Procurement-Based User Association for LTE-Advanced HetNets

Satendra Kumar, *Student Member, IEEE*, and Sudip Misra, *Senior Member, IEEE*

***Abstract*—This paper presents a procurement-based user association scheme for LTE-Advanced HetNets. The proposed scheme enables LTE Macrocell (LMC) to leverage capacity from Femtocells (FCs) and fulfills the excess demand of the registered LMC (rLMC) users. Additionally, we propose an energy-efficient and QoS-ensured resource allocation scheme for rLMC users. Although the existing schemes capacitate 2G/3G-based cellular networks to leverage resources from complementary networks such as WiFi, these schemes are not well suited for LTE-Advanced networks because of the different architecture. Moreover, the existing schemes centered on rate-adaptive resource allocation, which is limited in capability of offering guaranteed bit rate (GBR) services to the rLMC users. The proposed scheme has manifold notable features. First, unlike the existing schemes, it implicitly ensures QoS of offloaded rLMC users. Second, it encourages competition among FCs, which results in cost-effective rLMC user association. Third, it is adaptive to the stochastic nature of users' demands. Finally, for data rate calculation, we use lookup tables given in 3GPP standards for LTE-Advanced. This makes the proposed scheme practically implementable in LTE-Advanced HetNets. Numerical simulations and the corresponding insights of the results justify the suitability of the proposed scheme for LTE-Advanced HetNets.**

***Index Terms*—LTE-Advanced, femtocell, HetNets, procurement, user association.**

I. INTRODUCTION

In recent times, the magnitude of resource-demanding and power intensive mobile applications have increased excessively. Examples include video-calling, media-streaming, virtual reality (VR), and online gaming. According to a recent forecast [1] cellular traffic is expected to increase by nearly four times in the next four years. Traditional cellular networks (CNs) with LMC base station (BS) are unable to handle this massive data traffic due to their inherit constraint on coverage and network-capacity. To overcome this hurdle, cellular network providers (CNPs) opt for heterogeneous networks (HetNets) architecture [2], [3], wherein several FCs[1] are deployed under the blanket coverage of LMC. The FC is low-power, short-range, and cost-effective wireless access point which increases the network capacity and coverage, by decreasing the number of dead zones (i.e., regions with no wireless network coverage) [4]–[6].

The FCs are deployed by either CNPs or home-based users. Accordingly, the access mode of FC is divided into three categories — open, hybrid, and close [7]. In open access mode, any user within the coverage range of FC can avail the service. The main disadvantage of open access mode is the absence of access control mechanism which eventually increases the network congestion and decreases the user QoS. However, in case of closed access mode, only subscribed users (SUs) are allowed to use the FC service. Each FC maintains the information about SUs in the access control list [4]. Although closed access mode ensures better secrecy and QoS to SUs, the user access restriction causes under utilization of radio resources. To overcome the shortcoming of both open and closed access mode, the 3GPP standard [8] proposes hybrid access mode. In this mode, FC owners first use the radio resource to serve their own subscribed users and remaining resources are allocated to serve rLMC users.

The LTE-A QoS are mainly classified in two categories, namely: guaranteed bit rate (GBR) and non-guaranteed bit rate (non-GBR). The GBR service is applicable for real-time applications wherein the minimum data rate of user is ensured over the time-varying wireless channel [9]. However, the non-GBR services are suitable for delay-tolerant applications and may suffer packet loss under network congestion. In literature, there exists two dynamic resource allocation schemes – rate-adaptive (RA) and marginal-adaptive (MA) to deal with above mentioned service classes. In RA scheme, cells aim to maximize the throughput for a given power budget and bit error rate constraints [10]. On the contrary, in MA scheme, cells minimize the total transmitting power subject to the throughput and bit error rate requirement. Specifically, the MA adjusts the power to compensate the dynamic channel condition [11], [12]. Hence, for GBR services, the MA scheme is more effective compare to RA resource allocation scheme [10].

Nevertheless, to maintain user QoS, the CNP needs to deploy their FCs wherever possible. However, this increases the overall capital and expenditure cost of CNPs. Therefore, the CNP motivates the FC owners to operate in the hybrid mode, by offering incentives, and serve rLMC users. Above all, we aim to answer following queries from LMC perspective: *i) which FC should be selected for rLMC user association, ii) which set of rLMC users should be associated to the selected FCs, and iii) how much LMC should pay to the selected FCs*. In this paper, we study the user association problem for LMC, in which the FCs are motivated to serve the rLMC users in a QoS-ensured manner.

Various works on incentive-based user association [13]–[17] exist in the literature. Paris *et al.* [13] proposed a reverse auction framework wherein small cells compete to serve the macrocell excess user demand. Based on the same auction-based approach, Trakas *et al.* [15] proposed a scheme, in which multiple LMCs compete for the limited FC capacity to offload their excess demand. Recently, Qi and Wang [17]

The authors are with the Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur, India (Email: satendra.cse@iitkgp.ac.in; smisra@sit.iitkgp.ernet.in).

[1] The terms smallcell, hotspot, and femtocell mean the same in the context of the problem discussed in the paper, and will, thus, be used interchangeably.

used game theoretic scheme to motivate FC to adopt the hybrid access mode. The existing countermeasure schemes [13]–[15] are tailored for very idealistic situation. In particular, during FC selection, for rLMC user association, LMC excludes its own capacity which results in under utilization of radio resource. This also incurs excessive payment to FC owners. Further, in the existing approaches [13]–[17], the rate and SNR relationship are modeled using *Shannon's capacity theorem*, instead of lookup tables given in the 3GPP standard [8].

Considering these limitations, it is important to design a holistic user association scheme for LTE-Advanced HetNets, which not only motivates the FCs to serve the rLMC users, but also ensures QoS of the associated rLMC users while taking into account the serving capacity of LMC. The major contributions of the paper are as follows:

i) We study a market situation with single LMC base station and multiple FCs. The LMC aims to associate its registered users to the FCs with minimum cost.
ii) We formulate the user association problem as a multi-unit single-item procurement auction, wherein the FC builds a quotation table (QT) based on the power budget, and the number of physical RBs left after serving its SUs.
iii) The proposed scheme implicitly enforces healthy competition among FCs, and permits multiple bids based on the number of rLMC users FC is willing to associate.
iv) We propose an energy-efficient resource allocation scheme, which ensures the QoS requirement of rLMC users associated to different FCs.

II. RELATED WORKS

In the literature, there exists various studies on user association for HetNets. Based on the primary motive of study we classify them broadly into two categories — *Interference Avoidance* [**?**], [6], [18]–[20] and *Load balancing* [3], [21]–[24].

Ha and Le [18] proposed an uplink power control algorithm for user association and interference avoidance using game-theoretic approach. In particular, authors proposed a load-aware user association scheme followed by iterative uplink power allocation scheme for HetNets. The power updating schemes take into account the inference from other users and QoS requirements. Recently, Mishra and Murthy [19] addressed the interference avoidance problem by adjusting power spreading. In the proposed scheme, each FC optimizes its power spreading factor for maximizing energy efficiency and mitigating interference, while ensuring the SINR of its associated users. Further, Amine *et al.* [20] focused on the same problem from a game-theoretic perspective. The authors obtained mapping between users and FCs based on the SINR profile of the former using a many-to-many matching game. These countermeasure approaches are effective in co-channel interference avoidance and limited to ensuring minimum user SINR. Since GBR service requires sufficient resource allocation along with minimum SINR, none of the aforementioned work is adequate for GBR services.

In the context of load balancing, Son *et al.* [21] proposed a *soft load balancing* scheme and derived a closed-form expression of outage probability as a function of load ratio and user assigned bandwidth for CDMA and OFDMA-based networks. Further, the authors evaluated the optimal load fraction which minimizes the outage probability of the users. Kim *et al.* [22] proposed optimal resource allocation scheme for load balancing among LMCs and FCs. The resource allocation scheme motivates the FC to operate in hybrid mode and also ensures that the average throughput of the SUs is greater than that of the offloaded rLMC users. Recently, Park and Kim [23] proposed mute-based scheme for load balancing. The authors selected an optimal set of small cells, typically near to macrocell, to enhance the SINR of edge users associated to outer smallcells. The load is balanced by making the edge users attracted towards outer smallcell through improved SINR and power biasing techniques. The selection of smallcells for mute is based on user distribution and traffic demand statistics. The proposed scheme is effective for users who suffer from poor coverage. Further, Demirci and Korcak [24] addressed the problem of load balancing between WiFis and cellular network. In this work, the authors proposed various algorithms based on cell breathing, in which lightly (heavily) loaded WiFi nodes expand (shrink) their respective communication ranges. The WiFi nodes adjust the transmitting power of beacon packets, for a given user distribution, while ensuring full coverage. Further, the algorithms are extended to accept macrocell users based on the existing load of WiFi. These works are non-incentive based and usable in scenarios where macrocell and FC networks are owned by the same CNPs. However, these techniques are inefficient if the FCs are owned by third parties. For this, we need an incentive-based scheme, which motivates the third party owned FC to serve associated rLMCs.

In the existing literature, few works [13]–[16] investigated the problem of incentive-based load balancing through data offloading. Paris *et al.* [13] proposed a reverse auction framework to offload the macrocell excess demand. Here, multiple smallcells compete with each other to serve the macrocell's excess demand. The mechanism also ensures the integrity of the bid value received from the from smallcells. Likewise, Dong *et al.* [14] proposed a reverse auction-based solution for dynamic data offloading. In the proposed scheme, the service area is divided into different sectors and the macrocell estimates the future demand for each sector. Based on the estimated demand and bid received from smallcell, macrocell decides the amount of bandwidth to purchase from each smallcell. In contrast to these schemes, Trakas *et al.* [15] proposed an auction-based approach for small cell operators, in which multiple macrocells compete for the available smallcell capacity to offload their excess demand. Iosifidis *et al.* [16] proposed an iterative double-auction based mechanism, in which multiple macrocells compete to offload data over various smallcell networks. The macrocells try to maximize the amount of data to offload, whereas the smallcells try to minimize the cost it incurs to accept the offloading request. The authors achieved the social welfare by maximizing the difference between these two conflicting objectives.

Synthesis: Based on our analysis of the existing literature on user association, we infer that there is a need to design a scheme which associates rLMC users to available FC and should aligned with LTE-Advanced HetNets architecture. In particular, the scheme should enable LMC to motivate the FCs, owned by home-based-users, to serve the rLMC users in QoS-ensured manner.

III. System Model

We consider a two-tier heterogeneous wireless network consisting of a single LMC base station surrounded by $|\mathcal{F}|$ FCs, where $\mathcal{F} = \{F_1, \cdots, F_f, \cdots, F_{|\mathcal{F}|}\}$ is the set of FCs. The LMC is in-charge to serves rLMC users. We denote the set of rLMC by $\mathcal{R}_l$, which is characterized by homogeneous Poisson point process (PPP), namely $\phi_l(\lambda_l)$, where λ_l is the rLMC user density. We consider that the FCs are within the communication range of LMC, as shown in Fig. 1. The FC operates in hybrid access mode, and hence, serve two type of users — guest users (GUs) and subscribe users (SUs). The users registered with F_f^{th} FC are known as its SUs, and is denoted by $\mathcal{S}_{F_f}$. The deployment $\mathcal{S}_{F_i}$ is characterized by a homogeneous PPP, denoted by $\phi_{F_f}(\lambda_{F_f})$, where λ_{F_f} is the SU density. The rLMC users who are in the communication range of F_i^{th} FC, are known as GUs of the respective FC, which we denote by $\mathcal{G}_{F_f}$. It may be noted that, all the SUs can only be served by the corresponding FCs, whereas the GUs can either be served by the corresponding FC or the LMC. The FC invests its radio resources, namely resource block (RB) and power, to serve GUs and hence, receives the incentive from the LMC. Each FC serves its SUs before serving any GU.

Based on the existing literature [22], we consider that the LMC and the FCs operate in the orthogonal frequency spectrum. The bandwidth allocated to cells is further divided into small RBs, each of size 180 KHz. Let $\mathcal{N}^l$, $\mathcal{P}^l_{max}$, $\mathcal{N}^{F_f}$, and $\mathcal{P}^{F_f}_{max}$ denote the number of LMC RB, LMC power budget, number of F_f^{th} FC RB, and F_f^{th} FC power budget, respectively. Hence, the maximum power allowed in each RB of LMC and F_f^{th} FC are given by $p^l_{max} = \frac{\mathcal{P}^l_{max}}{\mathcal{N}^l}$ and $p^{F_f}_{max} = \frac{\mathcal{P}^{F_f}_{max}}{\mathcal{N}^{F_f}}$, respectively. For given channel condition and noise, the maximum power for RB defines the best modulation and coding scheme (MCS) that a cell can employ for its user. Further, each cell finds the discrete power level for desired MCS. We define the possible power levels for the c^{th} cell[2] as totally ordered set $\mathbf{P}^c = \{\mathbf{p}^c_1, \mathbf{p}^c_2, \cdots \mathbf{p}^c_{|\mathbf{P}_c|}\}$, where $|\mathbf{P}_c| \leq 15$ [25], *i.e.*, for a cell with worst channel condition, it is not possible to employ best MCS.

A. Channel Model

We assume that each user experiences invariant channel conditions for all the RBs, irrespective of its location inside the cell. We model the channel gain as a function of path-loss, fast fading, and shadowing. Therefore, the channel gain between user u and cell c can be expressed as:

$$\Gamma^c = K\beta^c\zeta^c(L^c_{avg})^{-\alpha^c} \tag{1}$$

where K is a system constant dependent on both transmitter and receiver antenna gains and heights, β^c is the fast fading gain occurring due to the multi-path propagation, ζ^c is the slow fading gain due to the shadowing effect, L^c_{avg} is the statistical average of distance of all the SUs (rLMC) in FC (LMC), and α^c is the path loss exponent. This may be noted that, the variation in the channel condition for a given user with respect to different RBs, results in a complex problem of sub-channel allocation [26]. The signal-to-noise-ratio (SNR) experience by the u^{th} user in the r^{th} RB of the c^{th} cell is given by–

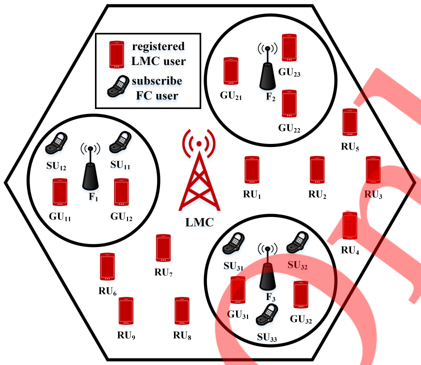


Fig. 1: Two-tier heterogeneous wireless networks

$$SNR^c_{r,u} = p^c_{r,u}\left(\frac{\Gamma^c}{\sigma^2}\right) \tag{2}$$

where $0 \leq p^c_{r,u} \leq p^c_{max}$ is the transmitting power in the r^{th} RB for the u^{th} user, and σ^2 is the additive white Gaussian noise (AWGN).

B. Throughput Model

The achievable throughput from the r^{th} RB, namely $T^c_{r,u}$, for the c^{th} cell and arbitrary user u, depends on the MCS employed for the corresponding user in the RB. In particular, the MCS $M^c_{r,u}$ defines the code rate, and number of bits per OFDM symbol as $\mathcal{C}^c_{r,u}$ and $\mathcal{B}^c_{r,u}$, respectively. We consider normal cyclic prefix with 7 OFDM symbol per sub-carrier. Out of the 84 resource elements, only $\frac{3}{4}^{th}$ of these, *i.e.* 63, are used for transmitting data and the rest are used for the control channel.

$$T^c_{r,u}(\mathbf{p}^c_{r,u}) = \frac{63 \times \mathcal{C}^c_{r,u} \times \log_2(\mathcal{B}^c_{r,u})}{466.67} \tag{3}$$

For given power allocation $\mathbf{p}^c_{r,u} \in \mathbf{P}^c$, we compute the SNR assuming that the channel gain and noise for the RBs are independent of the power allocated to RBs [27]. The SNR is mapped to the channel quality indicator (CQI), mentioned as follows:

$$CQI^c_{r,u} = \begin{cases} 1.17(SNR^c_{r,u}) - 40.65, & \text{if } c \in \mathcal{F} \\ 0.55(SNR^c_{r,u}) + 3.60, & \text{if } c \in \{l\} \end{cases} \tag{4}$$

The obtained CQI can be further mapped to the corresponding MCS using the lookup table given in the 3GPP specification [5], [8].

LTE-based cells employ the same MCS in all the RBs assigned to a particular user. Further, the RBs are indistinguishable for the users. A cell c can allocate the same power $\mathbf{p}^c_u$ in the RBs assigned to them. Hence, in order to satisfy

---

[2] We use c to index both the LMC and the FCs.

minimum data requirement $\mathbf{d}_u$, cell c allocates $\mathbf{r}_u^c$ number of RBs. Mathematically,

$$\mathbf{r}_u^c = \left\lceil \frac{\mathbf{d}_u}{T_u^c(\mathbf{p}_u^c)} \right\rceil \tag{5}$$

where $\mathbf{p}_u^c \leq p_{max}^c$. We also define the maximum throughput of a cell $\mathcal{T}_{max}^c$ as the sum of throughputs of all the RBs, when operated at the maximum allowed power level $\mathbf{p}_{|\mathbf{P}|}$. Mathematically, $\mathcal{T}_{max}^c = \mathcal{N}^c T_{r,u}^c(\mathbf{p}_{|\mathbf{P}^c|}^c)$.

IV. PROBLEM FORMULATION

We formulate LMC user association problem as a *multi-unit single-item procurement game*. The problem is double-pronged. Initially, the *potential FCs* discussed in Section V, build their quotation table (QT). After receiving the QTs from potential FCs, the LMC decides the portion of excess demand to offload to each FC. The collective throughput demand, $\mathcal{T}_{dem}^l$, from the rLMC users, for a given time period $[t, t+\delta]$, is time-varying. Moreover, during the peak time, it surpasses the maximum throughput, $\mathcal{T}_{max}^l$, of the LMC. To meet stochastic demands of such type, one possible solution for CNPs is to deploy their own small cells to offload the excess demand, $\mathcal{T}_{exc}^l = \mathcal{T}_{dem}^l - \mathcal{T}_{max}^l$, from the LMC. The deployed small cell's service schedule (ON/OFF) is optimized for a given demand pattern over time. However, this solution will be non-profitable, if the pattern of excess demand, $\mathcal{T}_{exc}^l$, is highly unpredictable and sparse over time [14].

We argue that offloading this excess demand to FCs is a promising solution for LMCs, where the LMCs motivate and pay the FCs to accept a fraction of $\mathcal{T}_{exc}^l$. If a FC agrees to accept, it liquidates power and RBs to serve GUs. Thus, the portion of $\mathcal{T}_{exc}^l$ which a FC accept is bounded by the collective demands of its GUs. Additionally, as the FC prioritizes its SUs, the portion of $\mathcal{T}_{exc}$ is also limited by the remaining power and the number of RBs after serving its SUs. We represent the remaining power and RB of F_f^{th} FC as $\mathcal{P}_{rem}^{F_f} = \mathcal{P}_{max}^{F_f} - \mathcal{P}_{sub}^{F_f}$ and $\mathcal{N}_{rem}^{F_f} = \mathcal{N}^{F_f} - \mathcal{N}_{sub}^{F_f}$, respectively, where $\mathcal{P}_{sub}^{F_f}$ and $\mathcal{N}_{sub}^{F_f}$ are the estimated power and number of RBs used for serving SUs.

We model the intercommunication among the LMC and the FCs as a *multi-unit single-item procurement game*, in which the LMC acts as a buyer and the FCs act as the suppliers.

1) The buyer procures $\mathcal{T}_{exc}^l$ amount of throughput.
2) A supplier $F_f \in \mathcal{F}$ is said to be a potential supplier, if it satisfies the following three conditions:
   i) $\mathcal{R} \cap \mathcal{G}^{F_f} \neq \phi$.
   ii) $\mathcal{N}_{rem}^{F_f} > 0$.
   iii) $\mathcal{P}_{rem}^{F_f} > 0$.

   We denote the set of such potential suppliers as $\mathcal{F}'$.
3) Each $F_{f'} \in \mathcal{F}'$ builds a QT $\mathcal{Q}^{F_{f'}}$ using Algorithms 2 and 3. The QT consists of $q(F_{f'}) \leq |\mathcal{G}^{F_{f'}}|$ quantity-price pairs, $\{[Q_1^{F_{f'}}, P_1^{F_{f'}}], [Q_2^{F_{f'}}, P_2^{F_{f'}}], \ldots, [Q_{q(F_{f'})}^{F_{f'}}, P_{q(F_{f'})}^{F_{f'}}]\}$.

After getting the QTs from all the potential suppliers, the LMC decides the quantity to be procured from them. The objective of LMC is to minimize the overall incentive paid to the FCs, for a given excess demand from rLMC. The objective function corresponding to a LMC is given in Equation (6a). Furthermore, constraints in Equations (6b) and (6d) restrict the LMC to select at most one bid for a given supplier. Constraint in Equation (6c) represents the cumulative procurement of throughput from FCs satisfying the rLMC excess demand.

$$\min \quad \sum_{f'=1}^{|\mathcal{F}'|} \sum_{q=1}^{q(F_{f'})} x_q^{F_{f'}} P_q^{F_{f'}} \tag{6a}$$

$$\text{s.t.} \quad \sum_{q=1}^{q_f} x_q^f \leq 1 \qquad \forall F_f \in \mathcal{F}' \tag{6b}$$

$$\sum_{f=1}^{|\mathcal{F}'|} \sum_{q=1}^{q(F_{f'})} x_q^{F_{f'}} Q_q^{F_{f'}} \geq \mathcal{T}^{exc} \tag{6c}$$

$$x_q^{F_{f'}} \in \{0,1\} \qquad \forall F_{f'} \in \mathcal{F}', 1 \leq q \leq q(F_{f'}) \tag{6d}$$

where $Q_q^{F_{f'}}$ and $P_q^{F_{f'}}$ are the q^{th} quantity-price in $\mathcal{Q}^{F_{f'}}$. $x_q^{F_{f'}}$ is a binary decision variable associated with the q^{th} entry of $\mathcal{Q}^{F_{f'}}$ and is set to unity, if the LMC decides to procure the corresponding quoted quantity from the supplier. The stated optimization problem is binary integer linear program (BILP) and hence, we use *branch-and-bound* method to obtain the optimal solution.

V. SOLUTION APPROACH

We propose a two-phase solution consisting of (a) *quotation building* and (b) *procurement process*. In the quotation building phase, each FC builds its QT and sends it to the LMC. After receiving the QT, the LMC decides the fraction of $\mathcal{T}^{exc}$ to be procured from each FC.

A. Quotation Building

Each FC estimates the available power and the number of RBs after serving its own SU. The FCs estimate the number of remaining RBs, if the best possible MCS is employed for each RB as discussed in Algorithm 1. The FC participates in the LMC initiated procurement phase with remaining number of RBs and power. It is worthy to note that Algorithm 1 does not give the final allocation of resources to SU. The FC may further reduce the power allocated to each SU (or lower level MCS), if FC has left over RBs after the procurement process.

With available power budget and number of RBs, the FC serves the GUs, in its coverage range. Further, based on power and RB usage, each FC builds the quotation table. The detailed procedure is given in Algorithm 2. In particular, the FC computes the number of RBs and corresponding power level assigned to each of the GUs. Initially, the FC assigns the lowest possible power level to the unserved GUs and compute the RBs required to serve their minimum throughput requirement. In case of unavailability of RBs, the FC selects a GU $g \in \mathcal{G}_{ser}$ having maximum number of assigned RBs ($\mathbf{r}_g$) and increases its current power level $\mathbf{p}_g$ by unity. Since, the increased power level serves the throughput requirement of GU in lesser number of RBs, the FC allocates the free RBs

Algorithm 1: Femtocell Power and RB Estimation

```
Inputs  : $\mathcal{S}$, $\mathbf{D}_{\mathcal{S}}$, $\mathcal{N}$, $\mathcal{P}$, $\mathbf{P}$
Outputs: $\mathcal{P}_{rem}$, $\mathcal{N}_{rem}$
$\mathcal{N}_{rem} = \mathcal{N}$; $\mathcal{P}_{rem} = \mathcal{P}$; $\mathcal{S}_{temp} = \mathcal{S}$
while $\mathcal{S}_{temp} \neq \phi$ or $\mathcal{N}_{rem} = 0$ do
    Choose user $s \in \mathcal{S}_{temp}$ with minimum $\mathbf{d}_s$
    Calculate number of RB, $\mathbf{r}_s$ at $\mathbf{p}_{|\mathbf{P}|}$
    if $\mathbf{r}_s \leq \mathcal{N}_{rem}$ then
        $\mathcal{N}_{rem} = \mathcal{N}_{rem} - \mathbf{r}_s$
        $\mathcal{P}_{rem} = \mathcal{P}_{rem} - \mathbf{p}_{|\mathbf{P}|}\mathbf{r}_s$
        $\mathcal{S}_{temp} = \mathcal{S}_{temp} \setminus \{s\}$
    else
        break;
```

to user $g \in \mathcal{U}_{unser}$. For each successive inclusion, of GU the FC updates the QT based on the power and RB usage. The FC continues this process until there is no unserved user around or FC has unavailability of power or RB.

Algorithm 2: Resource Allocation to Unserved Users

```
Inputs  : $\mathcal{G}$, $\mathbf{D}_{\mathcal{G}}$, $\mathcal{N}_{rem}$, $\mathcal{P}_{rem}$, $\mathbf{P}$
Outputs: $\mathsf{P}$, $\mathsf{R}$
$\mathcal{G}_{ser} = \phi$, $\mathcal{G}_{unser} = \mathcal{G}$
For each user $g \in \mathcal{G}$ set$\mathbf{p}_g = \mathbf{r}_g = 0$
while $\mathcal{G}_{unser} \neq \phi$ OR no change in $\mathcal{G}_{unser}$ do
    while $\mathcal{G}_{unser} \neq \phi$ or $\mathcal{N}_{rem}^{F_f} = 0$ or $\mathcal{P}_{rem}^{F_f} = 0$ do
        Select $g \in \mathcal{G}_{unser}$ with minimum $\mathbf{d}_g$
        Calculate $\delta_{\mathbf{p}_g}$, and $\delta_{\mathbf{r}_g}$ at $\mathbf{p}_g + 1$
        if $\delta_{\mathbf{p}_g} \leq \mathcal{P}_{rem}^{F_f}$ and $\delta_{\mathbf{r}_g} \leq \mathcal{N}_{rem}^{F_f}$ then
            Update $\mathcal{N}_{rem}$, $\mathcal{P}_{rem}$, $\mathbf{p}_g$, $\mathbf{r}_g$
            $\mathcal{G}_{ser} = \mathcal{G}_{ser} \cup \{g\}$, $\mathcal{G}_{unser} = \mathcal{G}_{unser} \setminus \{g\}$
            Update-Quotation-Table ($\mathcal{G}_{ser}$, $\mathsf{P}$, $\mathsf{R}$)
    while $\mathcal{N}_{rem}^{F_f} \leq 0$ do
        Select $g \in \mathcal{G}_{ser}$ with maximum $\mathbf{r}_g$ and
         $\mathbf{p}_g < \mathbf{p}_{|\mathbf{P}|}$
        Calculate $\delta_{\mathbf{p}_g}$, and $\delta_{\mathbf{r}_g}$ at $\mathbf{p}_g + 1$
        if $\delta_{\mathbf{p}_g} \leq \mathcal{P}_{rem}$ and $\delta_{\mathbf{r}_g} \leq \mathcal{N}_{rem}$ then
            Update $\mathcal{N}_{rem}$, $\mathcal{P}_{rem}$
```

Let $\mathcal{G}_{ser,q} \subset \mathcal{G}$ denotes the set of users FC willing to serve in q^{th} bid. Further, $\mathbf{r}_g$ and $\mathbf{p}_g$ denote the number of RBs and power allocated to user $g \in \mathcal{G}_{ser,q}$, respectively. Then, the price for the corresponding bid, denoted by P_q, is governed by –.

$$P_q = \varrho \sum_{g \in \mathcal{G}_{ser,q}} \mathbf{r}_g \mathbf{p}_g + \mathcal{R} \frac{\mathcal{N}}{\mathcal{N}_{rem}} \tag{7}$$

It may be noted that the price is monotonically increasing linear function of the power liquidated by the FC for serving guest users, whose slope depends on the cost per unit of power ϱ for FC. Furthermore, we model the price as a decreasing function of the remaining number of RBs. Thus, a FC with higher $\mathcal{N}_{rem}$ quotes lesser price for a given throughput. The intuition behind this modeling is to enforce healthy competition among the FCs. In the price function, we also consider the reluctant behavior of FC, denoted by $\mathcal{R} \in [0 \quad \infty]$, to serve the GUs, due to secrecy and privacy issues of their SUs. FCs with $\mathcal{R} = 0$ are considered to be generous, because they do not make profit for serving GUs. Using $\mathcal{G}_{ser,q}$ and the corresponding power allocation, FC updates the QT using Algorithm 3. Specifically, FC computes the overall power and RB usage corresponding to each bid and employs Equation (7) to calculate price (P_q) of the q^{th} bid.

Algorithm 3: Update Quotation Table

```
Inputs : $\mathcal{G}_{ser}$, $\mathsf{P}$, $\mathsf{R}$, $\mathcal{N}$, $\mathcal{N}_{rem}$, $\mathcal{R}$
Output: $\mathcal{Q}$
$q = |\mathcal{G}_{ser}|$, $Q_q = 0$, $Pow_q = 0$
for each $g \in \mathcal{G}_{ser}$ do
    $Q_q = Q_q + \mathbf{d}_g$
    $Pow_q = Pow_q + \mathbf{p}_g \mathbf{r}_g$
    Calculate price $P_q$ using Equation 7.
```

B. *Procurement Process*

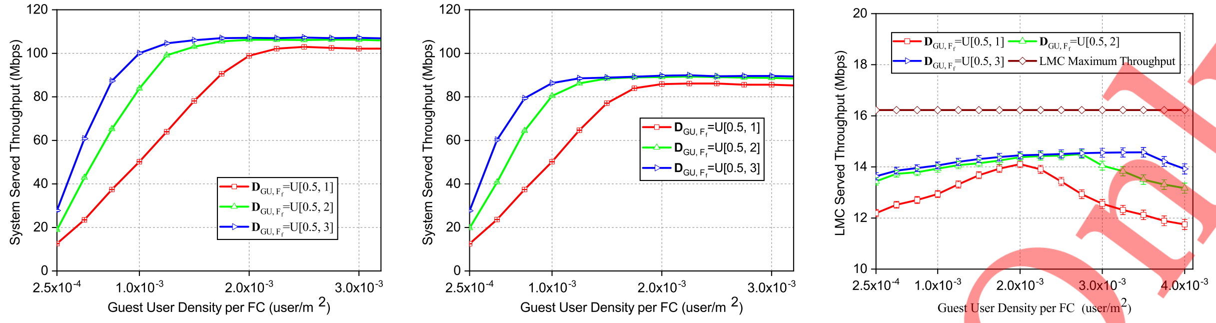
This phase is meant for an LMC, where it decides the particular bid for a given supplier to minimize the incentive to be paid for a given excess demand. The LMC solves the optimization problem defined in Section IV using the *branch-and-bound* method. In case of existence of an infeasible solution, *i.e.*, when the excess demand is exceeded beyond the total service of all the FCs, the LMC selects the last bid for each potential supplier.

VI. PERFORMANCE EVALUATION

Simulation results to evaluate the performance of the proposed scheme are excerpted in this paper. The system parameters used for simulation are aligned with the 3GPP specification [8]. Some of the vital parameters are reported in Table I. We consider the following performance metrics for the evaluation of the proposed scheme. 1) Throughput served by the system, which is the sum of the minimum throughput requirements of satisfied rLMC users served by both FCs and LMC, 2) The throughput served by LMC, 3) Resource block utilization of FC Networks (FCN), 4) Power usage of FCN networks. As mentioned in Section V, we employ a similar procedure for resource allocation, the RB and power usage of LMC results the same pattern that of FCN, 5) Number of served rLMC users, and 6) LMC payment towards the FCN.

A. *Experimental Setup*

We use the discrete event simulator, MATLAB, for extensive simulation of the proposed scheme. We model the two-tier network by deploying 9 non-overlapping FCs in the coverage range of LMC. For the entire simulation, we assume that the minimum requirement of SUs for each FC, i.e., $\mathbf{D}_{\mathcal{S},F_f} = U[0, 0.5]$. For an insight of the proposed scheme, we enforce each rLMC user to be the GU of one of the FCs and vary the density of the GU in the FC (λ_{GU,F_f}) instead of the rLMC



(a) System throughput @ $\lambda_{FCN} = 1 \times 10^{-3}$ (b) System throughput @ $\lambda_{FCN} = 2 \times 10^{-3}$ (c) LMC served throughput

Fig. 2: Throughput variation with guest user density

TABLE I: Experimental Setup

Parametrs	LMC	FC
Cell radius	1,000 m	50 m
Maximum Tx Power	46 dBm	20 dBm
Slow fading gain standard deviation	8 dB	4 dB
Fast fading gain distribution	Mean = 1	
Power density (White noise)	-174dBm/Hz	
Downlink Bandwidth	10 MHz	
Resource block	50 RBs; 180 KHz each	
Path loss exponent	4	
Antenna configutation	SISO	

user density. It may be noted that the proposed scheme is also valid for rLMC user density, which is given in Section III. Furthermore, we consider the minimum requirement of rLMC users (here GUs) $\mathbf{D}_{GU,F_f} \in \{U[0.5,1], U[0.5,2], U[0.5,3]\}$, and SU density in each cell $\lambda_{F_f} = \lambda_{FCN} \in \{1 \times 10^{-3}, 2 \times 10^{-3}\} \quad \forall F_f \in \mathcal{F}$. The data points are averaged over 100 runs of simulations and the corresponding 95% confidence interval is shown in Figs. [2]-[3].

B. Result and Discussion

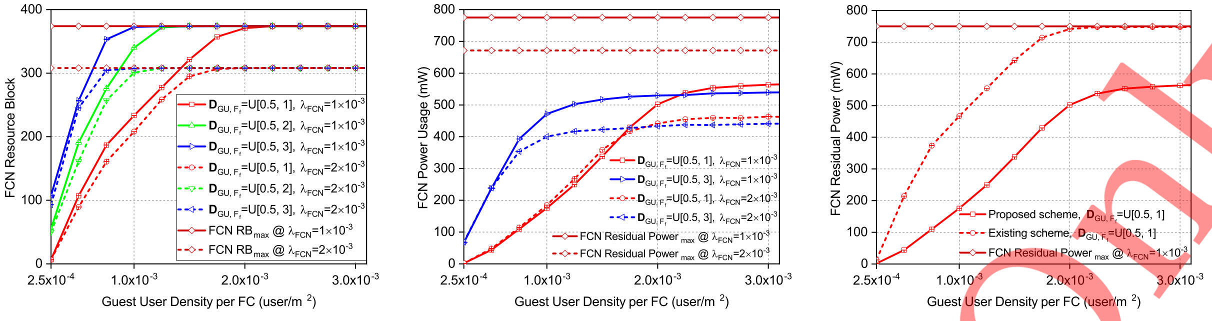
Fig. 2 depicts the variation of the served throughput with respect to the GU density per FC for different values of $\mathbf{D}_{GU,F_f}$ and λ_{FCN}. In Fig. 2(a), we keep $\lambda_{FCN} = 1 \times 10^{-3}$ and vary $\mathbf{D}_{GU,F_f}$. We observe that the system throughput increases linearly for a given $\mathbf{D}_{GU,F_f}$ up to a certain value of GU density λ_{GU,F_f} and saturates afterward. Also, the GU density, at which the system throughput gets saturated, is large (small) for lower (higher) value of $\mathbf{D}_{GU,F_f}$. This indicates that to fully utilize the system throughput, there is a need of high user density with lower throughput demands. Fig. 2(b) follows a similar pattern for $\lambda_{FCN} = 2\times10^{-3}$. In contrast, the convergence point of system served throughput for all $\mathbf{D}_{GU,F_f}$ is lower, as compared to that exhibited in Fig. 2(a). With the increase in SU density, the FCs are left with less power and RBs, according to Algorithm 1, and hence, are able to serve lesser throughput of their GUs.

From Fig. 2(c), it appears that the LMC served throughput is an increasing function of GU density (equivalently, rLMC user density) up to a certain point, which aligns with our intuition. Contrarily, the LMC served throughput starts decreasing afterward. As discussed in Algorithm 2, the LMC allocates power and RBs to rLMC users in a non-decreasing order of their requirement. For denser deployment, the large number of rLMC users with sparse demand engage the LMC resources and decreases the cumulative LMC served throughput. Further, we observe that the rate of decrement of served throughput is higher in case of $\mathbf{D}_{GU,F_f} = U[0.5,1]$ than $\mathbf{D}_{GU,F_f} = U[0.5,3]$ due to the presence of large number of user having sparse demand.

In Fig. 3, we analyze the power and RB usage of FCN against the GU density, for different combinations of $\mathbf{D}_{GU,F_f}$ and λ_{FCN}. In Fig. 3(a), the maximum RBs for FCN, $FCNRB_{max}$, is the sum of the remaining RBs after serving the SU, over all FCs. For a given $\mathbf{D}_{SU,F_f}$, the $FCN_{rb,max}$ possesses higher value for the lower value of λ_{FCN}. Further, for given λ_{FCN}, we observe that the RB usage is an increasing function of GU density, until it reaches the corresponding $FCN_{rb,max}$, where the rate of increase is governed by the GU requirements. Furthermore, it is interesting to note that, for a given $\mathbf{D}_{GU,F_f}$ and GU density, the FCN RB usage is more for lower SU density λ_{FCN} (or abundant remaining RBs). As discussed in Algorithm 2, each FC tries to reduce the power consumption by allocating more number of RBs, and justify the FCN RB usage pattern.

Fig. 3(b) reveals the FCN power usage for different values of GU density. We compare the power profile for different variants of $\mathbf{D}_{GU,F_f}$ and λ_{FCN}. We observe an immediate increase followed by saturating pattern of FCN power usage. Unlike FCN RB usage, we highlight two vital insights obtained from Fig. 3(b). First, we encounter lower power usage in case of λ_{FCN} for a given GU density and $\mathbf{D}_{GU,F_f}$. Indeed, FCN tries to save power at the cost of the increase in the number of RBs. Second, regardless of the value of $\mathbf{D}_{GU,F_f}$ and λ_{FCN}, the FCN power usage is less than the maximum FCN power, which is the sum of the residual power after serving SUs, over all FCs ($FCN_{pow,max}$). The proposed scheme intelligently increases the power allocated to (equivalently MCS of) each rLMC for a given $\mathbf{D}_{GU,F_f}$, which saves significant amount of power. Hence, it is power-efficient.

We compare the power usage of proposed scheme with the work of Qi and Wang [17], as shown in Fig. 3(c). In



(a) FCN RB usage

(b) FCN power usage

(c) FCN Power usage for proposed and existing approch

Fig. 3: Power and Resource block usage variation with guest user density

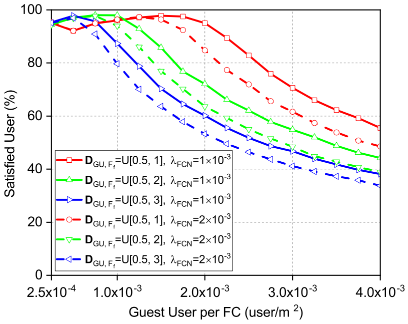
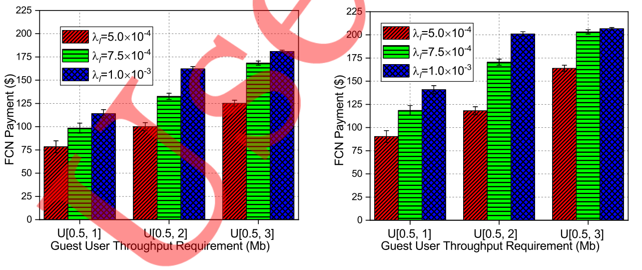


Fig. 4: Variation in satisfied user percentage with guest user density



(a) FCN payment @ $\lambda_{FCN} = 1 \times 10^{-3}$

(b) FCN payment @ $\lambda_{FCN} = 2 \times 10^{-3}$

Fig. 5: FCN payment variation with LMC user throughput demand

the existing scheme, the allocated RBs are employed with highest possible power which increases the overall power usage. However, in the proposed scheme, we allocate the minimum possible power to the RBs allocated to the user according to throughput requirement. Hence, the FCN power usage is lower in the proposed scheme compared to the existing scheme.

Fig. 4 depicts the change in the satisfied GU (with respect to GU density) for different combinations of $\mathbf{D}_{GU,F_f}$ and λ_{FCN}. As per the procurement procedure, the LMC associates more rLMC users to FCN with an increase in the excess demand of rLMC. Hence, we observe an increasing percentage of satisfied rLMC users. With further increase in rLMC user density, we observe a decrease in the percentage of satisfied rLMC users. The reason behind this decrease is the throughput bound of FCN, which is the sum of maximum throughput that each FC can serve after serving SUs. Beside this, we observe an initial decrease in the percentage of satisfied users for $\mathbf{D}_{GU,F_f} = U[0.5, 1]$. This is due to the under estimation of excess demand of rLMC users by LMC. It is noteworthy that LMC decides to procure throughput if the excess demand from rLMC users exceeds its maximum capacity $\mathcal{T}^c_{max}$.

Fig. 5 represents the FCN payment received from LMC, for different LMC users throughput requirement. For Figs. 5(a) and 5(b), we set the FCNs' subscribe user destiny as 1×10^{-3} and 2×10^{-3}, respectively. Further, we plot FCN payment for various GU density in each FCN. We observe that λ_{FCN} has an adverse effect on the price that LMC pays to FCN for serving LMC users. In other words, LMC pays more if FCs are heavily loaded with their SUs, by following Algorithm 1 and Equation (7). Additionally, we observe that the FCN payment increases with the increase in the guest user throughput requirement for given SU density. This immediate follows from the increase in power consumption of FCN to serve the associated guest users.

VII. CONCLUSION

In this paper, we proposed a multi-unit procurement-based scheme for user association in LTE-Advanced based HetNets. The proposed scheme enables LMC to exploit the power and RB of FCs to fulfill the rLMC user's excess demand, while ensuring both energy efficiency and QoS. We proposed a two-phase solution approach. First, each FC submits QT while taking its own user demands. Second, the LMC decides the procuring amount from potential FCs in a cost-effective manner. Results of the extensive simulation illustrates that the FCs' SU density has an adverse effect on rLMC user association. Furthermore, the FCs try to minimize power in the best possible manner to serve the rLMC users.

In the future, we plan to extend this work for more realistic ultra-dense FCs scenarios, in which the FCs may overlap with one another. Beside for meaningful procurement, we plan to design an admission control protocol for the case in which excess demand of rLMC users exceed the FC services.

REFERENCES

[1] "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016 - 2021," Tech. Rep., Mar. 2017.

[2] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, "Femtocells: Past, Present, and Future," *IEEE J. Sel. Areas in Commun.*, vol. 30, no. 3, pp. 497–508, Apr. 2012.

[3] R. Tao, W. Liu, X. Chu, and J. Zhang, "An Energy Saving Small Cell Sleeping Mechanism With Cell Range Expansion in Heterogeneous Networks," *IEEE Trans. on Wirel. Commun.*, vol. 18, no. 5, pp. 2451–2463, May 2019.

[4] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell Networks: A Survey," *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 59–67, Sep. 2008.

[5] P. K. Barik, C. Singhal, and R. Datta, "DAMS: D2D-Assisted Multimedia Streaming Service With Minimized BS Transmit Power in Cellular Networks," *Comp. Commun.*, vol. 144, pp. 149 – 161, 2019.

[6] D. Lynch, M. Fenton, D. Fagan, S. Kucera, H. Claussen, and M. ONeill, "Automated Self-Optimization in Heterogeneous Wireless Communications Networks," *IEEE/ACM Trans. on Netw.*, vol. 27, no. 1, pp. 419–432, Feb. 2019.

[7] M. Kamel, W. Hamouda, and A. Youssef, "Ultra-Dense Networks: A Survey," *IEEE Commun. Surv. Tut.*, vol. 18, no. 4, pp. 2522–2545, 2016.

[8] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures," The 3^{rd} Generation Partnership Project (3GPP), Technical Specification (TS) 36.213, 2011, version 10.1.0.

[9] H. Ekstrom, "QoS Control in the 3GPP Evolved Packet System," *IEEE Commun. Mag.*, vol. 47, no. 2, pp. 76–83, Feb. 2009.

[10] T. Mahmoodi, V. Friderikos, O. Holland, and H. Aghvami, "TCP-Aware Resource Allocation in OFDMA based Wireless Networks," in *Proc. of the 2^{nd} Int. Worksh. on Cross Layer Design*, Jun. 2009, pp. 1–5.

[11] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, "Downlink Packet Scheduling in LTE Cellular Networks: Key Design Issues and a Survey," *IEEE Commun. Surv. Tut.*, vol. 15, no. 2, pp. 678–700, Sec. 2013.

[12] E. Dahlman, S. Parkvall, J. Skold, and P. Beming, *3G Evolution: HSPA and LTE for Mobile Broadband*. Academic press, 2010.

[13] S. Paris, F. Martisnon, I. Filippini, and L. Clien, "A Bandwidth Trading Marketplace for Mobile Data Offloading," in *Proc. of IEEE INFOCOM*, Apr. 2013, pp. 430–434.

[14] W. Dong, S. Rallapalli, R. Jana, L. Qiu, K. K. Ramakrishnan, L. Razoumov, Y. Zhang, and T. W. Cho, "iDEAL: Incentivized Dynamic Cellular Offloading via Auctions," *IEEE/ACM Trans. on Netw.*, vol. 22, no. 4, pp. 1271–1284, Aug. 2014.

[15] P. Trakas, F. Adelantado, and C. Verikoukis, "A Novel Learning Mechanism for Traffic Offloading with Small Cell as a Service," in *Proc. of IEEE ICC*, Jun. 2015, pp. 6893–6898.

[16] G. Iosifidis, L. Gao, J. Huang, and L. Tassiulas, "A Double-Auction Mechanism for Mobile Data-Offloading Markets," *IEEE/ACM Trans. on Netw.*, vol. 23, no. 5, pp. 1634–1647, Oct. 2015.

[17] Y. Qi and H. Wang, "Incentive Pricing Mechanism for Hybrid Access in Femtocell Networks," *IEEE Commun. Lett.*, vol. 21, no. 5, pp. 1091–1094, May 2017.

[18] V. N. Ha and L. B. Le, "Distributed Base Station Association and Power Control for Heterogeneous Cellular Networks," *IEEE Trans. on Veh. Techn.*, vol. 63, no. 1, pp. 282–296, Jan. 2014.

[19] S. Mishra and C. S. R. Murthy, "Increasing Energy Efficiency via Transmit Power Spreading in Dense Femto Cell Networks," *IEEE Syst. J.*, vol. PP, no. 99, pp. 1–10, 2017.

[20] M. Amine, A. Walid, A. Kobbane, and S. Cherkaoui, "A Many-to-Many Matching Game in Ultra-Dense LTE HetNets," in *Proc. of THE 13^{th} Int. Wirel. Commun. and Mob. Comp. Conf. (IWCMC)*, Jun. 2017, pp. 1245–1250.

[21] H. Son, S. Lee, S. C. Kim, and Y. S. Shin, "Soft Load Balancing over Heterogeneous Wireless Networks," *IEEE Trans. on Veh. Techn.*, vol. 57, no. 4, pp. 2632–2638, Jul. 2008.

[22] D. Kim, T. Park, S. Kim, H. Kim, and S. Choi, "Load Balancing in Two-Tier Cellular Networks With Open and Hybrid Access Femtocells," *IEEE/ACM Trans. on Netw.*, vol. 24, no. 6, pp. 3397–3411, Dec. 2016.

[23] J. B. Park and K. S. Kim, "Load-Balancing Scheme With Small-Cell Cooperation for Clustered Heterogeneous Cellular Networks," *IEEE Trans. on Veh. Techn.*, vol. 67, no. 1, pp. 633–649, Jan. 2018.

[24] I. Demirci and O. Korcak, "Cell Breathing Algorithms for Load Balancing in Wi-Fi/Cellular Heterogeneous Networks," *Comp. Netw.*, vol. 134, pp. 140 – 151, 2018.

[25] X. Chen, H. Yi, H. Luo, H. Yu, and H. Wang, "A Novel CQI Calculation Scheme in LTE/LTE-A Systems," in *Proc. of Int. Conf. on Wirel. Commun. and Sig. Proces. (WCSP)*, Nov. 2011, pp. 1–5.

[26] D. Lpez-Prez, . Ladnyi, A. Jttner, H. Rivano, and J. Zhang, "Optimization Method for the Joint Allocation of Modulation Schemes, Coding Rates, Resource Blocks and Power in Self-Organizing LTE Networks," in *Proc. of IEEE INFOCOM*, Apr. 2011, pp. 111–115.

[27] D. Zhang, Z. Chang, T. Hmlinen, and F. R. Yu, "Double Auction Based Multi-Flow Transmission in Software-Defined and Virtualized Wireless Networks," *IEEE Trans. on Wirel. Commun.*, vol. 16, no. 12, pp. 8390–8404, Dec. 2017.

Satendra Kumar (S'16) received his M.Tech degree (Gold medalist) from National Institute of Technology (NIT), Raipur, India, in 2014. Prior to that, he obtained his B.Tech degree (Hons.) from Uttar Pradesh Technical University (UPTU), Lucknow, India, in 2012. Currently, he is pursuing his Ph.D. degree in Department of Computer Science and Engineering at Indian Institute of Technology (IIT) Kharagpur, India. His current research interests include resource allocation, network economics, mechanism design, and wireless caching.

Sudip Misra (SM'11) is a Professor in the Department of Computer Science and Engineering at the Indian Institute of Technology Kharagpur, India. He has published 9 books in the areas of wireless ad hoc networks, wireless sensor networks, wireless mesh networks, communication networks and distributed systems, and network reliability and fault tolerance, published by reputed publishers such as Cambridge University Press, Springer, Wiley, and World Scientific.

Prof. Misra is a fellow of the National Academy of Science (India). Currently, he is serving as the Associate Editor of the IEEE Transactions on Mobile Computing. He is also an Editor of the IEEE Transactions on Vehicular Technology.