On the Benefits and Implementation Costs of Multi-cell Selection in Heterogeneous Networks

Ararat Shaverdian and Catherine Rosenberg
Dept. of Electrical and Computer Engineering, University of Waterloo
Email: {ashaverd, cath}@uwaterloo.ca

***Abstract*—We consider a heterogeneous cellular network (HetNet) on the downlink and focus on *multi-cell selection* (MCS). MCS allows each user to associate with and receive data from multiple base stations (BTSs) at once, potentially boosting the network performance. We formulate a fully coordinated realistic MCS problem that provides an upper-bound on the network performance which is 20% above that of the state-of-the-art *single-cell selection* (SCS). This improvement comes at the cost of inter-BTS coordination which is not so easy to perform in practice. Hence, we focus on how to obtain this performance gain in two practical scenarios: 1) a conventional HetNet with no inter-BTS coordination, 2) a Centralized Radio Access Network (C-RAN) HetNet. In Scenario 1, we show that SCS with periodic individual opportunities for re-associations along with local Round Robin scheduling (which requires no inter-BTS coordination) can outperform the state-of-the-art SCS by 17% without incurring a huge cost in signaling. In Scenario 2 we show that, using a simple two-step heuristic, we can almost reach the upper-bound without paying for the complexity in coordinated scheduling.**

I. INTRODUCTION

Heterogeneous networks (HetNets) have been envisaged as a promising solution to the ever-increasing traffic demand in cellular networks. A HetNet consists of a set of macro cells, each comprised of a macro base station (MBTS) and a number of small cells (SCs). These SCs have lower transmit power, coverage area and maintenance costs than MBTSs and can increase the system throughput by offloading a fraction of the MBTSs users or filling the MBTSs coverage holes [1]. In the context of enabling technologies for 5G HetNets, we focus on *multi-cell selection* (MCS) where, upon arrival, a user may select multiple base stations (BTSs) for service[1]. Choosing the right user-to-BTS associations can potentially improve the system performance by allowing users to receive service from multiple BTSs without over-burdening them. In this paper, we consider the problem of MCS in a HetNet where, in each macro cell, the SCs are connected to the MBTS via backhaul links of sufficient capacity. Our first objective is to quantify the performance gains that could be obtained with MCS with respect to (w.r.t.) the state-of-the-art *single-cell selection* (SCS). The premise is that only if the gains are significant, should we then consider implementing MCS.

If there are significant performance gains in MCS, our second objective would be to study the possible complexities involved in its implementation. The following considerations are important when implementing an MCS scheme:

- **Performance improvement** compared to the state-of-the-art SCS scheme, whether the performance metric is the network throughput, user delay or fairness in throughput assignment to users.
- **Tractability of MCS problem** which involves the right assignment of the users to BTSs to, for example, avoid simultaneous transmission of co-channel BTSs to the same user and/or the ease of implementing an efficient scheduler for a given set of user-to-BTS assignments.
- **Low overhead** involved in signaling among BTSs if inter-BTS coordination through the backhaul links is required (e.g., for joint user scheduling), and/or in signaling between users-to-BTSs (e.g., for re-association).

We will consider the above-mentioned features in two practical scenarios. In Scenario 1, upon arrival, a user may associate with multiple BTSs each of which schedules its associated users locally without the cooperation of the other BTSs or the coordination of a central unit. This scenario corresponds to a conventional HetNet where inter-BTS cooperation is not permitted and, so, MCS may not be a feasible option. However, one possibility is to find a way to improve on the performance of an SCS scheme to achieve the performance bounds of MCS.

In Scenario 2, we consider the problem of MCS in a Centralized (or Cloud) Radio Access Network (C-RAN) HetNet. C-RAN provides a (fully or partially) centralized paradigm for inter-BTS coordination in HetNets [2]. The idea of C-RAN is to locate the baseband units from multiple BTSs into a central unit with high processing power and allocate baseband resources to remote radio heads at the BTSs via high-capacity and low-latency wirelines (a.k.a., fronthauls) between the central unit and radio heads. Lower antenna site installation costs, reduced energy consumption and better radio resource management are among many drivers for the deployment of C-RAN [3]. In addition, C-RAN enables all user-related signal processing tasks to be carried out at a central unit with greater computational power than conventional BTSs processors. This is particularly important in terms of resource management, as C-RAN can jointly manage the radio resources via the fronthauls yielding a potentially high performance gain.

We summarize our contributions in the following. We consider the downlink of a multi-macro cell HetNet where the underlying resource allocation scheme is orthogonal deployment, i.e., different types of BTSs (i.e., macro and small BTSs) in a given macro cell use orthogonal channels while the BTSs of the same type use the same set of channels.

---

[1] We will use the terms *cell selection* and *user association* interchangeably.

1) We formulate a joint MCS and scheduling convex optimization problem to obtain an upper-bound on the proportionally fair throughput performance of the system. Our problem can be reduced to a pure scheduling problem, since all users are allowed to associate with all BTSs, and is easy to solve using commercial solvers. Furthermore, it does not permit the possibility of simultaneous transmissions of two or more co-channel BTSs to the same user. The resulting upper-bound corresponds to an ideal MCS with the possibility of full coordination among BTSs (i.e., all information required for scheduling is available centrally). We show that MCS with full coordination significantly outperforms the sate-of-the-art SCS in [4]. We also show that a simple MCS scheme, called SMCS, in which a user associates with the MBTS and the SC offering the highest link rate is quasi-optimal. However, the main drawback of coordinated MCS is that the improvement in performance comes at the cost of an increased complexity in user scheduling compared to SCS as a result of inter-BTS coordination.

Since the implementation of a fully coordinated scheduler may not be a feasible option in today's HetNets, we further consider the problem of obtaining the gain in performance in practical scenarios. We consider the two scenarios described earlier.

2) In Scenario 1, we consider SMCS and show that this MCS scheme performs poorly *if deployed* along with local Round Robin scheduling (RR). Therefore, to achieve a performance close to the upper-bound, a more coordinated scheduling scheme is necessary which is difficult to devise. Consequently, we do not attempt to implement MCS in this scenario and, instead, propose to keep the state-of-the-art SCS scheme along with its optimal scheduling (which is shown to be the local RR in [1]) but to complement it with *individual* periodic opportunities for re-associations, i.e., to allow a user to independently and periodically check if the association she has is still the best. This scheme is extremely simple and performs close to the upper-bound. However, it incurs two implementation costs: periodic but simple computations at the user end (if the cell selection is device-centric) or at the network otherwise, and the re-association signaling between a user and her BTS *when a re-association effectively takes place*. Fortunately, as we will show, the signaling can be drastically reduced by limiting the number of re-associations per-user (at the cost of only a small loss in performance).
3) In Scenario 2, although the optimal (i.e., fully coordinated) user scheduling may be done centrally via C-RAN, it probably cannot be performed fast enough (e.g., on a per LTE-frame basis) for real-time user scheduling. Consequently, we propose a simple two-step algorithm inspired by the scheme in Scenario 1 and show that it performs quasi-optimally at no extra signaling cost.

The rest of the paper is organized as follows. Sections II and III outline the related work and system model. In Sections IV and V, we present our global MCS problem formulation and provide a few interesting insights into MCS and SCS. In Section VI, we present our proposed heuristics. We, then, conclude the paper with numerical results and concluding remarks in Sections VII and VIII.

II. LITERATURE REVIEW

The authors in [5] propose a fractional user association rule to maximize the proportional fairness (PF) utility, allowing users to associate with multiple BTSs. The fractional association problem results from relaxing the integrality constraint (which enforces assigning a maximum of one BTS to each user) of the SCS problem. Furthermore, they derive an SCS scheme by deploying a gradient projection method where users and BTSs cooperatively find near-optimal association via iteratively exchanging a number of parameters. They show that there is no significant gains achieved by fractional association compared to SCS in a static system. However, the complexity of their proposed iterative method may not be suitable for an online system. Similar results are shown in [6]. In [7], the authors formulate a joint power allocation and MCS problem in a co-channel HetNet and propose a near-optimal iterative algorithm to maximize the proportional fairness utility. They show that their MCS scheme outperforms SCS in terms of fairness and minimum rate (in bps/Hz) in a static system. However, they assume multiple BTSs can simultaneously serve one user on the same band. Moreover, they use the long-term average of downlink interference instead of the exact value in their calculations. [8] is another work which deals with joint transmission of multiple BTSs to the same user on the same band. It extends various SCS schemes to MCS where a user selects one SC and one MBTS using a simple algorithm and compares the performance of the schemes. However, this work does not focus on optimal scheduling and, instead, uses RR for simplicity. In [9], the authors consider a joint cell selection and scheduling problem which incorporates ON-OFF coordination between BTSs. They show that MCS is not easy to solve when ON-OFF coordination is incorporated in conventional HetNets and, so, they propose an iterative algorithm to solve it. They compare the performance of MCS to SCS in a static setup and show that the difference is very small. However, they do not evaluate their problem in an OFDMA-based system. All of the aforementioned works assume a static setup for performance evaluation without considering the dynamics of an online system. In the following, we outline the literature work in dynamic systems.

The authors in [4] show that pure physical layer-based SCS schemes can perform poorly in online systems. They propose a simple device-centric SCS scheme which incorporates a small amount of network-awareness, and optimizes the global objective of the network at the time of users' arrival. However, with their proposed scheme, the resulting network performance still maintains a large gap from the performance bound of the optimal SCS since a user keeps her original association even when the system state changes. To reduce this gap, the authors in [10] incorporate periodic re-association opportunities in

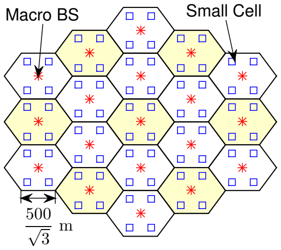

Fig. 1. HetNet configuration with a reuse factor (r) of 3. The neighboring interfering cells are highlighted in yellow.

their scheme where each device individually decides whether or not to change her association based on a simple formula. Although such frequent re-associations improve the system performance, they create a high level of user-to-BTS signaling which can cause implementation issues. [10] does not discuss such implementation issues and only focuses on SCS (and not MCS). The authors in [11] address the problem of maximizing PF utility while allowing MCS. They propose two near-optimal algorithms w.r.t. the global PF utility, with and without inter-BTS communication. Their simulation results suggest that their scheduling scheme outperforms their SCS PF scheme in terms of user throughput, throughput oscillation and fairness. However, the authors do not consider intercell interference.

III. SYSTEM MODEL

We consider the downlink of an OFDM-based multi-macro cell HetNet. The HetNet as a whole uses M' OFDM subchannels[2] (each of bandwidth b) and each macro cell is allotted $M = \frac{M'}{r}$ subchannels, where $r \geq 1$ is the reuse factor. We focus on one cell in the middle of the HetNet and consider the intercell interference from the neighboring macro cells. The cell in the middle consists of one MBTS, indexed by 0, overlaid with a set of B SCs, $\mathcal{B} := \{1, 2, ..., B\}$, indexed by j and a set of users indexed by i (see Fig. 1). We assume that a user will not associate with an MBTS or SC in another macro cell. The BTSs constantly transmit to their users while in the system. The transmit power of the MBTSs and SCs are denoted by P_0 and P_P, respectively; both of which are fixed and known. The underlying resource allocation scheme is orthogonal deployment, i.e., in a given macro cell, each of the SCs is given the same k channels while the MBTS is allotted $M-k$ dedicated channels orthogonal to the SC channels. We assume k is fixed and known and all macro cells use the same channel partitioning pattern and BTS power budgets.

Given that our study is focused on user association and the varying number of users within a system may effect the performance metric, we consider a dynamic system in which users come, according to a predefined random process, and depart after being served. We assume no two users arrive in the system at the same time. The time is slotted into frames and we assume that arrivals and departures occur at the beginning of a frame and that channel gains remain constant within a frame. All users are static (not mobile) and greedy in the sense that they wish to maximize their individual throughputs.

Let $g_{ji}(t)$ denote the channel gain from BTS j to user i at time t (i.e., at the beginning of frame t) which is a function of path loss, antenna gain and shadowing, and can be computed as recommended in [12]. We assume that channels are *flat within a frame* and *known to each BTS* through independent uplink feedback channels, allowing SINR computation at the BTS. We also assume that there is a function, $f(\cdot)$, that maps the SINR on each subchannel to a corresponding efficiency (in bits/symbol), and is known at each BTS. Let $\gamma_{ji}(t)$ denote the SINR seen by user i from BTS j at time t. Denoting the intercell interference from the co-channel SCs and MBTSs by $I_{\text{SC},i}(t),\ I_{M,i}(t)$, respectively, we have

$$\gamma_{ji}(t) = \begin{cases} \frac{(P_0/(M-k))\ g_{0i}(t)}{I_{M,i}(t)+N_0}, & j=0, \\ \frac{(P_P/k) g_{ji}(t)}{\sum_{l\in\mathcal{B}, l\neq j}(P_P/k) g_{li}(t)\ + I_{\text{SC},i}(t)+N_0}, & j\neq 0, \end{cases} \tag{1}$$

where N_0 denotes the additive white Gaussian noise power. The *link rate* $R_{ji}(t)$ seen by user i from BTS j is $k \times b \times f(\gamma_{ji}(t))$ for $j \in \mathcal{B}$ and $(M-k) \times b \times f(\gamma_{ji}(t))$ for $j = 0$. Lastly, we denote by $\mathcal{U}(t)$ the set of all users who see a higher SINR from the MBTS in the cell in the middle (i.e., MBTS 0) than from any other MBTS at time t.

To reduce the scheduling problem from a time-frequency domain to a pure time domain scheduling problem, we assume that a BTS transmits on all the subchannels allotted to it at a given time and distributes its power budget uniformly on its allotted subchannels.

IV. PROBLEM FORMULATION

Our goal is to perform joint user scheduling and multiple cell selection among all BTSs within the macro cell under consideration such that a global fairness in throughput assignment to users is guaranteed on a per frame basis. A commonly used fairness objective, which we use in our framework, is the sum of the logarithm of user throughputs, i.e., $\sum_{i\in\mathcal{U}} \log(\mu_i(t))$ where $\mu_i(t)$ is the user i's throughput [13]. This objective function ensures *proportional fairness* (PF) in throughput assignment to users. Since with MCS we do not place any constraint on the number of BTSs a user can select, the joint problem of cell selection and scheduling can be straightforwardly reduced to a pure global PF scheduling one.

Define a realization $\omega(t) := \{g_{\ell i}(t)\}_{\ell\in\mathcal{B}\cup\{0\}\cup\mathcal{I},\ i\in\mathcal{U}(t)}$ to be the set of channel gains between all (ℓ, i) pairs where $\mathcal{I}$ is the set of BTSs interfering the middle cell. Let $\beta_{ji}(t)$ denote the proportion of the time allotted to user i by BTS j in a given frame t where the realization is ω (we remove the time index for simplicity since we are dealing with a snapshot within a frame). Then, for the given frame, the global PF MCS problem can be formulated as follows.

$$\begin{aligned} \max_{\beta_{ji}(t)\geq 0} & \sum_{i\in\mathcal{U}(t)} \log\left(\sum_{j\in\mathcal{B}\cup\{0\}} R_{ji}(t)\beta_{ji}(t)\right) \\ \text{s.t.} & \sum_{i\in\mathcal{U}(t)} \beta_{ji}(t) \leq 1,\ \forall j \in \mathcal{B}\cup\{0\}, \\ & \sum_{j\in\mathcal{B}} \beta_{ji}(t) \leq 1,\ \forall i \in \mathcal{U}(t), \end{aligned} \tag{2}$$

---

[2] We use the terms channel and subchannel interchangeably.

where the first constraint ensures that the allotted schedules to the users associated with a BTS do not exceed one frame and the second constraint ensures that the allotted schedules to a given user from the SCs do not overlap one another. Note that the second constraint is absent from the MCS formulation in [6], making the results biased and over-optimistic when the number of users is small. This problem is a function of time and called by the system every time the realization ω changes, i.e., when users arrive, depart or channel gains change.

Remark 1. *The joint scheduling problem in* (2) *is a rather simple convex optimization problem. However, it has to be solved centrally since it requires information on all users in the cell and cannot be decoupled into per-BTS sub-problems, hence, incurring a high level of complexity in scheduling.*

Remark 2. *Because of the convexity of the problem and that the optimization is over a set of continuous variables, such a problem can be computed easily using commercial solvers such as* Minos *[14]. However, it is unclear whether it can be computed fast enough to be used online in today's HetNets, especially considering the fact that it has to be re-computed every time ω changes and cannot be (re-)computed locally.*

From Remarks 1-2, we observe that the problem in (2), called *MCS with full coordination*, may not be a feasible option for online systems. However, it still provides a very useful bound since, if the problem is solved as frequently as required, it will yield an *upper-bound* in performance for all other cell selection schemes. It is noteworthy that the joint SCS and scheduling problem (as described in [1]) is much harder to compute because of the integerality constraints which enforce that a user selects only one BTS at a time. In the following, we present a number of results which provide insights into the comparative performance of MCS and SCS.

V. A Few Insights into MCS and SCS

A. A static setting

We consider a snapshot problem with ($|\mathcal{U}(t)| =$) 20 users and ($M =$) 33 subchannels per BS, for all t, where the users are deployed i.i.d. uniformly in the middle cell. We generate a set Ω of 100 realizations and, for each realization $\omega \in \Omega$, we compute a solution to the problem in (2) for each value of $k \in \{1, 2, ..., 32\}$. From each solution, we obtain the geometric mean (GM) throughput corresponding to that realization by computing $(\prod_{i \in \mathcal{U}(t)} \mu_i(t))^{\frac{1}{|\mathcal{U}(t)|}}$. We then compute the average of these GM throughputs over the 100 realizations for each value of k. For the detailed description of the parameters used in the computation and simulation of the results, refer to Section VII-A. The following observation from our computations is in order.

Insight 1. *In a static setting, optimal SCS ([1]) and MCS with full coordination perform almost equally well for all k.*

However, this insight should not be wrongly interpreted as an implication that there is no need for MCS. Recall that in a static setting, each snapshot corresponds to one frame. That is, the result tells us that *if we allow all users to re-associate in every frame (or, more specifically, every time ω changes), only then can SCS perform as well as MCS.* Clearly, this is not a feasible option since solving the SCS problem is highly complex and time-consuming due to its integrality constraints.

B. A dynamic setting

In order to quantify the performance gap between MCS and SCS in an online system, we also consider a dynamic setting where the users arrive in the system according to a homogeneous Poisson point process of rate λ and choose their locations i.i.d. uniformly. To make the interpretation of the results easier, we assume that the users' channels are time-invariant, i.e., each user observes the same channel gain as upon arrival for her complete stay in the system. We assume that each frame duration is 10^{-2} seconds (s). Every arriving user is scheduled in the next frame and leaves the system when she downloads a file of a fixed size F. Clearly, the delay (or service-time) of a user depends on the old and new arrivals and departures. In Fig. 2, we show the average per-user delay as a function of arrival rate λ. The curve labelled *MCS with full coordination* corresponds to the case where problem (2) is computed every time ω changes and, hence, provides a lower-bound on the achievable delay.

Insight 2. *The state-of-the-art SCS proposed in [4], where cell selection for each user is performed only once at the arrival time, performs much worse than the lower-bound (see Fig. 2).*

As a benchmark for MCS schemes, we also consider the cell selection heuristic, namely *best SC and MBTS* (SMCS), where upon arrival each user i associates with the MBTS 0 and the SC j that offers the highest link rate R_{ji}.

Insight 3. *MCS and SMCS offer a similar performance when combined with fully coordinated scheduling (see Fig. 2).*

A question that arises from these insights is whether we can use an MCS scheme (e.g., SMCS), perform local[3] scheduling (e.g., local RR which is optimal for the state-of-the-art SCS [1]) in each BTS and obtain better results than the state-of-the-art SCS. One possibility would be to use *SMCS with RR*.

Insight 4. *SMCS performs poorly with local RR, implying an MCS scheme likely requires a coordinated scheduler or, at the best, devising a good local scheduler is non-trivial (see Fig. 2).*

To verify whether MCS with full coordination can indeed be used in online systems, we need to ensure that the computation time of the problem in (2) is less than a standard frame duration (e.g., 10^{-2} s). For this reason, we generate 100 random realizations, each comprising ($|\mathcal{U}(t)| =$) 30 i.i.d. users. For each realization, using the commercial solver Minos and a powerful server (Intel® Xeon® CPU E5-2697 v2 @ 2.70GHz with 12 cores and 48 CPUs), we solve the problem in (2).

---

[3] Note that by local, we mean uncoordinated per-BTS scheduling.

Insight 5. *The maximum, minimum and average computation times over 100 realizations are* 0.0216, 0.0019 *and* 0.0184 s*, respectively, indicating that the computation time typically takes longer than a standard LTE frame. Hence, a faster way to compute user schedules, without a huge loss in performance w.r.t. MCS with full coordination, is required.*

In the following, we use the Insights 1-5 to propose high-performing online heuristics for two scenarios: 1) a conventional HetNet where there is no coordination among BTSs and each BTS schedules its users locally, 2) a C-RAN HetNet where BTSs jointly coordinate their schedules via the C-RAN.

VI. HEURISTIC SCHEMES

A. Scenario 1: Conventional HetNet

In Scenario 1, since BTSs cannot jointly coordinate their schedules and local RR performs poorly with MCS, MCS is not considered an option to achieve the performance bound of MCS with full coordination. Instead, what we propose to do is to stay with SCS and periodically give opportunities to each user to re-associate. However, to implement such a scheme there is a need for a good device-centric SCS rule. The rule that we use is a device-centric SCS rule proposed in [4]. Specifically, our proposed scheme, called *SCS with re-association*, is as follows. Upon arrival, user u associates with BTS j_u^* such that

$$j_u^* = \arg \max_{j \in \mathcal{B} \cup \{0\}} \log \Big(\frac{R_{ju}}{U_j + 1} \times (\frac{U_j}{U_j + 1})^{U_j} \Big), \qquad (3)$$

where U_j is the number of users associated with BTS j. The network, then, gives an opportunity to the user to re-associate every T seconds. Upon receiving an opportunity to re-associate, the user will (virtually) dissociate from her current BTS j_u^* and, again, perform the rule in (3). We expect periodic re-associations to improve on the myopic decisions as the network evolves. In Section VII-B, we will show the impact of T on the performance of our heuristic, fully recognizing that if T becomes very small, the number of effective re-associations will increase and so will the signaling traffic between the users and BTSs. To try to quantify this signaling traffic, we will also show the number of effective re-associations as a function of T.

B. Scenario 2: C-RAN HetNet

In this scenario, since BTSs can jointly schedule their users via a C-RAN, it will be necessary to design a heuristic scheduling scheme only if the problem in (2) cannot be solved fast enough, e.g., within a frame duration. In that case, we propose to restrict a user's association to the MBTS and the SC that offers the highest link rate R_{ji} (a.k.a., SMCS) from her arrival until departure (to avoid co-channel transmission collisions from SCs). Furthermore, we propose the following two-step heuristic for scheduling, called *SMCS with C-RAN*. In the first step, the network decides which one of each user's two BTSs will schedule her in the next frame. This is done *internally* via C-RAN using the rule (3) at the arrival time of the user. In every frame when there is no arrival, the heuristic selects one user in the C-RAN at random and applies the rule (3) to this user to re-compute which BTS will schedule her next. In the second step, after all users in a frame are assigned to a BTS, each BTS schedules its users locally using RR. The rationale behind our two-step heuristic is to restrict each user to one of her two associated BTSs in a given frame so as to be able to use RR on a per-frame basis. In effect, we reproduce the heuristic in Scenario 1 without the shortcoming of the increase in signaling since all decisions are made internally.

Next, we evaluate the performance gains of our heuristics.

VII. NUMERICAL RESULTS

A. Parameter settings and simulation setup

We consider the HetNet with 19 hexagonal cells as shown in Fig. 1 with a reuse factor (r) of 3. Each cell, with radius $500/\sqrt{3}$ m, consists of one MBTS, located in the center, and four SCs located 230 m away, along the radius, from the MBTS. The physical layer parameters are based on the 3GPP evaluation methodology document [12] used for LTE HetNets. The rate function $f(.)$ is taken as the 15-rate modulation and coding scheme available in LTE, as shown in Table III of [1]. In the dynamic setting, for a given arrival rate λ, we run 5 simulations each with a different random seed. Each simulation runs for a period of at least 1000 s. For convergence, the period of simulation is increased at a step of 1000 s until the GM throughput is within 5% of the GM throughput before the increase. To compute the average per-user delay, for each user we record the total time that she spends in the system, i.e., until she completes downloading her file of size F. We compute the average per-user delay per-simulation by taking the arithmetic mean of the delays of all users who have departed from the system over the simulation period. We then take the arithmetic mean of these quantities over the set of 5 simulations (corresponding to different random seeds) to obtain the average per-user delay. The number of simulations are chosen such that the average delay corresponding to each simulation falls within 5% of that corresponding to every other simulation. To avoid biasing the results in favor of one scheme, we obtain the results for each scheme when the resource allocation parameter k is optimized, i.e., for each λ (for a given scheme) we run the 5 simulations per k and keep the k that yields the lowest average delay for that λ. We have set $M = 33$ and $F = 10$ Mbits.

B. Numerical results for Scenario 1

Fig. 2 shows the average per-user delay of various schemes as a function of λ. We have only shown the values of λ that result in an average of less than 30 users per macro cell for the optimal MCS. We observe that by incorporating periodic re-associations, we can considerably improve on the delay performance, particularly for high values of λ. As expected, by increasing the frequency of re-associations T, the performance approaches the lower-bound offered by the optimal MCS. SCS with re-association with a periodicity of $T = 0.5$ s

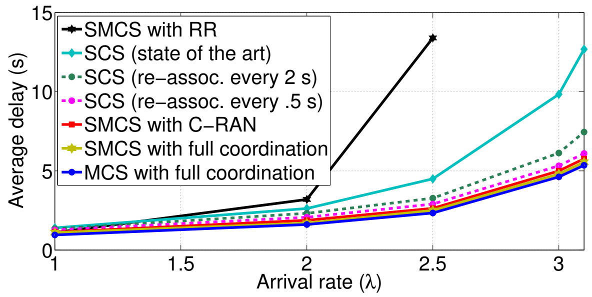

Fig. 2. Dynamic setting: Average per-user delay vs. arrival rate λ (users/s).

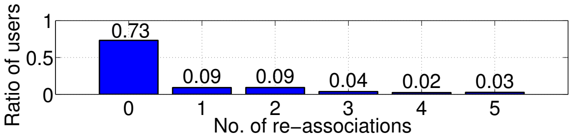

Fig. 3. Ratio of users vs. the number of times they re-associated over the simulation time. The heuristic used is SCS with re-assoc. ($T = 0.5$ s, $q = 5$).

can support up to 17.56% higher λ w.r.t. the state-of-the-art SCS and up to 33.04% higher λ w.r.t. SMCS with RR. At $\lambda = 3.1$, out of every $(5.35/0.5 \approx)$ 10.7 re-association opportunities given to each user on average, only ≈2.34 of them are effective but some users see a much higher number of effective re-associations. In order to reduce the incurred signaling costs of re-associating users, we decided to limit *the maximum number of re-associations allowed per-user, q,* to 5. With this modification, the number of effective re-associations reduces to 0.61 s (from 2.34) per user, while the average delay increases to 6.39 s (from 6.09). This implies that the signaling cost can be significantly reduced for only a small amount of loss in performance. Clearly, the performance exhibits a trade-off between q and the actual number of effective re-associations. We have shown the average delay and number of effective re-associations as a function of T and q in Table I. It is noteworthy that although SCS with re-association still maintains a small gap of around 3% with the lower-bound, it is computationally significantly faster than MCS with full coordination which requires the re-computation of the problem in (2) whenever ω changes. This is because the proposed heuristic only requires a user device to apply the rule in (3) and the BTSs to perform local RR, periodically.

Fig. 3 shows the ratio of users as a function of the number of times they re-associated during the simulation time where we have set $T = 0.5$ s, $q = 5$. The histogram shows that of all the users who departed from the system during the simulation time, only 9% re-associated more than twice.

C. Numerical results for Scenario 2

In Scenario 2, SMCS with C-RAN can support up to 19.23% and 34.78% higher λ w.r.t. the state-of-the-art SCS and SMCS with RR, respectively. Furthermore, it maintains a gap of only 1% w.r.t. the optimal MCS while eliminating all the costs of re-association signaling since all decisions are made internally via C-RAN and low-latency fronthauls.

TABLE I

AVERAGE DELAY AND NO. OF EFFECTIVE RE-ASSOCIATIONS VS. T AND q, FOR A FIXED ARRIVAL RATE OF $\lambda = 3.1$ (USERS/s).

Scheme	T (s)	q	Average delay (s)	No. of effective re-assoc. per user
MCS with full coo.	-	-	5.353221	None
SCS with re-assoc.	0.5	∞	6.09681738	2.34233705
SCS with re-assoc.	0.5	5	6.39294533	0.60766394
SCS with re-assoc.	2	5	6.9275982	0.3628347

VIII. CONCLUSION

We considered the problem of MCS in a conventional and C-RAN HetNet and showed that it significantly improves the system performance w.r.t. the state-of-the-art SCS. We proposed a heuristic for each scenario to achieve the performance bound of the optimal MCS. The proposed heuristics are simple and perform very well at a low signaling cost for the conventional case and at no signaling cost for the C-RAN case. As a part of our future work, we plan to adapt these schemes to scenarios with mobile users and time-varying channels.

REFERENCES

[1] D. Fooladivanda and C. Rosenberg, "Joint resource allocation and user association for heterogeneous wireless cellular networks," *IEEE Trans. on Wireless Communications*, vol. 12, no. 1, pp. 248–257, January 2013.
[2] C. M. R. Institute, *C-RAN The Road Towards Green RAN*, December 2013). [Online]. Available: http://labs.chinamobile.com/cran/wp-content/uploads/2014/06/20140613-C-RAN-WP-3.0.pdf
[3] A. Pizzinat, P. Chanclou, F. Saliou, and T. Diallo, "Things you should know about fronthaul," *Lightwave Technology, Journal of*, vol. 33, no. 5, pp. 1077–1083, March 2015.
[4] J. Ghimire, "Heterogeneous Cellular Networks: from Resource Allocation to User Association," Ph.D. dissertation, University of Waterloo, 2015. [Online]. Available: https://uwspace.uwaterloo.ca/handle/10012/9276
[5] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "User association for load balancing in heterogeneous cellular networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 6, pp. 2706–2716, June 2013.
[6] J. Ghimire and C. Rosenberg, "Resource allocation, transmission coordination and user association in heterogeneous networks: A flow-based unified approach," *Wireless Communications, IEEE Transactions on*, vol. 12, no. 3, pp. 1340–1351, March 2013.
[7] Y.-C. Chen and H.-Y. Hsieh, "Joint resource allocation and power control for comp transmissions in LTE-A hetnets with RRHs," in *Wireless Communications and Networking Conference (WCNC), 2014 IEEE*, April 2014, pp. 1368–1373.
[8] Y. Cao, H. Xia, and C. Feng, "Evaluation of diverse cell range expansion strategies applying comp in heterogeneous network," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*, Sept 2013, pp. 1962–1966.
[9] Q. Kuang, W. Utschick, and A. Dotzler, "Optimal joint user association and multi-pattern resource allocation in heterogeneous networks," *IEEE Trans. on Signal Processing*, vol. 64, no. 13, pp. 3388–3401, July 2016.
[10] J. Ghimire, A. Shaverdian, and C. Rosenberg, "Device-centric α-fair user association in heterogeneous cellular networks," *Submitted to IEEE Transactions on Wireless Communications*, 2016.
[11] H. Zhou, P. Fan, and J. Li, "Global proportional fair scheduling for networks with multiple base stations," *Vehicular Technology, IEEE Transactions on*, vol. 60, no. 4, pp. 1867–1879, May 2011.
[12] 3GPP Tech. Rep. TR 36.814. (2010) Evolved Universal Terrestrial Radio Access (E-UTRA); Further Advancements for E-UTRA Physical Layer Aspects. [Online]. Available: http://www.3gpp.org/dynareport/36814.htm
[13] F. Kelly, "Charging and rate control for elastic traffic," *European transactions on Telecommunications*, pp. 33–37, 1997.
[14] Minos 5.5. [Online]. Available: http://www.sbsi-sol-optimize.com/asp/sol_product_minos.htm