of the MBS and SDWN controller, the designed contract is self-revealing for SBSs, since that each type of SBS will receive the maximum net utility, which reflects the net revenue by offloading, if and only if it selects the right traffic offloading contract designed exactly for its type.

VI. SIMULATION RESULTS

In this part, we will use MATLAB 2016b to evaluate the proposed contract-based traffic offloading and resource allocation. First of all, we introduce the scenario setup of the simulations. In the following simulations, we assume a typical 4G/5G macrocell with a transmission radius of 500 m. The HetUDN consists of one MBS, $N = 100$ heterogeneous SBSs with $N = 100$ different types, and $M = 250$ MUEs. Both SBSs and MUEs are randomly distributed within the macrocell. The distribution of network elements in the simulation is shown as Fig. 2. In addition, we set $c_n = 0.6$, $w_n = 1$ and $\delta = 1$. The other main parameters of the HetUDN are shown in Table I.

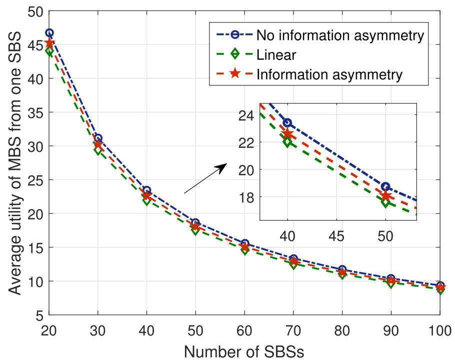


(a) Average traffic offloading versus the number of SBS types.

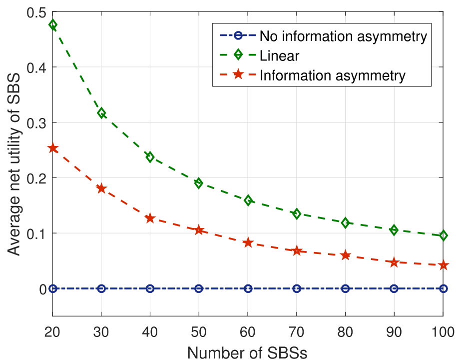


(b) Average payments versus the number of SBS types.

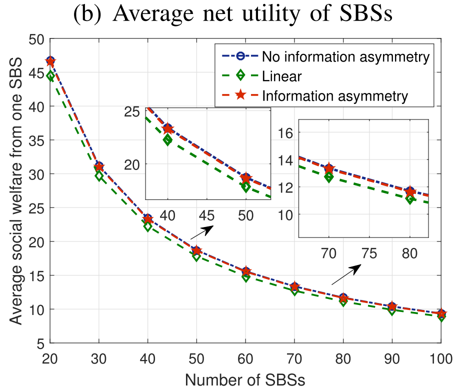
Fig. 5. The contract performance versus the number of different SBS types.



(a) Average MBS utility from one SBS



(b) Average net utility of SBSs



(c) Average social welfare from one SBS

Fig. 6. System performance versus the number of SBSs when applying different traffic offloading and resource allocation mechanisms.

To demonstrate monotonicity and incentive compatibility of the contract, the indexes of SBSs are sorted according to their values of type obtained by Definition 1. By applying the three different contracts designed in Section IV, we obtain the amount of traffic offloading requested by SBSs and payments required to the MBS, which are shown as Fig. 3(a) and Fig. 3(b), respectively. In Fig. 3, results illustrate that both the amount of traffic offloading and payment increase with the value of SBS type increasing, for the three different contracts, which reflects the fairness of the contracts. In addition, among the three contracts, the no information asymmetry contract requires the highest amount of traffic offloading and the highest payment for SBSs, followed by the incentive contract proposed in Section IV-A. The lowest traffic offloading and payment are requested by the linear pricing contract.

Moreover, the incentive compatibility of the contract designed in Section IV-A is verified by results shown in Fig. 3(c). Fig. 3(c) presents the net utility received by selecting $N = 100$ different contracts in the contract bundle for four sample SBSs $n = 1, 10, 50, 90$. The pentagram marks in Fig. 3(c) are the maximum net utility received for the four SBSs, and the corresponding horizontal axes points are the indexes of SBS types that contracts are designed for. Results indicate that for each type of SBS, the maximum net utility can be achieved only by selecting the right contract designed for this type.

By applying three different contracts, the system performance of HetUDN is shown as Fig. 4, which presents that the MBS utility, SBS net utility and social welfare increase monotonically with the value of SBS type growing. Results in Fig. 4(a) show that the contract for the scenario without information asymmetry brings the maximum utility for the MBS. Under the case that the SBS types are unavailable, the designed IC-based contract can only bring a approximate optimal utility for the MBS, which is upper bounded by the no information asymmetry situation. With information asymmetry, the linear pricing based contract does not treat differently to all types of SBSs. Therefore, without the knowledge of SBS type, the linear pricing performs worst on the MBS utility.

Since the MBS is selfish, when it is aware of the type of every BSS, the designed contract only need to satisfy the IR constraints when maximizing the MBS utility. Then every SBS can only get the utility equal to that of providing no offloading service, which means that the net utility is zero for every SBS, as shown in Fig. 4(b). By applying the contract with IC constraints, only the SBS with the lowest type value will receive zero-net utility, and SBSs with lower θ will receive less net utility than that obtained by linear pricing contract. However, for those SBSs with higher θ, they can receive more net utility than that obtained by linear pricing contract, which demonstrates the incentive compatibility of the IC based contract. The social welfare shown in Fig. 4(c) presents a similar result as Fig. 4(a). In addition, with the IC based contract for information asymmetry, the SBS with the highest θ brings the same social welfare as no information asymmetry.

Next, we study that how the contract and system performance change with the changing density of SBSs. Let the number of SBSs in the macrocell system varies from 20 to 100, and other parameters are set as before. The average traffic offloaded and the average payments for each SBS versus the nember of SBSs (types) are shown in Fig 5. The differences between the effects on the amount of traffic offloading by three traffic offloading contracts shown in Fig. 5(a) are similar to that of Fig. 3(a). In addition, the average amount of traffic offloading for every SBS decreases when the number of SBSs grows, which results that if the amount of total traffic request are fixed, distributing more SBSs will lighten the load of every BS. Fig. 5(b) indicates that the average payment for each SBS does not change by applying the contract for the no information asymmetry case and the linear pricing contract, no matter how many SBSs in the HetUDN. By applying the IC based contract for the information asymmetry case, the average payment obtained by per SBS decreases when the number of SBS increases. These results reflect the high effectiveness and efficiency of the contract designed in Section IV-A. Specifically, when there are more SBSs in the system, which means that more candidates can provide the traffic offloading service and the competitiveness among these SBSs tends to be weak, the average payment provided by the MBS for each SBS will be less than that in a more competitive market.

Due the same reason explained above, the average MBS utility, average SBS net utility and average social welfare by one SBS will all decrease with increasing number of SBS, by applying the three contracts, except that the SBS net utility obtained under the no information asymmetry case is always zero, as shown in Fig. 6. In addition, the social welfare shown in Fig. 6(c) also implies the approximate optimization property of the IC based contract for the information asymmetry case.

VII. Conclusion

In this work, we have proposed a contract-based traffic offloading and resource allocation mechanism for the SDWN-cased HetUDN. In the scenario with information asymmetry, the designed IC based traffic offloading contract has the incentive property to encourage every SBS to select the right contract designed personally to it, which specifies the amount of traffic offloading and the payment from the MBS. In addition, the SBS utility, MBS utility and social welfare can achieve an approximate optimization, comparing the situation without information asymmetry, and better than that achieved by linear pricing contract with information asymmetry. Furthermore, the definition of SBS type proposed in this work provides a valid index to measure the offloading performance of heterogeneous SBSs. Meanwhile, the SBS type definition also guarantee the monotonicity and incentive compatibility of contracts. In addition, the defined closed-form expression of SBS type makes this definition enforceable to be applied in HetNets with densely distributed BSs for resource management.

References

[1] X. Costa-Perez *et al.*, "5G-crosshaul: An SDN/NFV integrated fronthaul/backhaul transport network architecture," *IEEE Wireless Commun.*, vol. 24, no. 1, pp. 38–45, Feb. 2017.

[2] C. Jiang, H. Zhang, Y. Ren, Z. Han, K.-C. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 98–105, Apr. 2017.

[3] C. Jiang, B. Wang, Y. Han, Z.-H. Wu, and K. R. Liu, "Spatial focusing inspired 5G spectrum sharing," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*. New Orleans, LA, USA, Mar. 2017, pp. 3609–3613.

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

[5] J. Ding, R. Yu, Y. Zhang, S. Gjessing, and D. H. K. Tsang, "Service provider competition and cooperation in cloud-based software defined wireless networks," *IEEE Commun. Mag.*, vol. 53, no. 11, pp. 134–140, Nov. 2015.

[6] S. Kang and W. Yoon, "SDN-based resource allocation for heterogeneous LTE and wlan multi-radio networks," *J. Supercomput.*, vol. 72, no. 4, pp. 1342–1362, Apr. 2016.

[7] Y. Zhang, L. Liu, Y. Gu, D. Niyato, M. Pan, and Z. Han, "Offloading in software defined network at edge with information asymmetry: A contract theoretical approach," *J. Signal Process. Syst.*, vol. 83, no. 2, pp. 241–253, May 2016.

[8] H. Zhang, S. Huang, C. Jiang, K. Long, V. C. Leung, and H. V. Poor, "Energy efficient user association and power allocation in millimeter-wave-based ultra dense networks with energy harvesting base stations," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1936–1947, Sep. 2017.

[9] J. Du, C. Jiang, J. Wang, Y. Ren, S. Yu, and Z. Han, "Resource allocation in space multiaccess systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 2, pp. 598–618, Apr. 2017.

[10] E. Baccarelli, N. Cordeschi, A. Mei, M. Panella, M. Shojafar, and J. Stefa, "Energy-efficient dynamic traffic offloading and reconfiguration of networked data centers for big data stream mobile computing: Review, challenges, and a case study," *IEEE Netw.*, vol. 30, no. 2, pp. 54–61, Mar./Apr. 2016.

[11] Y. Mao, J. Zhang, and K. B. Letaief, "Dynamic computation offloading for mobile-edge computing with energy harvesting devices," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3590–3605, Dec. 2016.

[12] S. Zhang, N. Zhang, S. Zhou, J. Gong, Z. Niu, and X. Shen, "Energy-aware traffic offloading for green heterogeneous networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1116–1129, May 2016.

[13] Y. Wu, K. Guo, J. Huang, and X. S. Shen, "Secrecy-based energy-efficient data offloading via dual connectivity over unlicensed spectrums," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3252–3270, Dec. 2016.

[14] F. Mehmeti and T. Spyropoulos, "Performance analysis of mobile data offloading in heterogeneous networks," *IEEE Trans. Mobile Comput.*, vol. 16, no. 2, pp. 482–496, Feb. 2017.

[15] N. C. Nguyen, P. Wang, D. Niyato, Y. Wen, and Z. Han, "Resource management in cloud networking using economic analysis and pricing models: A survey," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 954–1001, 2nd Quart., 2017.

[16] L. Gao, G. Iosifidis, J. Huang, L. Tassiulas, and D. Li, "Bargaining-based mobile data offloading," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1114–1125, Jun. 2014.

[17] M. I. Kamel and K. M. Elsayed, "ABSF offsetting and optimal resource partitioning for eICIC in LTE-Advanced: Proposal and analysis using a Nash bargaining approach," in *Proc. IEEE Int. Conf. Commun. (ICC)*. Budapest, Hungary, Jun. 2013, pp. 9–13.

[18] B. C. Chung and D.-H. Cho, "Mobile data offloading with almost blank subframe in LTE-LAA and Wi-Fi coexisting networks based on coalition game," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 608–611, Mar. 2016.

[19] T. M. Ho, N. H. Tran, L. B. Le, W. Saad, S. A. M. Kazmi, and C. S. Hong, "Coordinated resource partitioning and data offloading in wireless heterogeneous networks," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 974–977, May 2016.

[20] A. F. Hanif, H. Tembine, M. Assaad, and D. Zeghlache, "Mean-field games for resource sharing in cloud-based networks," *IEEE/ACM Trans. Netw.*, vol. 24, no. 1, pp. 624–637, Feb. 2016.

[21] S. Samarakoon, M. Bennis, W. Saad, M. Debbah, and M. Latva-Aho, "Ultra dense small cell networks: Turning density into energy efficiency," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1267–1280, May 2016.

[22] X. Xu, C. Yuan, J. Li, and X. Tao, "Energy-efficient active offloading with collaboration communication and power allocations for heterogeneous ultradense networks," *Trans. Emerg. Telecommun. Technol.*, vol. 28, no. 9, Sep. 2017, doi: 10.1002/ett.3152.

[23] C. Jiang, Y. Chen, Q. Wang, and K. J. R. Liu, "Data-driven auction mechanism design in IaaS cloud computing," *IEEE Trans. Serv. Comput.*, to be published.

[24] E. Gelenbe, "Analysis of automated auctions," in *Computer and Information Sciences—ISCIS*. Berlin, Germany: Springer, 2006, pp. 1–12.

[25] E. Gelenbe, "Analysis of single and networked auctions," *ACM Trans. Internet Technol.*, vol. 9, no. 2, 2009, Art. no. 8.

[26] E. Gelenbe and L. Györfi, "Performance of auctions and sealed bids," in *Computer Performance Engineering*. Berlin, Germany: Springer, 2009, pp. 30–43.

[27] E. Gelenbe and K. Velan, "An approximate model for bidders in sequential automated auctions," in *Agent and Multi-Agent Systems: Technologies and Applications*. Berlin, Germany: Springer, 2009, pp. 70–79.

[28] E. Gelenbe and K. Velan, "Modelling bidders in sequential automated auctions," in *Proc. 8th Int. Conf. Auto. Agents Multiagent Syst. (AAMAS)*, Budapest, Hungary, 2009, 208–218.

[29] K. Velan and E. Gelenbe, "Analysing bidder performance in randomised and fixed-deadline automated auctions," in *Agent and Multi-Agent Systems: Technologies and Applications*. Berlin, Germany: Springer, 2010, pp. 42–51.

[30] A. Di Ferdinando, R. Lent, and E. Gelenbe, "A framework for autonomic networked auctions," in *Proc. Workshop Innov. Serv. Technol.*, 2007, Art. no. 3.

[31] G. Sakellari, T. Leung, and E. Gelenbe, “Auction-based admission control for self-aware networks,” in *Computer and Information Sciences II*. London, U.K.: Springer, 2012, pp. 223–230.

[32] G. Iosifidis, L. Gao, J. Huang, and L. Tassiulas, “A double-auction mechanism for mobile data-offloading markets,” *IEEE/ACM Trans. Netw.*, vol. 23, no. 5, pp. 1634–1647, Oct. 2015.

[33] D. Zhang, Z. Chang, and T. Hamalainen, “Reverse combinatorial auction based resource allocation in heterogeneous software defined network with infrastructure sharing,” in *Proc. IEEE 83rd Veh. Technol. Conf. (VTC Spring)*, Nanjing, China, May 2016, pp. 15–18.

[34] Y. Zhang, L. Song, W. Saad, Z. Dawy, and Z. Han, “Contract-based incentive mechanisms for device-to-device communications in cellular networks,” *IEEE J. Sel. Areas Commun.*, vol. 33, no. 10, pp. 2144–2155, Oct. 2015.

[35] A. Asheralieva and Y. Miyanaga, “Optimal contract design for joint user association and intercell interference mitigation in heterogeneous LTE-A networks with asymmetric information,” *IEEE Trans. Veh. Technol.*, vol. 66, no. 6, pp. 5284–5300, Jun. 2017.

[36] M. Peng, X. Xie, Q. Hu, J. Zhang, and H. V. Poor, “Contract-based interference coordination in heterogeneous cloud radio access networks,” *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1140–1153, Jun. 2015.

[37] J. Du, E. Gelenbe, C. Jiang, H. Zhang, and Y. Ren. *Appendix for Contract Design for Traffic Offloading and Resource Allocation in Heterogeneous Ultra-Dense Networks*, accessed: Sep. 26, 2017. [Online]. Available: http://www.jiangchunxiao.net/appendix.pdf

Jun Du (S'16) received the B.S. degree in information and communication engineering from the Beijing Institute of Technology, Beijing, China, in 2009, and the M.S. degree in information and communication engineering from Tsinghua University, Beijing, in 2014, where she is currently pursuing the Ph.D. degree. From 2016 to 2017, he was a Sponsored Researcher, and she visited Imperial College London, U.K. Her research interests are mainly in resource allocation and system security of heterogeneous networks and space-based information networks.

Erol Gelenbe (S'67–M'70–SM'79–F'86–LF'11) is currently a Professor with the Dennis Gabor Chair, Department of Electrical and Electronic Engineering, Imperial College London, where he is involved in energy savings, system security, and network quality of service. He previously founded the System Performance Evaluation groups at INRIA and French universities, chaired the ECE Department at Duke University, and created the School of Electrical Engineering and Computer Science at University of Central Florida. In his theoretically oriented research, he invented mathematical models, such as G-Networks for computer and network performance, and the random neural network to mimic biological spiking neural networks and machine learning.

Chunxiao Jiang (S'09–M'13–SM'15) received the B.S. degree (Hons.) in information engineering from Beihang University in 2008, and the Ph.D. degree (Hons.) in electronic engineering from Tsinghua University in 2013. From 2013 to 2016, he held a post-doctoral position with the Department of Electronic Engineering, Tsinghua University, during which he visited the University of Maryland College Park and University of Southampton. He was a recipient of the IEEE Communications Society Young Author Best Paper Award in 2017.

Haijun Zhang (M'13–SM'17) was a Post-Doctoral Research Fellow with the Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC, Canada. He is currently a Full Professor with the University of Science and Technology Beijing, China. He has published more than 100 papers and authored two books. He serves as an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE 5G TECH FOCUS, and the *Journal of Network and Computer Applications*, and serves/served as a Leading Guest Editor for the *IEEE Communications Magazine* and the IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTING. He serves/served as the General Co-Chair of 5GWN'17 and GameNets'16, the Track Chair of ScalCom2015, the Symposium Chair of GameNets'14, the TPC Co-Chair of INFOCOM 2018 Workshop on Integrating Edge Computing, Caching, and Offloading in Next Generation Networks, the General Co-Chair of ICC 2018 (ICC 2017, Globecom 2017) Workshop on 5G Ultra Dense Networks, and the General Co-Chair of Globecom 2017 Workshop on LTE-U. He received the IEEE ComSoc Young Author Best Paper Award in 2017.

Yong Ren (SM'16) received the B.S., M.S., and Ph.D. degrees in electronic engineering from the Harbin Institute of Technology, China, in 1984, 1987, and 1994, respectively. He held a post-doctoral position with the Department of Electronics Engineering, Tsinghua University, China, from 1995 to1997, where he is currently a Professor with the Department of Electronics Engineering and the Director of the Complexity Engineered Systems Laboratory. He holds 12 patents and has authored or co-authored over 100 technical papers in the behavior of computer network, P2P network, and cognitive networks.