

Mobile Cloud Storage Over 5G: A Mechanism Design Approach

Richa Siddavaatam, *Student Member, IEEE*, Isaac Woungang , *Senior Member, IEEE*,
Glaucio H. S. Carvalho, *Member, IEEE*, and Alagan Anpalagan , *Senior Member, IEEE*

Abstract—In order to meet the increasing demand for the data storage, 5G wireless networks embodying mobile edge computing (MEC) features arise as a compelling solution. In this paper, the dense heterogeneous network (HetNet) and the MEC infrastructure are exploited to propose a mobile cloud storage framework that minimizes the data transmission delay. The proposed framework is composed of two parts: A data management with error correction (DMEC) scheme, and a radio resource management (RRM) scheme. The DMEC scheme, derived from the redundant array of inexpensive disks (RAID) technology, is implemented in the user equipment (UE) side, and it intelligently exploits the overlapping coverage of HetNet to minimize the transmission delay. On the other hand, the RRM scheme, based on mechanism design, presents the physical resource block allocation problem as a graph coloring problem and performs the radio resource allocation in multiuser scenario to maximize the network performance. The RRM scheme also comprises a pricing algorithm, which calculates the price a UE needs to pay for the resources. The proposed RRM scheme exhibits several desirable characteristics such as incentive compatibility, efficiency, and truthfulness, all derived from the Vickrey–Clarke–Groves mechanism. Simulation results are presented, showing that the proposed framework when compared to baseline techniques, minimizes the transmission delay by 10²%, which places our proposal as effective and efficient solution for the mobile cloud storage problem.

Index Terms—5G wireless network, mobile cloud storage, mobile edge computing (MEC), mobile cloud computing (MCC), radio resource management (RRM) design.

I. INTRODUCTION

AN INHERENT human desire to record and save everyday moments allied to the IoT-driven business demand for data analytics and storage has led to a mobile data deluge. Global cloud IP traffic will reach 19.5 ZB (1.6 ZB per month) by the end of 2021, up from 6.0 ZB per year (499 EB per month) in 2016 [1]. According to Cisco, this big data phenomenon might be translated into business opportunities for another storage technology, namely the mobile edge computing (MEC). In a

nutshell, MEC is the next generation of mobile cloud computing solutions featured by a one-hop away small cloud data center. It is meant to enable mobile network operators (MNO) to successfully cope with the massive mobile data transfer at the edge without the need to overwhelm the core network and the Internet with unnecessary data transmissions. To amplify the transmission capacity from mobile devices to the MEC through base stations (BSs), the MNOs are shifting their access networks to a dense 5G architecture where thousands of small cells armed with their MEC infrastructures are overlaid by a macro cell. This 5G-based mobile cloud storage ecosystem arises as an efficient and effective response to the big data phenomenon at the edge.

Recently, the MEC paradigm has received a great deal of attention in the literature. The need to supply mobile devices with computational and storage resources in a timely manner has sparked the research on resource management schemes [20], [21]. However, most of the breakthroughs have focused on the problem of latency minimization during the computational task offloading procedure in which front end and back end must coordinate their actions to support advanced mobile cloud services. In this sense, there are a lack of initiatives in the field of mobile cloud storage. Our potential to explore such initiative is an attempt to narrow this gap in the literature.

In an earlier work, we have addressed the problem of mobile cloud storage for big data transmissions in [10] and [15]. In both works, the goal was to minimize the latency while intelligently exploiting the overlapping nature of a dense 5G wireless networks to expedite the big data transmissions from a mobile device to the MEC. In this sense, the solution proposed in [10] was to divide the big data file into multiple chunks along with two data correction syndromes and to concurrently transmit them over the available data paths. This solution was further improved in [15] by considering that the size of the data chunks are proportional to the capacity of the data link used for transmission and by devising an error correction technique to deal with the eventual errors during transmission.

This paper extends from our previous work by considering a multiple users scenario along with a mechanism design-driven resource management approach. In a multiple users setting, each user has multiple transmissions opportunities provided by the overlapping coverage of a dense 5G heterogeneous network (HetNet) to simultaneous transfer its big data through multiple wireless links. In order to prevent an abusive user from selfishly consuming the wireless resources, a mechanism design framework is proposed to control the allocation of these resources

Manuscript received June 8, 2018; revised December 31, 2018 and February 20, 2019; accepted March 16, 2019. Date of publication April 30, 2019; date of current version November 22, 2019. The work of I. Woungang was supported by the National Science and Engineering Research Council of Canada (NSERC) under Grant RGPIN-2017-04423. (*Corresponding author: Isaac Woungang.*)

R. Siddavaatam, I. Woungang, and G. H. S. Carvalho are with the Department of Computer Science, Ryerson University, Toronto, ON M5B 2K3, Canada (e-mail: richa.siddavaatam@ryerson.ca; iwoungang@scs.ryerson.ca; glauciohscarvalho@gmail.com).

A. Anpalagan is with the Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON M5B 2K3, Canada (e-mail: alagan@ee.ryerson.ca).

Digital Object Identifier 10.1109/JSYST.2019.2908391

TABLE I
DEFINITION OF NOTATIONS

Notation	Definition
P, Q	Error correction syndromes
D_i	i^{th} bit of a data chunk
PD_i	i^{th} bit of a padded data chunk
\hat{P}	Received power
gf	Galois field generator
N	Number of UEs in the network
t_i	Type:Private information declared to mechanism by i^{th}

among the multiple users. A mechanism design is a branch of game theory that merges computer science and economy and leads to a solution that exhibits several desirable features such as incentive compatibility, efficiency, and truthfulness. Overall, the contributions from this paper can be summarized as follows.

- 1) A 5G-driven mobile cloud storage solution for the mobile data deluge is proposed, which exploits the overlapping nature of a dense 5G HetNet to simultaneously transmit the mobile big data from the device to the MEC. In this respect, a solution framework is proposed, which minimizes the transmission delay from the device to the MEC storage center. The proposed algorithm consists of two parts: data management scheme and radio resource management scheme. Based on the redundant array of inexpensive disks (RAID) concepts, a distributed data transmission management scheme with error correction technique is implemented at the individual UE. On the other hand, the radio resource management problem of a HetNet implementing this data management scheme and supporting multiple UEs is formulated as a mechanism design problem, where multiple UEs compete for limited radio resources. The proposed mechanism design based radio resource management scheme is composed of PRB allocation and a pricing algorithm. Numerical results are presented, showing that the proposed technique minimizes the transmission delay in $1 \times 10^2\%$ when compared to a greedy algorithm and the multiple UE implementation version of the DSMU algorithm proposed in [10].
- 2) The PRB allocation problem in the considered network implementing the proposed data transmission management scheme is presented as a graph coloring problem (GCP). Also, the relevant network condition parameters (i.e., the number of UEs and a set of UE pairs with mutual interference lower than the interference threshold) are transformed into an undirected graph. This graph helps in making the radio resource reuse decisions in order to maximize the network performance.
- 3) A pricing algorithm that calculates the price a UE has to pay for the radio resources that it receives from the radio resource management algorithm is proposed. This algorithm is designed to be socially fair and to assign the rewards in such a way that the coalition is successful and valid for a longer duration time.

The rest of this paper is organized as follows. Table I defines the notations used. In Section II, some related work is

presented. In Section III, the considered system model is described. In Section IV, the first part of the proposed framework, i.e., the data management scheme with error correction technique, is described. In Section V, the second part of the proposed framework, i.e., the radio resource management scheme is described. In Section VI, the mathematical analysis of the proposed solution is presented. In Section VII, simulation results are presented. Finally, Section VIII concludes the paper.

II. RELATED WORK

The literature on the MEC network performance primarily focuses on improving the computational task offloading and the radio resource management. Xu *et al.* [23] applied the reinforcement learning paradigm to guide the task offloading decision between the edge computing and the central cloud computing centers, with primary focus on energy conservation. Since both of these tasks are highly co-related, we manage them concurrently as similarly done in [21]. Sardellitti *et al.* [21] considered the joint optimization of radio and computational resources in a network comprising a single MEC server serving multiple UEs using the MIMO antenna technology.

It has been reported in [22] that the computation of a centralized optimal solution for the problem of computational tasks offloading in multiuser multichannel environment is NP-hard. Algorithmic game theory, a fusion of micro-economics game theory and computer science, is being used as an efficient approach to design an algorithm in strategic environment. Jin *et al.* [20] formulated the resource allocation problem as an auction problem while in [22] and [24], the same multiuser computational offloading problem has been formulated as a noncooperative game. In our proposed solution, the mechanism design principles are followed as an engineering approach toward designing the rules of the game or the incentive structure to achieve the desired outcome.

All above-mentioned work have considered only the computational intensive task offloading as main objective. To the best of our knowledge, the problem of storage intensive task offloading has not yet been considered thus far. In this paper, we are proposing a solution in the form of algorithmic framework inclusive of radio resource management, for improving network performance while offloading the storage intensive tasks onto the mobile cloud storage centers. Moreover, our proposed solution is designed in such a way as to exploit the features of the coverage from multiple BSs that are available in the deployed dense HetNet.

III. SYSTEM MODEL

HetNet: The considered 5G network architecture (with numerology = 0) consists of a HetNet composed of multiple femto and macro BSs with overlapping coverage areas, a public cloud system, and several MEC storage centers as shown in Fig. 1. In this model, each BS supports a MEC storage center and a available back-haul connection, which enables each BS to access the public cloud system. It is assumed that there are N mobile devices [user equipments (UEs)], each of which has access to multiple BSs. The network is designed by 5G standards [2], [3] further details of system parameters is given in Section VII-A.

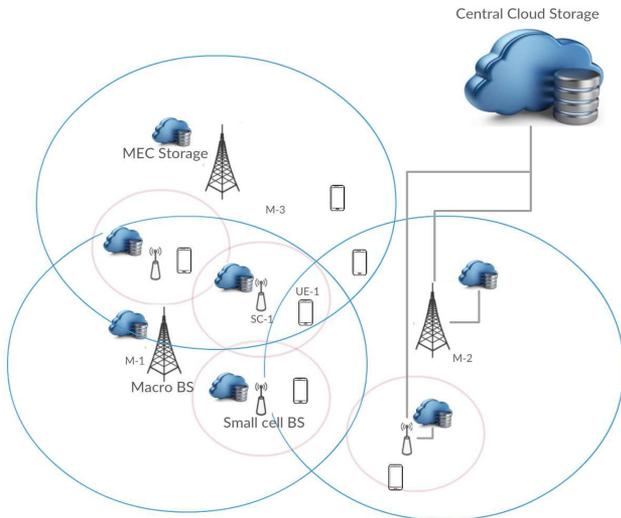


Fig. 1. System Model.

1) *Multiple Data Paths Available to the UEs:* In the proposed HetNet deployment (see Fig. 1), it is likely that small cell BSs will transmit at lower power compared to macro BSs since the signal strength from a macro BS is likely to be much larger than that from a small cell BS. For this reason, a cell range expansion mechanism such as the one proposed in [4]—which exploits the cell selection bias—is considered in this paper to extend the usability of the small cell BSs. In this setting, the BS association decision with a UE depends on the signal to interference and noise ratio (SINR) experienced by the user on the down-link transmission link, knowing that the UE can be associated to multiple BSs.

Typically, a UE will be associated to only those BSs from which it has experienced a SINR higher than a prescribed SINR-threshold value set by the network designer. Moreover, as per the long-term evolution (LTE) standard [4], the SINR threshold γ_{femto} for the association decision with a femto BS is obtained as $\gamma_{\text{femto}} = \gamma_{\text{macro}} - \alpha$, where α is the cell selection bias and γ_{macro} is the SINR threshold for the BS association decision with a macro BS. Therefore, at any instant t , a UE can be associated with multiple BSs on the condition that each of these BSs provides a SINR higher than the respective $\gamma_{\text{femto}}/\gamma_{\text{macro}}$ value.

Because of the fact that a UE can be associated with multiple BSs, multiple paths between the UE and the cloud storage system may be available for data transfer purpose. These paths are also referred to as data links. For example, the eight different data paths available to UE-1 from Fig. 1 are depicted in Fig. 2.

Without loss of generality, one can assume that the network condition does not change during a single storage execution.

2) *LTE Uplink Physical Layer:* The orthogonal frequency division multiple access (OFDMA) is implemented on the LTE physical layer downlink transmission while the single carrier frequency division multiple access is implemented for the uplink transmission. OFDMA allows the data to be directed to/from multiple users on a subcarrier basis for a specified number of symbol periods. In the LTE specifications, these are referred to as PRBs, where a PRB is the smallest element of the resource

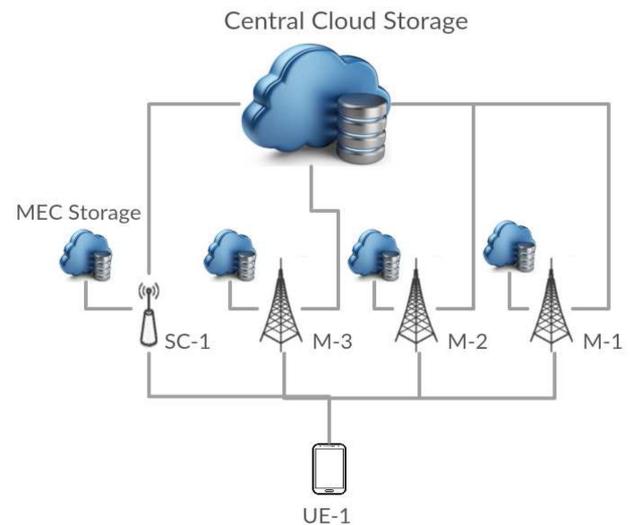


Fig. 2. Data paths available between UE-1 and the cloud system in Fig. 1.

allocation assigned to the UE/BS by the RAN [6]. As per the LTE uplink specifications, the number of PRBs available with the total transmission bandwidth of 20 MHz are 100. These PRBs need to be allocated prudently in order to minimize the average delay experienced by all the active UEs in the network.

IV. PROPOSED DATA MANAGEMENT SCHEME

The procedural details of the complete proposed solution comprising of DMEC scheme and RRM scheme are given in Fig. 3. As stated above, in the considered network, each UE has access to the cloud system (public cloud and MEC storage centers) via multiple data links due to overlapping coverage from the dense deployment of small cell networks. With multiple UEs operational in the network that need to transfer big data files to the cloud system, the data transmission management and radio resource management are complex tasks. As proposed in [10] and [15], these data paths can be exploited to increase the speed of data transfer. A big data file is split into multiple chunks and each of these chunks are transmitted via available data-paths concurrently, hence reducing the transmission delay.

The major constraint on the successful functioning of the transmission technique proposed in [10] is the limited uplink radio resources. As per the LTE-A standards, the uplink spectrum offers 120 PRBs, which needs to be shared among all active UEs. In the considered network scenario, due to the overlapping coverage of multiple BSs, the majority of the active UEs have multiple data paths available to transmit the data chunks simultaneously. The data link is a wireless link between the UE and the BS. However, the number of PRBs available to these data links for the uplink data transmission are finite. Therefore, an intelligent allocation and reuse of the available PRBs to these data paths is necessary in order to achieve the best possible network performance.

In summary, the problem addressed in this paper is to allocate limited PRBs to all active UEs in the network where each UE has a different size of data file to transfer. Moreover, each UE has a different number of data links available and may experienced varied SINR.

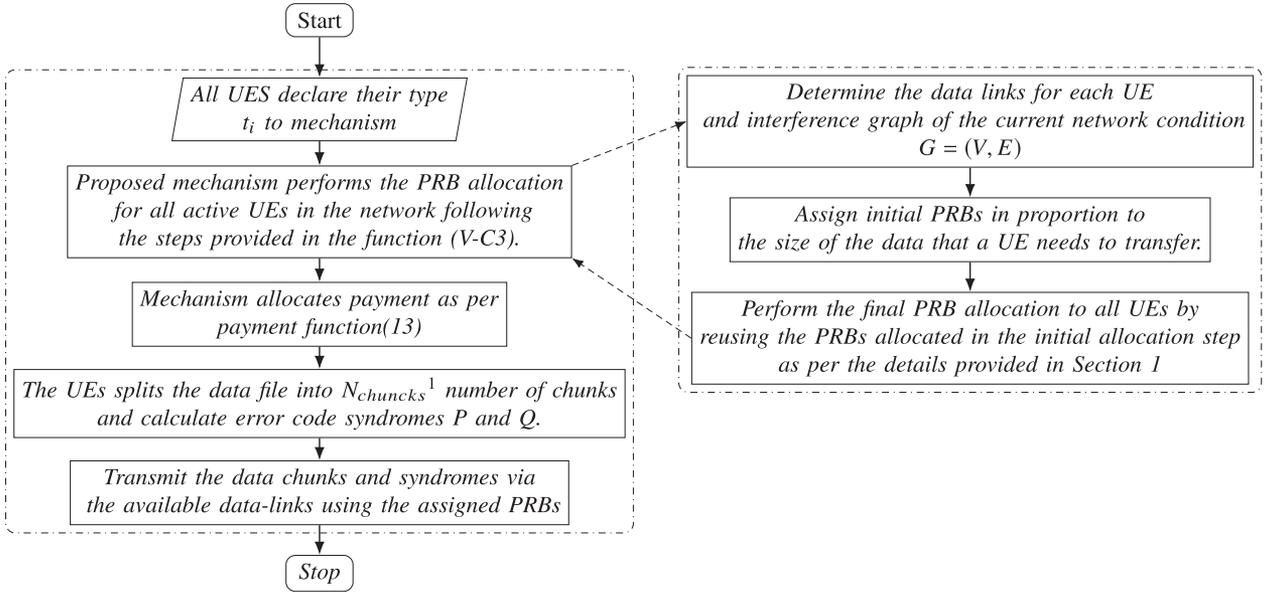


Fig. 3. Procedural details of the proposed solution comprising of DMEC and RRM schemes.

A. Error Correction Technique

In our proposed solution framework, multiple MEC servers are expected to share a distributed storage, therefore fault tolerance should be taken into account [8] since it is likely that any of these devices may fail. Moreover, the data chunks are transmitted over erroneous wireless links, which may corrupt the data while transmission. With high probability of error in data retrieval, the data recovery is critical. This paper adopts the data correction method proposed in [9]. In doing so, it is assumed that the size of the data chunk is proportional to the uplink capacity of the data link on which it is to be transmitted.

For the syndromes P and Q , it is considered that the size of P (resp. Q) is the size of the largest data chunk (resp. the smaller data chunk) are padded with zeros as per (2). These syndromes are using (3), (4) and are saved to the cloud system along with the data chunks of the file. Two of the available data links are reserved for transmission of these syndromes

$$Z = \max \text{size}(D_i) \quad (1)$$

$$PD_i(k) = \begin{cases} D_i(k) & \text{for } k \leq \text{size}(D_i) \\ 0 & \text{for } \text{size}(D_i) < k \leq Z \end{cases} \quad (2)$$

$$P_1 = PD_1 \text{ where } D_i \text{ is } i\text{th data chunk.}$$

$$P_i = P_{i-1} \oplus PD_i, \text{ for } 2 \leq i \leq N_i^{\text{chunks}}$$

$$P = \{P_1, \dots, P_{N_i^{\text{chunks}}}\} \quad (3)$$

$$Q_1 = gf^1 \cdot PD_1$$

$$Q_i = Q_{i-1} \oplus gf^1 \cdot PD_i$$

$$Q = \{Q_1, \dots, Q_{N_i^{\text{chunks}}}\}. \quad (4)$$

In (4), the algebra of Galois field $GF(2^8)$ is used, where gf denotes the generator. Also, a linear feedback shift register is implemented as field generator; and it is assumed that in case a data chunk is loss, the UE to which it pertains will be made aware

of the size of this loss item. Therefore, the zero padded calculated chunk can be truncated in order to recover the original lost data chunk. The proposed method provides protection against up to loss of two data chunks. In case a single data chunk is lost, the syndrome P and the rest of the error-free chunks can be used to recover the lost chunk since P is a X-OR parity. In case two data chunks are lost, the syndromes P and Q and rest of the error-free chunks are to calculate the lost data chunks [9].

V. PROPOSED RADIO RESOURCE MANAGEMENT SCHEME

A. Introduction to Mechanism Design

Mechanism design is a sub-field of the game theory and microeconomics, which has been used to model, analyze, and solve many design problems in distributed environment, where participants can not assumed to follow an algorithm, but their own self interest [18]. The participants who are termed as agents are capable of manipulating the algorithm, hence, the designer of the algorithm should ensure that the interests of the agents are best served by following the designer's expectations. Agents are rational or selfish in the game theoretic sense of making the decisions consistently in pursuit of their own individual goals. It is assumed that each agent is intelligent enough to know everything about the underlying game that a game theorist knows and each agent can make similar inferences about the game that a game theorist can make. Mechanism design can be viewed as reverse engineering of game theory where the rules of a game are designed by the game theorist to achieve a specific desired outcome.

A typical mechanism design optimization problem consists of the following elements [11].

- 1) A game is played among N selfish, intelligent, and strategic agents, and each of these agents has some private information [denoted as type (\hat{t}_i)]. The rest of the game information is a public knowledge.

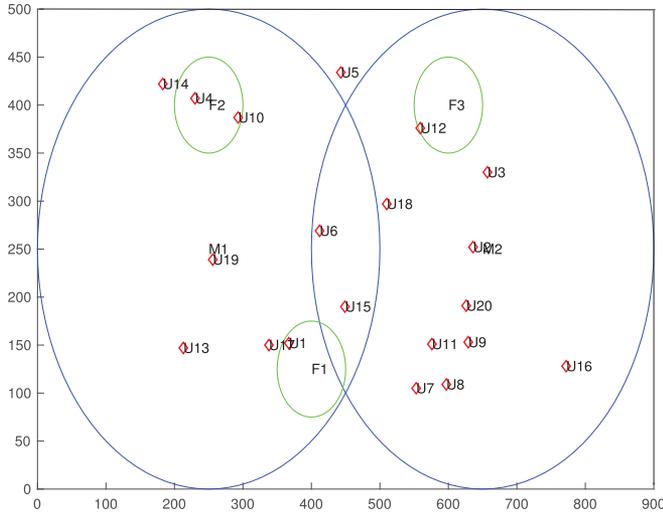


Fig. 4. Simulated network scenario.

- 2) Each agent i declares its type t_i to the mechanism, which may not be same as its true type \hat{t}_i .
- 3) The *output specification* is an objective function $g(O, T)$ that maps the type vector $T = \{t_1, \dots, t_n\}$ to an output vector $O = \{o_1, \dots, o_n\}$. The goal of the mechanism is to find an output vector that minimizes the function $g(O, T)$.
- 4) The *valuation* v_i is a quantification of the i th agent's value from the output o_i .
- 5) Along with the output, a mechanism hands in a payment value p_i to the agent, calculated by using a payment function $p(t_i, O)$. Consequently, the agent's utility is determined as $u_i = v_i(o_i, t_i) - p_i$. The mechanism designer's aim is to design an output vector and payment function, which maximizes the utility of each agent.

B. Radio Resource Allocation Problem

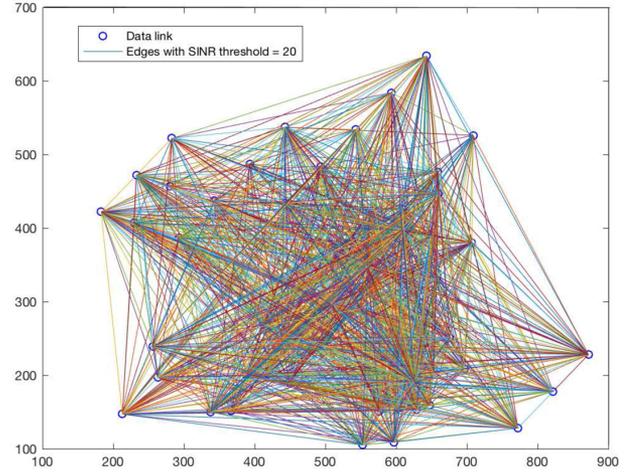
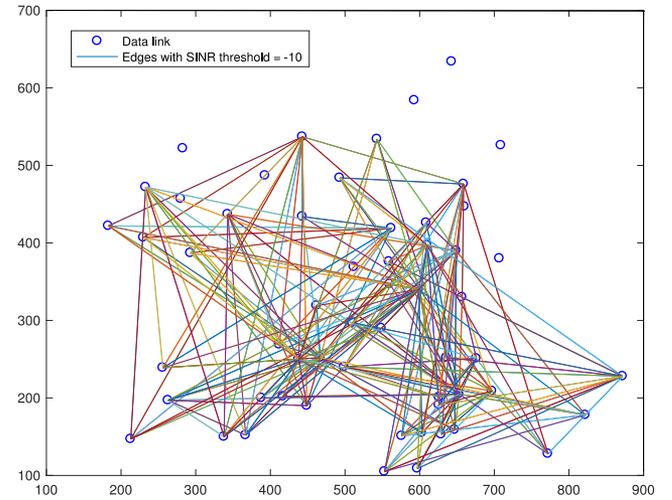
In the considered scenario, the PRB allocation problem can be formulated as a mechanism design problem as follows.

- 1) A game is played among N selfish, intelligent, and strategic agents who are active UEs in the network, and each active UE has a data file that it needs to store in the cloud system; hence, it has a type $\hat{t}_i = D_i \geq 0$, where D_i is the size of the data file.
- 2) Each UE declares a type t_i which might not be equal to its true type \hat{t}_i .
- 3) The output function $g(O, T)$ is defined as

$$g(O, T) = \frac{1}{N} \cdot \sum_{i=1}^N \frac{1}{\lambda_i} \quad (5)$$

$$\lambda_i = \frac{t_i}{\beta_i} \quad (6)$$

where O is the PRB allocation vector assigned by the mechanism, λ_i is the transmission delay experienced by the i th UE while transmitting a data file of size t_i , and β_i is the maximum data rate assigned to the i th UE transmission.

Fig. 5. Interference graph with $\gamma_{\text{thresh}} = 20$ dB.Fig. 6. Interference graph with $\gamma_{\text{thresh}} = -10$ dB.

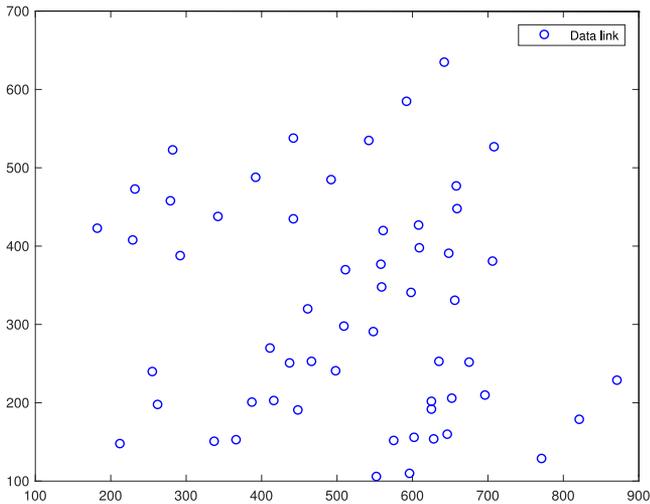
- 4) The valuation $v_i = \frac{1}{\lambda_i} = \frac{t_i}{\beta_i}$ is the reciprocal of the delay experienced by the i th UE. The aim of the mechanism is to discover the O (PRB allocation vector) that minimizes the output function $g(O, T)$.
- 5) In addition to the PRB allocation, the mechanism assigns the payment value p_i to the UE, which represents the price that the i th UE needs to pay to the cloud service provider. This value is directly proportional to the amount of PRBs assigned to it and the total amount of interference that the UE causes to the rest of the UEs in network.
- 6) A part of public knowledge is the interference graph representation of the current network condition (as described in Section V-C1).

C. Proposed Algorithmic Solution to the Radio Resource Allocation Problem

1) *PRB Allocation Problem as GCP*: Once the mechanism have acquire all the required knowledge regarding the network scenario, it transforms the bandwidth allocation problem (namely a PRB allocation problem) in the considered network into a graph coloring problem (GCP). It generates the

Algorithm 1: PRB Allocation Algorithm.

- 1: Determine the data links and generate the interference graph following the details from the Section V-C1
- 2: **procedure** PRB allocation $t_i, G = (V, E)$
- 3: Perform the initial assignment of PRBs to all UEs, where the number of PRBs assigned to a UE is proportional to the size of the data file that it needs to transmit, i.e., $O_i^{\text{initial}} = \{o_1, \dots, o_n\}$ and $O_i^{\text{initial}} \propto t_i$.
- 4: Split equally O_i^{initial} among data-links available to the i^{th} UE.
- 5: **for** all 2-tuple of data-links (k, j) in the network **do**
- 6: **if** $E_{k,j} \neq 1$ (if edge is not present between the i^{th} UE and k^{th} UE) **then**, reuse the initial PRBs
- 7: $O_j = O_k^{\text{initial}} \cup O_j^{\text{initial}}$
- 8: $O_j = O_k^{\text{initial}} \cup O_j^{\text{initial}}$ where $E_{k,j}$ is equal to 1 if an edge exists between the k^{th} UE and the j^{th} UE and 0 otherwise.

Fig. 7. Interference graph with $\gamma_{\text{thresh}} = -40$ dB.

interference that would be used by the proposed allocation Algorithm 1. Following the technique proposed in [5], the above-mentioned PRB allocation problem can be represented as a k -graph coloring problem, where k is the number of available PRBs. In the graphs depicted in Figs. 5–7, $G = (V, E)$, where V is the set of vertices and E is the set of edges. The k -coloring of G is a function

$$c: V \rightarrow \{1, 2, \dots, k\}$$

such that $c(i) \neq c(j)$, where E contains an edge $[v_i, v_j]$ connecting the vertices $v(i)$ and $v(j)$. In order to formulate the PRB allocation problem as a graph coloring problem, UE-BS are represented as vertices and a PRB is represented as a color. Here onwards, the terms data link and vertex will be used interchangeably as well the terms color and PRB. In addition, the minimum SINR performance expected at the user terminals is set to γ_{thresh} . The corresponding interference threshold (I_{thresh}) can be calculated as $I_{\text{thresh}} = \min(\hat{p}_1, \hat{p}_2) - \gamma_{\text{thresh}}$ [in decibel], where \hat{p}_1 and \hat{p}_2 are the powers received by data link1 and data link2, respectively. The presence of an edge is restricted to

those users whose mutual interference is higher than I_{thresh} . The presence of an edge connecting two users signifies that the users, if allocated the same PRB, will cause a mutual interference that is higher than I_{thresh} . The network scenario depicted in Fig. 4 are represented by the graphs shown in Figs. 5–7, where γ_{thresh} is set to the values 20 dB, -10 dB, and -40 dB, respectively.

The PRB allocation problem is transformed into a graph coloring problem. For a particular instance of the network condition, the number and position of the vertices in the graph remain constant, but the number of edges varies with the value of the SINR threshold (γ_{thresh}). When γ_{thresh} increases, it is clear that I_{thresh} decreases since $I_{\text{thresh}} = \min(\hat{p}_1, \hat{p}_2) - \gamma_{\text{thresh}}$ [in decibel], where \hat{p}_1 is the power received by user 1 and \hat{p}_2 is the power received by user 2. An edge is present between two vertices if the mutual interference is higher than I_{thresh} ; hence, the number of edges increases with an increase in γ_{thresh} as inferred to by Figs. 5–7. The higher the number of edges, the higher the number of constraints on the PRB allocation, hence, the complexity of PRB allocation problem increases with an increase in the value of the SINR threshold.

2) *Proposed PRB Allocation and Pricing Algorithm:* Mechanism design is a methodology of designing the rules of a game to achieve the desired goal of the game. As a matter of fact, the Vickrey–Clarke–Groves (VCG) mechanism has been widely used in multiagent systems to devise an efficient resource allocation when the agents have private independent valuations [12]. This is because it exhibits several desirable properties, including incentive and compatibility. It also ensures that the resulting resource allocation is efficient (i.e., the social welfare is maximized) and individually rational (i.e., it ensures that all the participating players derives a nonnegative utility from the mechanism) [16]. For this reason, we have considered a PRB allocation based on the VCG mechanism.

Here, the VCG mechanism is applied to the utilitarian mechanism design problems.

Definition 1: A maximization mechanism design problem is called utilitarian if its objective function is the sum of all agents' valuations, i.e., $g(O, T) = \sum_{i=1}^n v_i(t_i, O)$ [11].

Lemma 1: The maximization mechanism design problem under consideration is utilitarian.

Proof: As stated earlier in Section V-B, the agent's valuation (v_i) in the considered scenario is the reciprocal of transmission delay experienced by the i th UE (λ_i). As per (5), the objective function of the considered mechanism design problem is the average value of the valuations of all active UEs in the network, which is proportional to the sum of the valuations of the individual UEs. This proves that the problem under consideration is utilitarian; hence, the VCG mechanism can be applied to the considered mechanism design problem.

For any mechanism $m(O(t), h(O, t))$ to be classified as a VCG-based family of mechanisms, it should exhibit the following properties [11].

- 1) The mechanism must be utilitarian

$$O(t_i) \in \arg \max_O \left\{ \sum_{i=1}^n v_i(\cdot) \right\}. \quad (7)$$

- 2) It must define the payment functions for the agents as per the following equation:

$$p(t_i) \sum_{j \neq i} v^j(t^j, O(t_i)) + h(t^{-1})$$

where $h(t^{-1})$ is an arbitrary function of t^{-1} . (8)

The intuition behind the payment function of a VGC mechanism is that a player pays for its “social cost”, i.e., the damage it imposes on the society. The first term of (8) corresponds to the optimal welfare for other players if the i th player is not participating in the game, and the second term of (8) corresponds to the welfare of other players with the current output O . Our proposed mechanism complies to these requirements.

3) *Output Function*: The goal of our proposed framework is to minimize the average delay experienced by all active UEs in the network. Once all UEs declare their types to the mechanism, the optimization problem can be defined [as per (10)] and the output (i.e., the PRB allocation) is determined. The optimization problem can be described as follows:

$$O = \arg \max_{\tilde{O}} \{v_i\} \quad (9)$$

$$O = \arg \max_{\tilde{O}} \left\{ \frac{1}{N} \sum_{i=1}^N \frac{t_i}{\beta_i(\tilde{O})} \right\} \quad (10)$$

$$\beta_i = W_i \times \log_2 \left(\frac{\hat{P}_i^J}{\tau_i \times [1 + (\eta \div 100)]} \right) \quad (11)$$

$$\tau_i = \left(\sum_{j=1, j \in o_i, j \neq i}^{N_{BS}^{\text{total}}} \hat{P}_i^j \right) \quad (12)$$

where \tilde{O} is a set of sets representing the PRB allocation to the requesting UEs. $O = \{o_1, \dots, o_N\}$, o_i is a set of PRBs assigned to the i th UE, N_{BS}^{total} is the number of BSs in the network, W_i is the amount of bandwidth available to the i th UE, and \hat{P}_i^j is the power received by the i th UE from the j th BS.

4) *Payment Function*: Once the PRB allocation has been formulated, the payment, i.e., the price a UE needs to pay for the PRBs for the data transfer purpose is calculated by the payment function $p(t_i)$ (13). Any mechanism can be stated to be efficient if it has an efficient allocation rule/output function. The VCG mechanism is an efficient mechanism, which uses the payment function $p : T \Rightarrow \mathbb{R}$. As stated earlier, the payment function p_i is the difference between the optimal welfare for other players if the i th player is not participating and the welfare of other players with the current type T . An intuition behind this payment function is that a player should pay for the damage that it imposes on the society [14].

Following the considered VCG mechanism in the considered network scenario, an efficient payment function would be $p(t_i) \propto \bar{\lambda}_j^{-1} - \bar{\lambda}_j$, where $\bar{\lambda}_j^{-1}$ is the average delay experienced by other UEs if the i th UE disappears from the network; and $\bar{\lambda}_j$ is the average delay experienced by all UEs with all active UEs present with the same declared type T . The difference between these two terms is dependent upon two factors namely,

- 1) Additional PRBs rest of the UEs would be received if the i th UE disappears from the network.
- 2) There would be reduction in the interference received by other UEs if the UE disappears from the network.

Considering these factors, the payment function of our proposed method is obtained as

$$p(t_i) = 2^{\|W_i\|} + pw_1 \cdot \|(\psi_i)\| + pw_2 - 1 \quad (13)$$

where the effect of the first factor is considered in the first term of (13), where the payment value is directly proportional to the normalized amount of PRBs assigned to the UE ($\|W_i\|$). The effect of the second factor is considered in the second term of (13) where ψ_i is the total interference caused by the i th UE to the rest of the active UEs in the network as a result of the same output function. The constants pw_1 and pw_2 are the weights of the payment function, which are decided by the MNOs as per their billing requirements.

VI. ANALYSIS OF OUR PROPOSED METHOD

Our proposed method inherits the desirable economic properties of the VGC mechanism as follows.

A. Incentive Compatible

Definition 2: A mechanism said to be incentive compatible in *ex post* Nash equilibrium if the best response strategy for all players is to reveal its true type even if it has the complete information about other agent's types [14] and [16].

Definition 3: According to [16], the utility of the i th UE is defined as

$$u_i = v_i - h(t_i) \quad (14)$$

$$v_i = \frac{W_i}{t_i} \cdot \log_2 \left(\frac{\hat{P}}{\tau_i (1 + 0.01 \cdot \eta)} \right). \quad (15)$$

The valuation a player/UE gets from the game is proportional to the delay experienced by a UE when transferring its big data file. In (14), if we ignore the terms which are independent of the declared type (t_i), the utility of a UE is obtained as

$$u_i = \frac{W_i}{t_i} \cdot \log_2 \left(\frac{\hat{P}}{\tau_i (1 + 0.01 \cdot \eta)} \right) - 2^{\|W_i\|} - \|(\psi_i)\|. \quad (16)$$

Lemma 2: The proposed mechanism composed of an output and a payment function is incentive compatible in *ex post* Nash equilibrium.

Proof: Every VCG mechanism is incentive compatible [13]. With the close examination of (16), the interference caused to another active UEs (ψ_i) is proportional to the number of interfering UEs, which are using at least one PRB similar to the PRBs assigned to the i th UE. The number of UEs reusing the PRBs is proportional to the number of edges connected to a vertex in the network graph representing the UE. The interference graph is independent of the type (t_i) declared by the UEs. Similarly τ_i is independent of t_i , hence the utility of the UE reduces to the term, which depends on the UE's declared type

$$u_i = \frac{W_i}{t_i} - 2^{\|W_i\|}. \quad (17)$$

Let \hat{t}_i be the type, i.e., the data file size declared by the i th UE. If $\hat{t}_i > t_i$, i.e., if a UE has declared a data file size larger than the real data file size, more PRBs would be assigned to the UE. This would linearly increase the first term of (17) while exponentially increasing the payment [second term of (17)], hence the net utility will be reduced. On the other hand, if $\hat{t}_i < t_i$, i.e., a UE has declared a data file size lower than the real data file size, then less number of PRBs would be assigned to that UE, which is undesirable by that UE.

B. Efficiency

Definition 4: An allocation rule (or output function) of a mechanism is said to be efficient if it maximizes the social welfare [12]

$$O \in \arg \min_{\hat{O}} \forall_i \lambda_i. \quad (18)$$

Lemma 3: The proposed mechanism is efficient since it allocates the PRBs with the aim to maximize the average transmission delay experienced by each active UE in the network as per (10).

Proof: In the considered problem, the social welfare is equivalent to the inverse of the delay experienced. As the proposed mechanism is incentive compatible *ex post* Nash equilibrium, it encourages the players (UEs) to declare their true types. Once each UE in the game has declared the true size of the data file it needs to transfer, the output function decides on the PRB allocation as well as the payment values for the UEs, which in turn maximizes the average value of the data rates experienced by all of active UEs in the network.

C. Individually Rational

Definition 5: Each player in the game is assumed to be “intelligent”. The mechanism should encourage a selfish agent to join it rather than opt out of it by ensuring that the utility derived by an agent within the mechanism is never less than that of the agents that are exterior to it. It is assumed that the utility that an agent gains from not joining the mechanism is 0. One needs to prove that the utility an UE gains in the mechanism is always ≥ 0 [12].

Lemma 4: The proposed mechanism is individually rational. The output/PRB allocation and payment functions have been designed in such a way that the mechanism encourages a selfish UE and participating MNOs to participate in the coalition.

Proof: The proposed mechanism ensures that even with the worst network conditions, each participating UE receives at least one PRB, which is equivalent to $W_i = 180$ MHz. With this minimum possible amount of bandwidth, the proposed mechanism ensures that a UE can achieve a minimum possible utility [as per (16)], which is always higher than the utility that a UE would achieve by not participating in the game.

VII. SIMULATION RESULTS

A. Simulation Settings

For evaluating the performance of our proposed algorithmic framework, we have executed it on a simulated HetNet

following the LTE-A standards. Since our focus is on the faster transfer of a big data file to the cloud storage system, we have simulated the uplink data transmission from UE to the cloud computing center. Moreover, we have simulated the downlink transmission power since the cell association decision is based on the SNIR experienced by the UE from multiple BSs. In order to simulate the power and interference received by the UE and the macro/femto BS, we have considered the path loss model proposed in [5]. For the downlink channel simulation, the power received by a UE is calculated as per (19) and (20). On the other hand, for the uplink channel simulation, the power received by a BS is calculated as per (21) and (22). The path loss PL is calculated by using (23)–(26), depending on the location of the UE and the type of BS

$$\hat{P} = P_{tx}^m - \text{PL (dB)} \quad (\text{for macro base station}) \quad (19)$$

$$\hat{P} = P_{tx}^f - \text{PL (dB)} \quad (\text{for femto base station}) \quad (20)$$

$$\hat{P} = P_{tx}^{UE} - \text{PL (dB)} \quad (\text{for macro base station}) \quad (21)$$

$$\hat{P} = P_{tx}^{UE} - \text{PL (dB)} \quad (\text{for femto base station}) \quad (22)$$

where P_{tx}^f is the transmit power of the femto BS, P_{tx}^m is the transmit power of the macro BS, P_{tx}^{UE} is the transmit power of the UE, \hat{P} is the received power by a UE in case of downlink transmission, and \hat{P} is the received power by a BS in case of uplink transmission.

Besides, the path loss $\text{PL}_f^{\text{indoor}}$ (in dB) between the indoor users and the femtocell BS is obtained as

$$\text{PL}_f^{\text{indoor}} = 37 + \alpha_{in} \times 10 \times \log_{10}(d) + 18.3d_f^{d_f^2 + 3d_f + 1.54}. \quad (23)$$

The path loss $\text{PL}_m^{\text{indoor}}$ between the indoor users and the macrocell BS is obtained as

$$\text{PL}_m^{\text{indoor}} = 40 + \alpha_{out} \times 10 \times \log_{10}(d) + pl_{bp}. \quad (24)$$

The path loss $\text{PL}_m^{\text{outdoor}}$ between the outdoor users and macrocell BS is obtained as

$$\text{PL}_m^{\text{outdoor}} = 40 + \alpha_{out} \times 10 \times \log_{10}(d). \quad (25)$$

The path loss $\text{PL}_f^{\text{outdoor}}$ between outdoor users and femto BS is obtained as

$$\text{PL}_f^{\text{outdoor}} = 37 + \alpha_{out} \times 10 \times \log_{10}(d) + pl_{bp} \quad (26)$$

where d is the distance between the UE and the BS and the other parameters are given in Table II whose values are determined following 5G network standards [2], [3]. The maximum possible data rate is calculated according to (11) following the Shannon–Hartley theorem.

The proposed algorithmic framework is executed on the simulated network model provided by MATLAB (see Section VII-A). A typical simulated network scenario is shown in Fig. 4. The goal of our framework is to efficiently transfer the big data files from mobile UEs to the cloud system with minimum data transmission delay. For comparison purpose, a greedy algorithm and along with a data split multiple UE (DSMU) algorithm are designed, implemented, and executed on the same

TABLE II
SYSTEM PARAMETERS

Abbreviation	Parameter Description	Value.
P_{tx_m}	Transmit power of macrocell BS	45 dBm
P_{tx_f}	Transmit power of femtocell BS	20 dBm
η	Noise experienced by a UE (% of total interference experienced by UE)	10%
N_{PRB}	Number of PRBs available	100
B_{ch}	Channel bandwidth	20 MHz
N_{SC}^{total}	Bandwidth per PRB	180 KHz
N_{UE}^{in}	Number of indoor users	0
N_{UE}^{out}	Number of outdoor users	20
N_{UE}^{total}	Number of total users	20
N_m	Number of macrocells	2
N_{FAPS}	Number of femtocells	3
d_f	Indoor penetration factor	1
α_{out}	Outdoor path loss exponent	3.5
α_{in}	Indoor path loss exponent	2.5
pl_{pp}	Building penetration loss	8
γ_{thresh}^{macro}	SINR threshold for cell association for macro BS	28 dBm
γ_{thresh}^{femto}	SINR threshold for cell association for femto BS	-13 dBm
η_m	Macro BS Back-haul connection speed	500 Mbps
η_f	Femto BS Back-haul connection speed	100 Mbps

network model. The performance of our proposed algorithm is compared against that of the greedy and DSMU algorithms.

B. Greedy Algorithm

In the implementation of the greedy algorithm (GA), each UE associates itself to the BS with the maximum downlink SINR and the data file is transmitted to the public cloud via this BS. The uplink PRB allocation in the GA algorithm is performed as follows. There are two macro BSs present in the simulated network, namely M1 and M2. The total available PRBs are equally assigned to all the UEs associated with either macro BS without any reuse. The available PRBs are again reused for assignment to the UEs associated to the femto BSs. On the other hand, the total PRBs are equally split among the UEs associated with any femto BS.

C. Data Split Multiple UE Algorithm

The DSMU algorithm is an enhanced version of a recently proposed algorithm [10], in which multiple active concurrent UEs have been considered. Typically, in this enhanced version, each UE has multiple data links (i.e., wireless links between the UE and its associated BS) associated with it. All data links from multiple UEs that are associated with macro BS M1 would be assigned the first half of the available PRBs and the rest of the PRBs would be assigned to the data links associated with the macro BS M2. The whole set of PRBs is reused for assignment to the data links, which are associated with the femto BSs. The available PRBs are equally assigned to all data links, which are associated with any femto BS.

First, it is assumed that all active UEs in the network have a data file of equal size to transmit to the cloud. The data file size is varied from 1 to 10 Gbits and the impact of this variation on the average delay experienced by all UEs is investigated. The results are captured in Fig. 8. It is observed in Fig. 8 that the time required by our algorithm for complete file transfer is

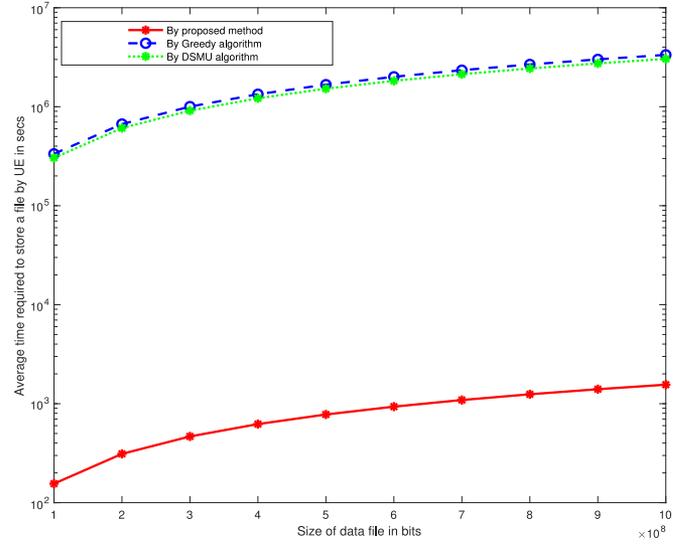


Fig. 8. Time required by the UE to complete the transfer of the file versus size of data file (on a logarithmic scale).

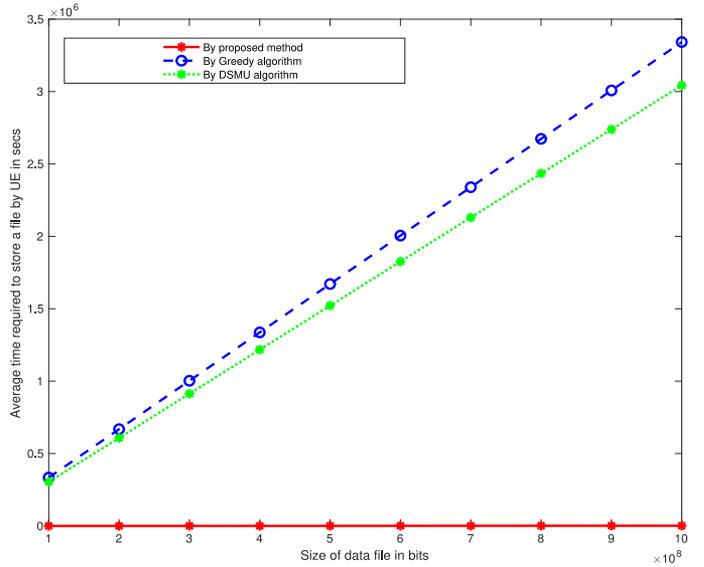


Fig. 9. Time required by the UE to complete the transfer of the file versus size of data file.

significantly lower compared to that generated by the greedy and DSMU algorithms.

Indeed, to transfer a data file of size 9 Gbits, the greedy and DSMU algorithms require approximately 10^8 s while the proposed algorithm transfers the file in 10^3 s. Additionally, in Fig. 9, it can be observed that as the size of the data file increases, the time required to transfer it increases linearly in all cases. However, the proposed algorithm is less sensitive to variation in the data file size, which is evident from the fact that the slope of the greedy algorithm is $3.5e^{-2}$, that of the DSMU algorithm is approximately $2.8e^{-2}$, and that of the proposed algorithm is $6.3e^{-6}$.

Second, the number of active UEs in the network is varied and the impact of this variation on the average delay experienced by all UEs is investigated. The results are captured in Fig. 10. It can be observed that for the proposed scheme and DSMU

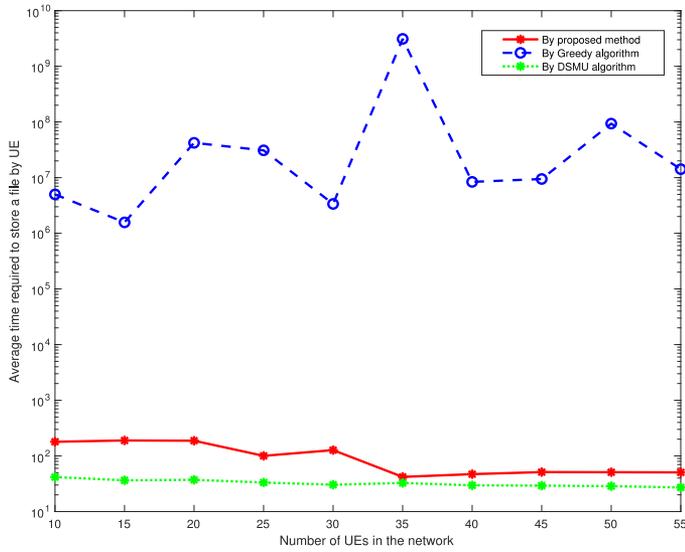


Fig. 10. Time required by the UE to complete the transfer of the file versus number of active UEs in the network (on a logarithmic scale).

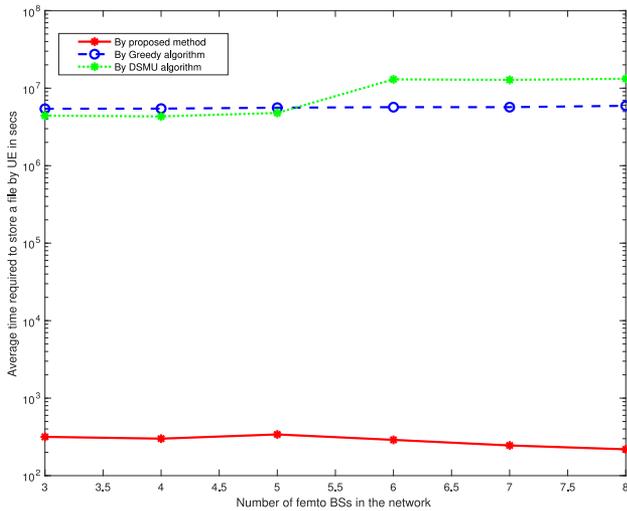


Fig. 11. Time required by the UE to complete the transfer of the file versus number of small cell BSs in the network.

algorithm, the average delay experienced by a UE is approximately 10^4 times less than that experienced by the greedy algorithm. It can be further observed that the proposed and DSMU algorithms are less sensitive to changes in the number of active UEs while the performance of the greedy algorithm fluctuates with the number and locations of the active UEs in the network. The rationale behind this stable performance of the proposed algorithm with the variations in the number and locations of the UEs is that the proposed algorithm utilizes multiple BSs covered by a UE and the PRB allocation is performed considering this data transmission strategy. Moreover, the PRB allocation algorithm also considers the network conditions, including the locations and number of active UEs in the network.

Third, the number of small cell BSs in the network is increased and the impact of this increase on the average delay experienced by all UEs is investigated. The results are captured in Fig. 11. It can be observed that the data transmission delay generated

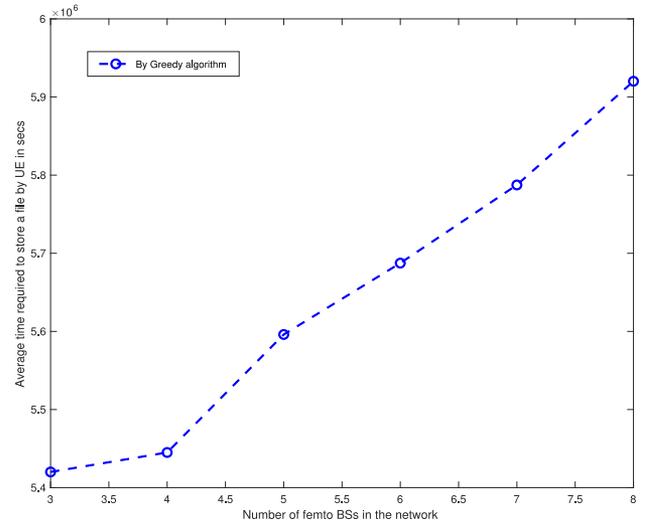


Fig. 12. Time required by the UE to complete the transfer of the file versus number of femto small cell BSs in the network.

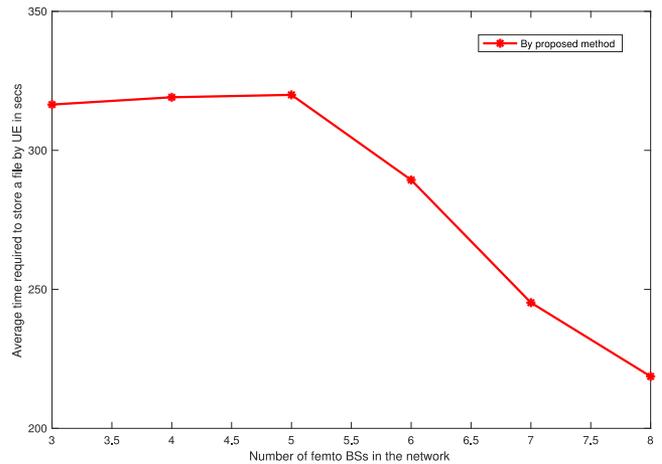


Fig. 13. Time required by the UE to complete the transfer of the file versus number of femto small cell BSs in the network (on a logarithmic scale).

using our proposed algorithm is about 10^5 times lower than that generated by the greedy and DSMU algorithms.

In Fig. 12, which is a magnified view of Fig. 11, it can be observed that as the HetNet becomes more denser, the delay experienced by the greedy algorithm increases, meaning that its performance deteriorates. On the other hand, with our proposed algorithm as depicted in Fig. 13, the average delay experienced by a UE decreases. This is due to the fact that our proposed algorithmic framework exploits the intrinsic HetNet infrastructure and continues to improve the network performance even when the number of small cell BSs is increased in the network.

Fourth, the heatmap from Fig. 14 depicts the values of the proposed payment function under varying normalized number of PRBs allocated to a UE among all active UEs in the network and varying values of the normalized interference that the same UE causes to the rest of the UEs in the network. It can be observed that the payment value is directly proportional to the amount of PRBs allocated to a UE and the amount of interference a UE causes to the rest of the active UEs in the network.

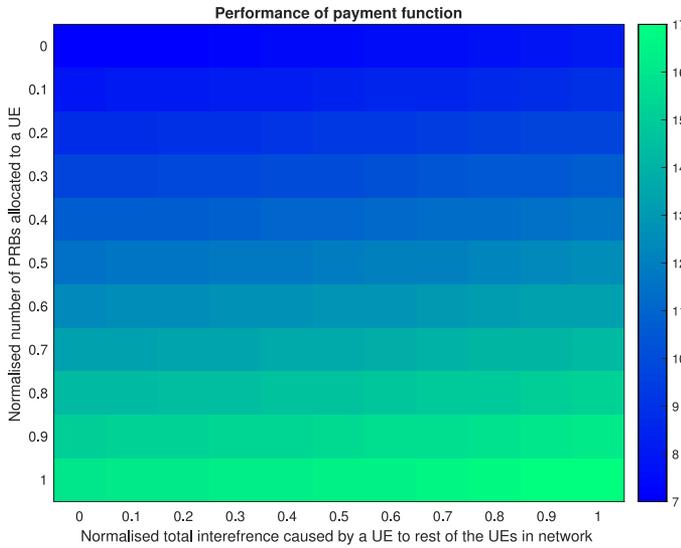


Fig. 14. Payment function distribution.

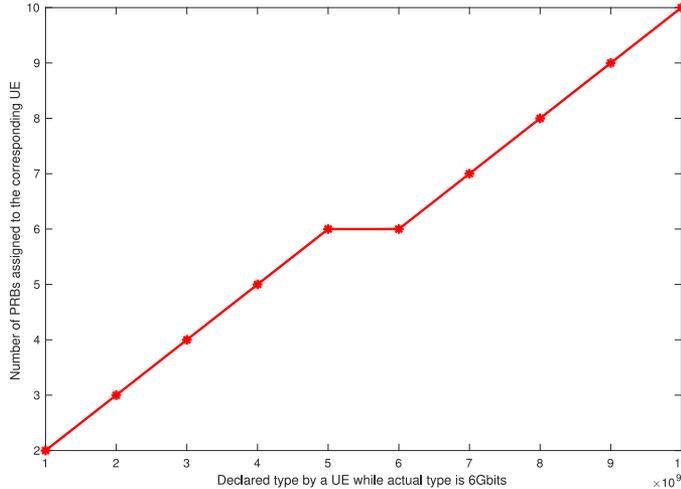


Fig. 15. Initial number of PRBs assigned to a UE by the proposed algorithm for various declared type (t_i).

D. Incentive Compatible *ex post* Nash Equilibrium

To validate that our proposed algorithmic framework is incentive compatible *ex post* Nash equilibrium, we have performed some simulations as follows. In the proposed framework, the type t_i is the size of data file that a UE needs to transfer to the cloud system. It is assumed that all UEs have the same data files of size 6 Gbits, and one of the active UEs has declared a data file size/type varying from 1 to 10 Gbits, but different from 6 Gbits (i.e., it is not the true type value). All other simulation parameters are inherited from Table II.

First, the declared type values is varied and the impact of this variation on the initial number of PRBs assigned to the UE under consideration without the application of frequency reuse is investigated. The results are captured in Fig. 15. It is observed that as the declared type by the UE increases, the number of PRBs assigned to the UE initially by the algorithm increases linearly.

Second, the final number of PRBs assigned to a UE by the proposed algorithm when applying the frequency reuse tech-

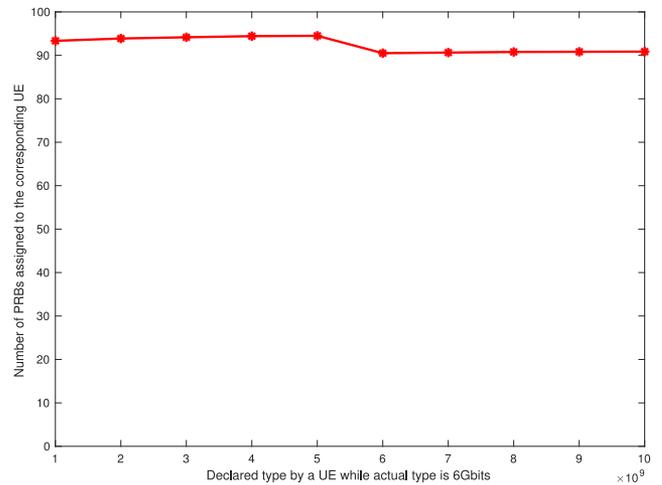


Fig. 16. Final number of PRBs assigned to a UE by the proposed algorithm for various declared type (t_i) using frequency reuse.

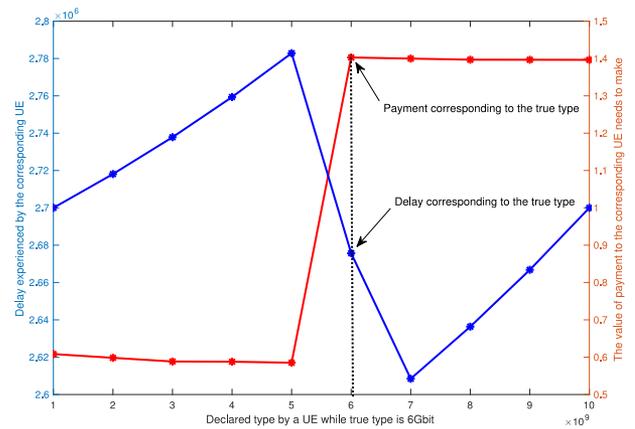


Fig. 17. Delay experienced by the UE and the payment value a UE needs to pay for various values declared type (t_i) declared.

nique is shown in Fig. 16. It can be observed that the initially assigned PRBs are reassigned to the UEs who do not share an edge with the current UE from the corresponding GCP. When the declared type value is increased, the number of PRBs assigned to a UE slightly decreases. This is due to the fact that the frequency reuse factor is dependent on the interference graph, thereby on the network conditioning parameters such as locations of the UEs and base stations; and it is independent of the declared type values. The PRBs assigned to a UE by frequency reuse constitutes the majority of PRBs assigned to the UE. This is the rationale behind utilizing the amount of initial PRBs without frequency reuse (instead of the number of final PRBs) in the payment function.

Third, the number of PRBs assigned to a UE is directly proportional to the throughput that a UE can achieve. Indeed, the delay experienced by the UE under consideration when transferring its data file and the proportional payment value that it needs to pay for various declared types values, are depicted in Fig. 17. It can be observed that when the declared type value is increased, the delay experienced by the UE slightly decreases until it reaches its true value, then it suddenly increases, and then remains constant when the declared type values further in-

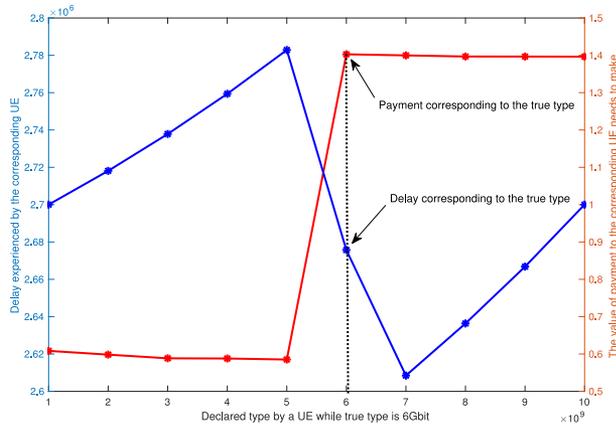


Fig. 18. Net utility of a UE for various values of type (t_i) declared.

creases. On the other hand, the payment value increases when the declared type value increases and suddenly dips when the declared type is closer to its true value.

Fourth, the Net utility of the UE into consideration under varying declared type (t_i) values is investigated, where the Net utility of the UE is a function of the delay experienced by the UE in transferring the data file and the amount interference caused by the UE to the rest of the UEs in network. The results are captured in Fig. 18. It can be observed that the utility of a UE is maximum (with experimental errors) when the declared data file size/type is closer to the true type size (6 Gbits). Therefore, the proposed mechanism is successful in encouraging a UE to declare its true type, i.e., it is incentive compatible *ex post* Nash equilibrium.

VIII. CONCLUSION

We have proposed a mechanism design-based algorithmic framework to transfer big data files with minimum transmission delay in a network composed of HetNet, public cloud systems, and MEC storage centers. Our proposed solution consists of a data transmission management scheme, a PRB/bandwidth allocation function, and a pricing function using the principle of algorithmic mechanism design. The proposed RRM scheme is shown to provide several desirable characteristics such as incentive compatibility, efficiency, truthfulness, all derived from the VCG mechanism. Simulations results have shown that compared to baseline techniques, the proposed framework minimizes the transmission delay by about $10^2\%$. Moreover, the performance of proposed algorithm is shown to be stable, irrespective of the increase in the number of active UEs in the network. It also yields promising results in terms of the delay experienced by the UE for completing its data transfer in the cloud in the presence of an increased number of small cell BSs in the network. As future work, one can investigate the design of a revenue sharing scheme to complement the proposed framework, in which multiple MNOs operating various macro and small cell BSs would be sharing the revenue generated from their coalition with fairness so that the coalition would function successfully for a longer duration.

REFERENCES

- [1] "Cisco global cloud index: Forecast and methodology," 2016–2021 White Paper, Feb. 1, 2018.
- [2] LTE-A network standards, May 12, 2016. [Online]. Available: <http://www.3gpp.org/technologies/keywords-acronyms/97-lte-advanced>
- [3] S. Y. Lien, S. L. Shieh, Y. Huang, B. Su, Y. L. Hsu, and H. Y. Wei, "5G New radio: Waveform, frame structure, multiple access, and initial access," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 64–71, Jun. 2017.
- [4] S. Deb, P. Monogioudis, J. Miernik, and J. P. Seymour, "Algorithms for enhanced inter-cell interference coordination (eCIC) in LTE HetNets," *IEEE ACM Trans. Netw.*, vol. 22, no. 1, pp. 137–150, Mar. 2013.
- [5] R. Siddavaatam, A. Anpalagan, I. Woungang, and S. Misra, "Ant colony optimization based sub-channel allocation algorithm for small cell HetNets," *Wireless Pers. Commun.*, vol. 77, no. 1, pp. 411–432, 2014.
- [6] J. Zyren and W. McCoy, "Overview of the 3GPP long term evolution physical layer," Freescale Semiconductor, Inc. White paper, Tech. Rep. 3GPPEVOLUTIONWP, Rev 0, Jul. 2007.
- [7] C. U. Castellanos *et al.*, "Performance of uplink fractional power control in UTRAN LTE," in *Proc. IEEE Veh. Tech. Conf.*, Singapore, 2008, pp. 2517–2521.
- [8] J. S. Plank, "A tutorial on Reed-Solomon coding for fault-tolerance in RAID-like systems," *Softw., Pract. Exper.* vol. 27, no. 9, pp. 995–1012, 1997.
- [9] H. P. Anvin, "The mathematics of RAID-6," 2007. [Online]. Available: <http://kernel.org/pub/linux/kernel/people/hpa/raid6.pdf>
- [10] R. Siddavaatam, I. Woungang, G. Carvalho, and A. Anpalagan, "Efficient ubiquitous big data storage strategy for mobile cloud computing over HetNet," in *Proc. IEEE Global Commun. Conf.*, Washington, DC, USA, 2016, pp. 1–6.
- [11] N. Nisan and A. Ronen, "Algorithmic mechanism design," in *Proc. 31st Annu. ACM Symp. Theory Comput.*, 1999, pp. 129–140.
- [12] R. K. Dash, A. Rogers, N. R. Jennings, S. Reece, and S. Roberts, "Constrained bandwidth allocation in multi-sensor information fusion: A mechanism design approach," in *Proc. IEEE 8th Int. Conf. Inf. Fusion*, 2005, vol. 2, pp. 1185–1192.
- [13] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, Eds., *Algorithmic Game Theory (Vol. 1)*. Cambridge, U.K.: Cambridge Univ. Press, 2007.
- [14] V. Krishna, *Auction Theory*. New York, NY, USA: Academic, 2009.
- [15] R. Siddavaatam, I. Woungang, G. Carvalho, and A. Anpalagan, "An efficient method for mobile big data transfer over HetNet in emerging 5G systems," in *Proc. IEEE 21st Int. Workshop Comput. Aided Model. Des. Commun. Links Netw.*, Toronto, ON, USA, 2016, pp. 59–64.
- [16] R. K. Dash, A. Rogers, and N. R. Jennings, "A mechanism for multiple goods and interdependent valuations," in *Proc. Int. Workshop Agent-Mediated Electron. Commerce*, Springer, Berlin, Germany, Jul. 2004, pp. 15–29.
- [17] V. Krishna, *Auction theory, (ser. Mechanism Design)*. New York, NY, USA: Academic, 2009, pp. 61–79.
- [18] N. Nisan and A. Ronen, "Algorithmic mechanism design," *Games Econ. Behav.*, vol. 35, no. 1/2, pp. 166–196, 2001.
- [19] Y. Liu, M. J. Lee, and Y. Zheng, "Adaptive multi-resource allocation for cloudlet-based mobile cloud computing system," *IEEE Trans. Mobile Comput.*, vol. 15, no. 10, pp. 2398–2410, Oct. 2016.
- [20] A. L. Jin, W. Song, and W. Zhuang, "Auction-based resource allocation for sharing cloudlets in mobile cloud computing," *IEEE Trans. Emerg. Topics Comput.*, vol. 6, no. 1, pp. 45–57, Jan.–Mar. 2018.
- [21] S. Sardellitti, G. Scutari, and S. Barbarossa, "Joint optimization of radio and computational resources for multicell mobile-edge computing," *IEEE Trans. Signal Inf. Process. Over Netw.*, vol. 1, no. 2, pp. 89–103, Jun. 2015.
- [22] X. Chen, L. Jiao, W. Li, and X. Fu, "Efficient multi-user computation offloading for mobile-edge cloud computing," *IEEE/ACM Trans. Netw.*, vol. 24, no. 5, pp. 2795–2808, Oct. 2016.
- [23] J. Xu, L. Chen, and S. Ren, "Online learning for offloading and autoscaling in energy harvesting mobile edge computing," *IEEE Trans. Cogn. Commun. Netw.*, vol. 3, no. 3, pp. 361–373, Sep. 2017.
- [24] H. Cao and J. Cai, "Distributed multiuser computation offloading for cloudlet-based mobile cloud computing: A game-theoretic machine learning approach," *IEEE Trans. Veh. Tech.*, vol. 67, no. 1, pp. 752–764, Jan. 2018.