HAL
open science

Network-aware User Association in Public Safety Oriented Mobile Networks

Jad Oueis, Razvan Stanica, Fabrice Valois, Catherine Rosenberg

To cite this version:

Jad Oueis, Razvan Stanica, Fabrice Valois, Catherine Rosenberg. Network-aware User Association in Public Safety Oriented Mobile Networks. I-TENDER 2017 - 1st Workshop on ICT Tools for Emergency Networks and Disaster Relief, Dec 2017, Seoul, South Korea. pp. 1-6. hal-01625595

HAL Id: hal-01625595

https://hal.science/hal-01625595v1

Submitted on 27 Oct 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Network-aware User Association in Public Safety Oriented Mobile Networks

Jad Oueis, Razvan Stanica, Fabrice Valois
Univ Lyon, INSA Lyon, Inria, CITI
F-69621 Villeurbanne, France
Email: {jad.oueis, razvan.stanica, fabrice.valois}@insa-lyon.fr

Catherine Rosenberg
Department of Electrical and Computer Engineering
University of Waterloo, Canada
Email: cath@uwaterloo.ca

***Abstract*—In many disaster scenarios, communication infrastructure fails to provide network services for both civilians and first responders. One solution is to have rapidly deployable mobile networks formed by interconnected base stations, that are easy to move, deploy, and configure. Such public safety-oriented networks are different from classical mobile networks in terms of scale, deployment, and architecture. In this context, we revisit the user association problem, for two main reasons. First, the backhaul, formed by the links interconnecting the base stations, must be accounted for when deciding on the association, since it may present a bottleneck with its limited bandwidth. Second, the mission-critical nature of the traffic imposes strict guaranteed bit rate constraints, that must be respected when associating users. Therefore, we propose a network-aware optimal association that minimizes the bandwidth consumption on the backhaul, while still respecting the stringent performance requirements.**

I. Introduction

A major challenge in the aftermath of natural and man-made disasters is communication. Ensuring reliable communication, notably between first responders, such as paramedics, firefighters, and police officers, is crucial for saving people's lives [1]. However, communication failures are common in emergency situations, as recently seen with the tropical storm Harvey, which caused service outage in up to 90% of the cell sites in some of the affected regions [2].

As a result, we recently witnessed an increased investment in public safety (PS) oriented networks, tailored to meet the stringent requirements of first responders, in terms of reliability, resiliency, and security [3]. LTE has already been adopted as the building technology for these networks, due to its large existing ecosystem, high bandwidth, low latency, and real-time communication capabilities [3]. One of the foreseen solutions is to have easily and rapidly deployable networks, replacing the communication infrastructure when the latter is damaged or destroyed, or in isolated areas where no infrastructure preexists [4]. The key elements of rapidly deployable networks are small-sized base stations (BSs), that are easy to move, configure and deploy, potentially in an autonomous manner. Deployment platforms include aerial platforms (e.g., balloon fleets [5], drones [6]), and terrestrial networks (e.g., BSs mounted on vehicles or wearable as backpacks [7]).

Several differences exist between classical and PS mobile networks. In classical mobile networks, each BS of the radio access network (RAN) usually has a dedicated backhaul link towards the core network. This architecture requires a standalone core network hardware and backhaul provisioning. In PS networks, the architecture must be simplified, since a dedicated core network may not be available, and the BSs are supposed to be as autonomous as possible. Hence, the core network functionalities are to be co-located with the BSs [4], through function virtualization, for example [8]. In this case, all traffic, usually forwarded to the standalone core network, is now routed locally on the links interconnecting the BSs and forming the backhaul. Moreover, PS networks are relatively small-sized. They interconnect only a few BSs, generally less than 10, depending on the needed coverage area, and serve a limited number of users (e.g., first responders), depending on the scale of the emergency situation in question. Furthermore, the traffic in PS networks is mostly intra-cellular, i.e., between users belonging to the same network. Indeed, even when communication with external networks is possible, it is usually rare.

In this paper, we first revisit the user association problem, i.e., assigning each user to a BS, in the context of a rapidly deployable PS mobile network. User association is usually based on RAN metrics, with the assumption of an over-provisioned backhaul. However, the backhaul in our scenario may represent a bottleneck, especially since the links between BSs, forming this backhaul, have a limited bandwidth. With multiple BSs potentially sharing the same backhaul link to forward their traffic, ignoring the backhaul will eventually lead to poor network performance. On top of that, traditional association schemes do not take user requests into consideration when deciding on the association. Nevertheless, with the stringent performance requirements imposed in PS networks, users must be granted services with a guaranteed bit rate. Indeed, video streaming lag during a critical rescue operation cannot be tolerated. Moreover, we revisit the user attachment problem, i.e., assigning the flows of each user to a specific gateway. Usually, each user is attached to one default gateway responsible of locally routing all of its flows. We argue that an optimized attachment would further reduce the traffic on the backhaul by reducing the signaling load caused by gateway selection and management.

The contributions of this paper are twofold. First, we formulate a network-aware user association optimization problem. The problem balances different constraints: the limited

resource availability on the RAN, the guaranteed throughput requested by each flow between two users, and the data and signaling traffic routed on the backhaul. The problem is solved with the objective of minimizing the backhaul bandwidth consumption.

Second, we separate the user association from the user attachment. We propose different attachment schemes, and formulate the corresponding optimization problems. Each problem returns both the optimal user association and the optimal attachment to a gateway, with the same objective described above.

Results show that our network-aware association significantly reduces traffic on the backhaul in comparison with a traditional RAN-based association scheme, leading to a backhaul bandwidth consumption reduction of 40%. Moreover, when the attachment is optimized alongside the association, the backhaul consumption reduction is even more important, with the backhaul signaling traffic further reduced.

The remainder of the paper is organized as follows. An overview on related work is presented in Sec. II. In Sec. III, the network model is described. We present the user association and attachment optimization problems in Sec. IV, and discuss numerical results in Sec. V. Sec. VI concludes the paper.

II. Related Work

The conventional association policy in mobile networks is based on the downlink (DL) signal strength, where a user associates to the BS from which he gets the maximum signal to interference plus noise ratio (SINR) on the DL [9]. Several studies deemed this best SINR approach as ineffective for several reasons, including its poor load balancing [10], and its complete disregard for the uplink (UL) [11] and the backhaul [12].

A multitude of association policies outperforming the best SINR approach have been proposed, for different network architectures, and with different objectives [9], such as optimizing rates or delay in homogeneous networks [10], balancing BS loads in heterogeneous networks (HetNets) [13], and optimizing coverage in self-organizing networks [14]. The majority of the existing works tackle the problem from a RAN point of view, completely ignoring the backhaul. Nevertheless, some backhaul-aware association policies were proposed, notably for backhaul-limited HetNets [12].

To the best of our knowledge, the user association problem has not been studied yet in the context of rapidly deployable networks. Besides the different network architecture, the major difference with prior works is that we consider a guaranteed throughput approach, where a user is granted the exact throughput it asks for. Moreover, we jointly account for both the RAN and the backhaul when deciding on user association.

III. Network Model

We consider a PS mobile network, where $\mathcal{J}$ is the set of rapidly deployable BSs, $\mathcal{L}$ is the set of bidirectional links interconnecting the BSs, and $\mathcal{U}$ is the set of user equipments (UEs).

Backhaul- The core network entities are co-located with the BSs. We focus on two main entities: the mobility management entity (MME), handling network management such as paging, authentication, and gateway selection, and the serving gateway (S-GW), handling local data routing. We consider that network management is ensured by a single MME entity, co-located with one of the BSs, and that each BS of the network is co-located with a S-GW. All signaling traffic passes through the MME, and all data traffic passes through the S-GW. Hence, all traffic is routed locally via the links interconnecting the BSs, referred to as inter-BS backhaul links since they form the backhaul network. Without loss of generality, we consider that, regardless of the inter-BS links technology, there is no contention between the links for resource utilization. We assume that potentially interfering links are operating on distinct channels, allowing interference-free parallel transmissions on the backhaul [15].

Traffic- We model data traffic as bidirectional symmetric flows between two UEs. We only consider intra-cellular flows, where both UEs of a flow belong to the same network. Let $\mathcal{F}$ be the set of flows, and $f = \{u, v\} \in \mathcal{F}$ a bidirectional flow between UEs u and v. A flow exists between two UEs with a given probability $p \in [0, 1]$. A single UE can have several simultaneous flows with different UEs.

We denote by d_f the requested data rate of a flow f, in bits/second, in each direction. Each flow f is accompanied by signaling traffic: between the MME and the BS to which the UE is associated, and between the MME and the S-GW to which the UE is attached.

Routing- Two BSs may or may not be directly connected via a backhaul link. Traffic transported on the backhaul between the entities is routed either directly, if the two end-entities are at one hop from each other, or through the interconnected BSs in a multi-hop fashion. In the latter case, we adopt a minimum-hop routing policy, where the shortest path in terms of number of hops is selected for routing between two BSs. We define $Z^l_{j,j'}$ as a boolean, such that $Z^l_{j,j'} = 1$ if link $l \in \mathcal{L}$ belongs to the shortest path between BSs j and j', and $Z^l_{j,j'} = 0$, otherwise.

RAN- We consider a time-frequency system (e.g., OFDMA) with a total of K_j orthogonal channels allocated to each BS j. The system is in Frequency Division Duplexing (FDD) mode, such that distinct channels are used for DL and UL, with an FDD ratio $\alpha \in (0, 1)$, such that $K^{DL}_j = \alpha \cdot K_j$ and $K^{UL}_j = (1 - \alpha) \cdot K_j$, respectively. We assume that all BSs are identical in terms of transmit power and antenna gain. We adopt a physical layer model similar to the one used in [16]. We denote by $R^{DL}_{u,j}$ and $R^{UL}_{u,j}$ the maximum per channel physical rate UE u can get from BS j, in bits/second, on the DL and the UL, respectively. We define $\theta(\cdot)$ as a function mapping a given SINR to the corresponding physical per channel rate, such that $R^{DL/UL}_{u,j} = \theta\left(SINR^{DL/UL}_{u,j}\right)$ [16]. Each flow is granted exactly the data rate it requests. For each of its flows, a UE is granted a fraction of the available channels on the BS it is associated to, depending on the

physical per channel data rate it gets from that BS. The sum of channel fractions allocated to all users associated to a BS cannot exceed the maximum number of channels allocated to that BS.

IV. Optimizing User Association

We first propose a network-aware user association policy that takes into consideration the RAN, the backhaul, and the flow properties. Our goal is to determine to which BS each UE must associate in such a way that the bandwidth consumption on the backhaul is minimized, in order to avoid backhaul saturation. We determine, in the same problem, the optimal position of the MME, i.e., with which BS the MME must be co-located. An optimal MME placement is key, since all signaling traffic passes through that MME [17].

Each BS of the network must be served by the MME co-located with one of the BSs. Each UE must be associated to one and only one BS, and attached to at least one S-GW. In this work, we study three attachment policies, leading to three different problem formulations:

$\mathcal{P}\mathbf{1}$: user attachment follows association. A UE is attached to the S-GW co-located with the BS it is associated to.

$\mathcal{P}\mathbf{2}$: user attachment does not necessarily follow association. A UE can be attached to any S-GW co-located with any BS, even if different from the BS it is associated to. In this case, we optimize attachment in addition to the association.

$\mathcal{P}\mathbf{3}$: attachment per flow. A UE is attached to one S-GW per flow. A UE can attach to different S-GWs for its different flows. In other words, UEs at both ends of a flow are attached to the same S-GW. The attachment is optimized on a per-flow basis in addition to the association.

Each of these problems is formulated as a mixed integer quadratic optimization problems (MIQP). We limit this work to the offline problem, where UEs are already in the network, with known positions, flows and respective data rates.

A. $\mathcal{P}1$: *user association optimization*

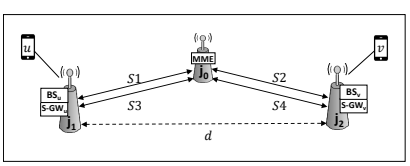

Fig. 1. Data and signaling traffic between BSs for a bidirectional flow in problem $\mathcal{P}1$.

We first formulate problem $\mathcal{P}1$, in which attachment follows association. Fig. 1 shows the data and signaling traffic paths for a bidirectional flow between UEs u and v, respectively associated to BSs j_1 and j_2. If u (resp. v) is associated to BS j_1 (resp. j_2), then u (resp. v) is attached to S-GW j_1 (resp. j_2). Hence, when two UEs are associated to the same BS, the data they exchange is not routed on the backhaul.

$$\mathcal{P}\mathbf{1}: \min_{W,X} \sum_{l\in\mathcal{L}} \left(C_l^d + C_l^{S1} + C_l^{S2} + C_l^{S3} + C_l^{S4} \right) \tag{1}$$

$$\sum_{j\in\mathcal{J}} W_j = 1 \tag{2}$$

$$\sum_{j\in\mathcal{J}} X_{u,j} = 1\ ;\ \ \forall u \in \mathcal{U} \tag{3}$$

$$\sum_{u\in\mathcal{U}} \frac{X_{u,j}}{K_j^{DL}\cdot R_{u,j}^{DL}} \sum_{f\in\mathcal{F}/u\in f} d_f \le 1\ ;\ \ \forall j \in \mathcal{J} \tag{4}$$

$$\sum_{u\in\mathcal{U}} \frac{X_{u,j}}{K_j^{UL}\cdot R_{u,j}^{UL}} \sum_{f\in\mathcal{F}/u\in f} d_f \le 1\ ;\ \ \forall j \in \mathcal{J} \tag{5}$$

$$C_l^d = \sum_{f\in\mathcal{F}} \Big(\sum_{j_1\in v} X_{u,j_1} \sum_{j_2\in\mathcal{J}} X_{v,j_2}\cdot Z_{j_1,j_2}^l \cdot 2d_f \Big)\ ;\ \forall l \in \mathcal{L} \tag{6}$$

$$C_l^{S1} = \sum_{f\in\mathcal{F}} \Big(\sum_{j_1\in\mathcal{J}} X_{u,j_1} \sum_{j_0\in\mathcal{J}} W_{j_0}\cdot Z_{j_1,j_0}^l \cdot S1_f \Big)\ ;\ \ \forall l \in \mathcal{L} \tag{7}$$

$$C_l^{S2} = \sum_{f\in\mathcal{F}} \Big(\sum_{j_2\in\mathcal{J}} X_{v,j_2} \sum_{j_0\in\mathcal{J}} W_{j_0}\cdot Z_{j_2,j_0}^l \cdot S2_f \Big)\ ;\ \ \forall l \in \mathcal{L} \tag{8}$$

$$C_l^{S3} = \sum_{f\in\mathcal{F}} \Big(\sum_{j_1\in\mathcal{J}} X_{u,j_1} \sum_{j_0\in\mathcal{J}} W_{j_0}\cdot Z_{j_0,j_1}^l \cdot S3_f \Big)\ ;\ \ \forall l \in \mathcal{L} \tag{9}$$

$$C_l^{S4} = \sum_{f\in\mathcal{F}} \Big(\sum_{j_2\in\mathcal{J}} X_{v,j_2} \sum_{j_0\in\mathcal{J}} W_{j_0}\cdot Z_{j_0,j_2}^l \cdot S4_f \Big)\ ;\forall l \in \mathcal{L} \tag{10}$$

We define association vector X, with $X_{u,j}$ a boolean, such that $X_{u,j} = 1$ if UE u associates to BS j. We define vector W, such that $W_j = 1$ if the MME is co-located with BS j. As constraints, there is only one MME in the network (Eq. 2), and a UE is associated to one BS (Eq. 3). For the RAN, the total flows received from UEs associated to a BS on the DL should not exceed the DL BS capacity (Eq. 4), and the total flows sent by UEs associated to a BS on the UL should not exceed the UL BS capacity (Eq. 5). If, for a given set of users, flows and data rates, there are no sufficient resources on the RAN to support all the flows (i.e., Eq. 4 or Eq. 5 not satisfied), then the problem is unfeasible. In practice, this corresponds to rejecting a new flow request due to insufficient RAN resources.

The data path of the bidirectional flow f, with rate d_f in each direction, goes from j_1 to j_2, and vice-versa (Fig. 1). In this case, a data rate $2\cdot d_f$ (in Mb/s) is consumed by f on each link l on the shortest path between j_1 and j_2, that is on each link l with $Z_{j_1,j_2}^l = 1$. In Eq. 6 we compute C_l^d for each backhaul link l, which is the total bandwidth consumed by the data traffic of all the flows $f \in \mathcal{F}$ on the given link l. Signaling traffic exists for each flow f between MME j_0 and BS j_1, MME $j0$ and BS j_2, MME j_0 and S-GW j_1, and MME j_0 and S-GW j_2. The corresponding signaling bit rates consumed on each link l, belonging to each of those respective paths, are denoted as $S1_f$, $S2_f$, $S3_f$, and $S4_f$, respectively.

From Eq. 7 to Eq. 10, we compute C_l^{S1}, C_l^{S2}, C_l^{S3}, and C_l^{S4}, respectively representing the total bandwidth consumed by the corresponding signaling traffic of all the flows $f \in \mathcal{F}$ on a single link l . Eventually, the bandwidth consumed on each backhaul link l by all the flows is the sum of the bandwidth consumed by both the data and the signaling traffic. Our objective is to minimize the total bandwidth consumed on the backhaul, as formulated in Eq. 1.

B. $\mathcal{P}2$: joint user association and attachment per user optimization

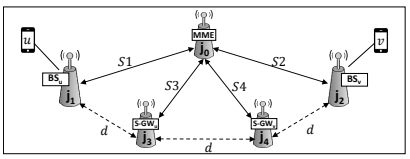

Fig. 2. Data and signaling traffic between BSs for a bidirectional flow in problem $\mathcal{P}2$.

We formulate $\mathcal{P}2$ as an extension of $\mathcal{P}1$, where user attachment is jointly optimized in addition to user association. The MME placement is also optimized.

With S-GWs distributed in the network and co-located with the BSs, the closer the S-GW of a UE is to the MME, the less signaling traffic is consumed on the backhaul. If the S-GW and the MME are co-located, the signaling traffic between them is not routed on the backhaul at all. However, the data traffic must always traverse the S-GW, which is not necessarily placed on the shortest backhaul path anymore.

Fig. 2 shows the updated data and signaling paths for a bidirectional flow between u and v, respectively associated to j_1 and j_2. The objective function formulated in Eq. 1 remains the same. Likewise, the constraints in Eq. 2, Eq. 3, Eq. 4, and Eq. 5 remain unchanged. However, there are a few key differences with the problem formulated in $\mathcal{P}1$. In $\mathcal{P}2$, we define the attachment vector Y, with $Y_{u,j}$ a boolean such that $Y_{u,j} = 1$ if UE u is attached to S-GW j. We add a constraint stating that a UE is attached to only one S-GW, such that:

$$\sum_{j \in \mathcal{J}} Y_{u,j} = 1 \; ; \;\; \forall u \in \mathcal{U} \tag{11}$$

As the BS and the S-GW are not necessarily co-located, the data path is now as follows: BS j_1, S-GW j_3, S-GW j_4, BS j_2, and vice-versa. This prompts a change in the C_l^d computation from Eq. 6, which becomes:

$$C_l^d = \sum_{f \in \mathcal{F}} 2d_f \Bigg(\sum_{j_1 \in \mathcal{J}} X_{u,j_1} \sum_{j_3 \in \mathcal{J}} Y_{u,j_3} \cdot Z^l_{j_1,j_3} + \sum_{j_4 \in \mathcal{J}} Y_{v,j_4} \Big(\sum_{j_3 \in \mathcal{J}} Y_{u,j_3} \cdot Z^l_{j_3,j_4} + \sum_{j_2 \in \mathcal{J}} X_{v,j_2} \cdot Z^l_{j_4,j_2} \Big) \Bigg) \; ; \;\; \forall l \in \mathcal{L} \tag{12}$$

Following the same reasoning for the signaling paths, Eq. 7, and Eq. 8 remain unchanged from $\mathcal{P}1$. However, Eq. 9 and Eq. 10, computing C_l^{S3} and C_l^{S4} in $\mathcal{P}1$, are replaced by Eq. 13 and Eq. 14, respectively.

$$C_l^{S3} = \sum_{f \in \mathcal{F}} \Big(\sum_{j_3 \in \mathcal{J}} Y_{u,j_3} \sum_{j_0 \in \mathcal{J}} W_{j_0} \cdot Z^l_{j_0,j_3} \cdot S3_f \Big) \; ; \;\; \forall l \in \mathcal{L} \tag{13}$$

$$C_l^{S4} = \sum_{f \in \mathcal{F}} \Big(\sum_{j_4 \in \mathcal{J}} Y_{v,j_4} \sum_{j_0 \in \mathcal{J}} W_{j_0} \cdot Z^l_{j_0,j_4} \cdot S4_f \Big) \; ; \forall l \in \mathcal{L} \tag{14}$$

C. $\mathcal{P}3$: joint user association and attachment per flow optimization

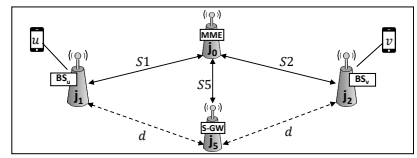

Fig. 3. Data and signaling traffic between BSs for a bidirectional flow in problem $\mathcal{P}3$.

Usually, each UE is attached to a default S-GW, and all flows of that UE are forwarded to that S-GW to handle their routing locally. For each flow, there is signaling traffic between the MME and the default S-GWs of each UE.

In this section, we propose an attachment policy that handles attachment flow by flow. For each flow, both UEs are attached to the same S-GW. The notion of a default S-GW for each UE no longer stands, since a UE can attach to different S-GWs for its different flows. Adding more flexibility to the attachment by allowing UEs to attach to multiple S-GWs at a time, in addition to each flow being handled by only one S-GW, reduces the backhaul consumption caused by signaling.

We formulate $\mathcal{P}3$, the joint user association and attachment per flow optimization problem. Fig. 3 shows the updated data and signaling paths for a bidirectional flow between u and v, respectively associated to j_1 and j_2. The objective function in Eq. 1 remains the same, as well as the constraints in Eq. 2, Eq. 3, Eq. 4, and Eq. 5. The key differences in $\mathcal{P}3$ with respect to $\mathcal{P}1$ and $\mathcal{P}2$, is the fact that each flow must be attached to one S-GW. We define flow attachment vector T, with $T_{f,j}$ a boolean such that $T_{f,j} = 1$ if flow f is attached to S-GW j. A constraint stating that a flow is attached to only one S-GW replaces the constraint in Eq. 11, such that:

$$\sum_{j \in \mathcal{J}} T_{f,j} = 1 \; ; \;\; \forall f \in \mathcal{F} \tag{15}$$

The data path of the bidirectional flow f is as follows: BS j_1, S-GW j_5, BS j_2, and vice-versa. This prompts a change in the computation of C_l^d in Eq. 6, which becomes:

$$C_l^d = \sum_{f \in \mathcal{F}} 2d_f \Big(\sum_{j_5 \in \mathcal{J}} T_{f,j_5} (\sum_{j_1 \in \mathcal{J}} X_{u,j_1} \cdot Z^l_{j_1,j_5} + \sum_{j_2 \in \mathcal{J}} X_{v,j_2} \cdot Z^l_{j_5,j_2}) \Big) \; ; \;\; \forall l \in \mathcal{L} \tag{16}$$

Following the same reasoning for the signaling paths, Eq. 7, and Eq. 8 remain unchanged from $\mathcal{P}1$. However, Eq. 9 and Eq. 10 are replaced by Eq. 17, a single equation, since there is one S-GW per flow. The signaling bit rate between the MME and this S-GW is denoted $S5_f$. In Eq. 17, we compute C_l^{S5} as the total bandwidth consumed by the signaling traffic of all the flows $f \in \mathcal{F}$ on a given link l belonging to the signaling path between MME j_0 and S-GW j_5.

$$C_l^{S5} = \sum_{f \in \mathcal{F}} \Big(\sum_{j_5 \in \mathcal{J}} T_{f,j_5} \sum_{j_0 \in \mathcal{J}} W_{j_0} \cdot Z^l_{j_0,j_5} \cdot S5_f \Big) \; ; \; \forall l \in \mathcal{L} \tag{17}$$

V. Numerical Results

A. Network setup

We base our study on a setup well-suited for PS networks deployed in disaster relief scenarios. We consider a random network topology of 5 BSs randomly distributed in a square area of 1 km^2, as shown in Fig. 4. The network is modeled as a connected random geometric graph, such that a backhaul link exists between 2 BSs if and only if the distance between them is smaller than a given radius $\eta = 0.4$ km. We suppose there are 35 randomly distributed UEs, with intra-cellular flows of average data rate $d = 1$ Mb/s.

As the volume of the signaling traffic and its bit rate are highly dependent on the service in question, and may differ from one particular flow to another, we do not limit our study to pre-defined values of the signaling bit rates. Instead, we evaluate each of our proposals function of different signaling bit rates, in order to emphasize their impact. To that end, we consider the signaling bit rate to be a linear function of the flow data rate, such that $Si_f = \beta \cdot d_f, \; \forall i \in [1, 5]$, with β given as a percentage. We set $K_j = 200$ channels, equally divided between DL and UL, with an FDD ratio $\alpha = 0.5$.

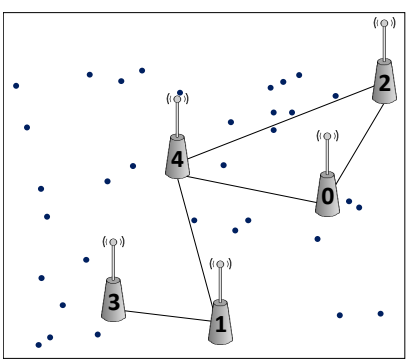

Fig. 4. Network topology.

B. Backhaul bandwidth consumption

We set as benchmark the traditional RAN-based association policy, in which the UE associates to the BS from which it gets the maximum SINR and the attachment follows the association, and compare it to each of the proposed association and attachment schemes in $\mathcal{P}1$, $\mathcal{P}2$, and $\mathcal{P}3$.

Each of the optimization problems is solved using the commercial solver CPLEX, despite the problems complexity. This is possible because of the small sizes of the problems in question, a direct consequence of the limited size of PS networks in terms of number of BSs and UEs. Each solution returns user association, user attachment, and MME placement. As our goal is to reduce the backhaul bandwidth consumption, we use as evaluation metric the relative backhaul bandwidth consumption reduction with respect to the best SINR policy, denoted δ, and expressed as a percentage. If C_{SINR} is the total backhaul bandwidth consumption with a best SINR policy, and $C_{\mathcal{P}i}$ is the total backhaul bandwidth consumption in $\mathcal{P}i$, then:

$$\delta = 100 \cdot \frac{C_{SINR} - C_{\mathcal{P}i}}{C_{SINR}} \tag{18}$$

Fig. 5 shows the relative backhaul bandwidth consumption reduction with respect to the best SINR policy δ (Eq. 18), for $\mathcal{P}1$, $\mathcal{P}2$, and $\mathcal{P}3$. Metric δ is shown function of the signaling bit rate, represented by the percentage β. For $\beta = 0$, there is no signaling traffic. We note that, for the 4 scenarios in question, when the topology shown in Fig. 4 is studied, the optimal position of the MME returned by each of the optimization problems is BS 4.

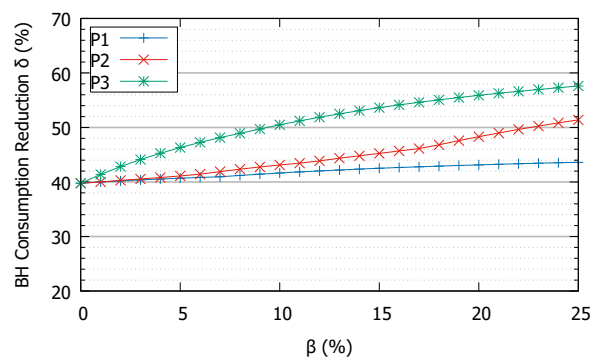

Fig. 5. Relative backhaul bandwidth consumption reduction with respect to the best SINR policy for $\mathcal{P}1$, $\mathcal{P}2$, and $\mathcal{P}3$, function of the signaling traffic.

Let us first evaluate the results of $\mathcal{P}1$. We notice that, for $\beta = 0$, $\mathcal{P}1$ reduces the backhaul bandwidth consumption by 40%. In other words, a gain of 40% in backhaul capacity is achieved when UEs are optimally associated, which is a relatively important gain. If we take the signaling traffic into consideration, we notice that, as the signaling traffic increases (β), the gain does not increase significantly. Indeed, $\mathcal{P}1$ does not optimize attachment, and the results suggest that signaling traffic on the backhaul is not sufficiently reduced when only association is optimized. In other words, the gain achieved by optimizing association in $\mathcal{P}1$ is mainly due to reducing data traffic on the backhaul, and not signaling.

Moving to $\mathcal{P}2$, where attachment is also optimized, it is clear that, for $\beta = 0$, the results are exactly the same as in $\mathcal{P}1$. This is expected because when signaling traffic is not accounted for, the only traffic on the backhaul comes from

data following this path: BS j_1, S-GW j_3, S-GW j_4, BS j_2, and vice-versa (See Fig. 2). In this case, the optimal solution is clearly for each UE to attach to the BS it is associated to, similarly to what happens in $\mathcal{P}1$, meaning no need for j_1 and j_3 (resp. j_2 and j_4) to be different. However, when signaling is accounted for, the previous statement does not hold anymore, since signaling between S-GW and MME is further routed on the backhaul. We notice that, as the signaling traffic increases, the backhaul consumption reduction also increases, achieving higher gains than in $\mathcal{P}1$ once the signaling bit rate is more than 5% of the flow data rate. This proves that, in case the signaling traffic is significant, optimizing attachment further reduces the overall traffic on the backhaul by reducing both signaling and data traffic, leading to improved gains in backhaul capacity.

Finally, we evaluate $\mathcal{P}3$, where we apply an optimized attachment per flow policy, rather than per user. Similarly to $\mathcal{P}1$ and $\mathcal{P}2$, for $\beta = 0$, the backhaul consumption reduction is at 40 % with respect to the best SINR case, proving once more that, when signaling is not accounted for or very low, attachment optimization is not necessary. Nevertheless, the backhaul consumption reduction increases with the increase of signaling traffic. The achieved gains on the backhaul capacity are higher than both $\mathcal{P}1$ and $\mathcal{P}2$ for the same signaling bit rates. For instance, even with signaling rates at $\beta = 1\%$, there is a slight gain on backhaul capacity in comparison with $\mathcal{P}1$. For $\beta = 10\%$, the gain is 50% with respect to a best SINR strategy. Thus, the attachment per flow concept proves itself to be interesting in terms of backhaul bandwidth consumption reduction, notably when signaling traffic is important in the network.

C. Association and attachment distribution

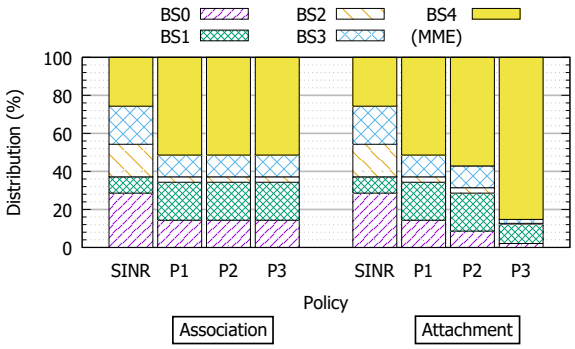

Fig. 6. User association and attachment distribution among the BSs for the different association/attachment policies, for signaling traffic at $\beta = 5\%$.

In Fig. 6, we show the user association distribution, as well as the attachment distribution among the BSs for the best SINR association policy and each of the proposed policies. The shown results correspond to a signaling traffic at $\beta = 5\%$ of the data traffic. We first note that user association distribution is the same in $\mathcal{P}1$, $\mathcal{P}2$ and $\mathcal{P}3$. This is expected, as the difference between these policies mainly lies in the attachment criteria and not the association. In comparison with the best SINR policy, we notice that more users are associated to BS 4 than any other BS. This is one of the reasons of the achieved backhaul bandwidth consumption reduction observed in Fig.5, since BS 4 is the BS co-located with the MME, as returned by the optimal solution. For the user attachment, the observation for $\mathcal{P}1$ is redundant with user association, since there is no difference between association and attachment in this case. However, it is clear that when attachment is optimized in $\mathcal{P}2$, more users are attached to the S-GW of BS 4, which is also co-located with the MME. Similarly, this is the case in $\mathcal{P}3$, where the majority of the flows are also attached to BS4.

These observations indicate that, for most UEs, it is more beneficial to associate and/or attach to the BS co-located with the MME, if this is an option, as long as that BS can handle all the users. While this could raise questions on the achieved load balancing, we remind that RAN constraints are also accounted for, meaning no BS can take more users than allowed by its capacity.

VI. Conclusion

In conclusion, we revisited in this paper the user association problem, in the context of PS rapidly deployable networks. We proposed an optimal user association that minimizes the bandwidth consumption on the backhaul, while respecting the stringent user guaranteed bit rate requirements, and accounting for RAN constraints. This association significantly reduces traffic on the backhaul in comparison with a traditional RAN-based association, achieving 40% less backhaul bandwidth consumption. Moreover, we showed that jointly optimizing user association and user attachment further reduces bandwidth consumption on the backhaul when signaling traffic is significant. While this study is limited to the offline problem, it can serve as a basis for more practical online user association schemes, in which a backhaul-aware optimal association decision is made in real-time for each user.

References

[1] B. S. Manoj, A.H. Baker, "Communication Challenges in Emergency Response", *Communications of the ACM*, 50(3), 2007.

[2] Federal Communications Comission, "Communications Status Report for Areas Impacted by Tropical Storm Harvey", Aug. 2017.

[3] J. Oueis ,V. Conan, D. Lavaux, R. Stanica, F. Valois, "Overview on LTE Isolated E-UTRAN Operation for Public Safety", *IEEE Communications Standards Magazine*, 1(2), July 2017.

[4] 3GPP TR 23.797, "Study on architecture enhancements to support isolated Evolved Universal Terrestrial Radio Access Network operation for public safety (Release 13)", June 2015.

[5] Google Loon project, https://x.company/loon/

[6] I. Bor-Yaliniz, H. Yanikomeroglu, "The New Frontier in RAN Heterogeneity: Multi-Tier Drone-Cells", *IEEE Communications Magazine*, 54(11), Nov. 2016.

[7] Air Lynx, https://www.air-lynx.com/

[8] K. Gomez, L. Goratti, T. Rasheed, L. Reynaud, "Enabling Disaster-Resilient 4G Mobile Communication Networks", *IEEE Communications Magazine*, 52(12), Dec. 2014.

[9] D. Liu, L. Wang, Y. Chen, M. Elkashlan, K.Wong, R. Schober, L. Hanzo, "User Association in 5G Networks: A Survey and an Outlook", *IEEE Communications Surveys & Tutorials*, 18(2), Apr. 2016.

[10] H.Kim, G.Veciana, X.Yang, M.Venkatachalam, "Distributed α-optimal User Association and Cell Load Balancing in Wireless Networks", *IEEE/ACM Transactions on Networking*, 20(1), Feb. 2012.

[11] H. Boostanimehr, V. Bhargava, "Joint Downlink and Uplink Aware Cell Association in HetNets with QoS Provisioning", *IEEE Transactions on Wireless Communications*, 14(10), Oct. 2015.

[12] N. Sapountzis, T. Spyropoulos, N. Nikaein, U. Salim, "Optimal Downlink and Uplink User Association in Backhaul-limited HetNets", *IEEE INFOCOM*, San Francisco, CA, USA, Apr. 2016.

[13] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C., Caramanis, J. Andrews, ' 'User Association for Load Balancing in Heterogeneous Cellular Networks", *IEEE Transactions on Wireless Communications*, 12(6), June 2016.

[14] A. Fehske, H. Klessig, J. Voigt, G. Fettweis, "Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-organizing Radio Networks", *IEEE Transactions on Vehicular Technology*, 62(5), June 2013.

[15] B.Aoun, R.Boutaba, Y.Iraqi, G.Kenward, "Gateway Placement Optimization in Wireless Mesh Networks with QoS Constraints", *IEEE Journal on Selected Areas in Communications*, 24(11), Nov. 2006.

[16] D. Fooladivanda, C. Rosenberg, "Joint Resource Allocation and User Association for Heterogeneous Wireless Cellular Networks", *IEEE Transactions on Wireless Communications*, 12(1), Jan. 2013.

[17] J. Oueis, V. Conan, D. Lavaux, R. Stanica, F. Valois, "Core Network Function Placement in Mobile Networks", *IEEE PIMRC*, Montreal, QC, Canada, Oct. 2017.