On Selective Activation in Dense Femtocell Networks

Michael Lin, Simone Silvestri, Novella Bartolini, and Thomas F. La Porta

***Abstract*— Over-provisioned femtocell networks can be used to serve indoor locations that see high peak loads, such as airports or train stations. However, networks designed for high peak loads are mostly under-utilized, which is wasteful from an energy-use perspective. This paper introduces a femtocell *selective activation* problem. We motivate the use of selective activation in femtocell networks using real femtocell power measurements. We formally define the selective activation problem, and introduce GREENFEMTO, a distributed femtocell selective activation algorithm. We prove that GREENFEMTO converges to a locally Pareto optimal solution. Detailed simulations of an LTE wireless system are used to demonstrate the performance of GREENFEMTO. We find that GREENFEMTO uses up to 55% fewer femtocells to serve a given load, relative to an existing femtocell power-saving technique. Furthermore, we show that GREENFEMTO comes within 15% of a globally optimal solution. We conclude that selective activation can be successfully applied to femtocell networks to both reduce power consumption, and reduce outage probabilities.**

***Index Terms*— Communication technologies, wireless communications, wireless networks, Mathematics, algorithms, distributed algorithms.**

I. INTRODUCTION

THE use of mobile broadband wireless devices is experiencing explosive growth. According to recent studies [1], global mobile broadband subscriptions will reach 4.4 billion by the end of 2016, and global mobile data traffic has been growing at a rate of 150% yearly. This increased use has led to a higher reliance on these devices for indoor and outdoor high speed wireless data access. This has resulted in an increase in the occurrence of temporary, high-density concentrations of mobile users with high traffic demands [2], often indoors. Such density increases can occur intermittently, for example, during sporting events or concerts, or more regularly, such as when trains arrive at high-traffic stations or at an airport gate. These increases temporarily place a very high, localized load on wireless networks, reducing service quality and leading to outages.

Provisioning wireless networks to manage the peak number of users in a congested area can reduce the occurrence of such problems, but is expensive and wasteful during normal loads. Temporary, mobile wireless base stations [3] are not suitable for areas which may see regularly occurring congestion, such as subway stations or airports. For these reasons, there has been significant effort from the industry [2], [4] and the research community [5]–[14] to design solutions based on the new technology of *femtocell networks*.

Femtocells are small, short-range wireless base stations that rely on user-provided Internet for backhaul. Provisioning a short-range small cell network for peak hotspot load is easier and cheaper than provisioning a long-range macrocell network for the same load. This makes small cells well-suited for managing indoor hotspots in locations that see periodic traffic, such as public transportation hubs, meeting halls, and classrooms.

In order to be able to successfully handle the high load generated by user hotspots, the femtocell network needs to be deployed at high density. Such dense deployment may result in a non-negligible network energy consumption. For example, assuming femtocells consume 6 W of power, a network of 1000 cells covering a university campus would consume more than 50 MWh/year. Solutions for designing high-density wireless networks mainly focus on user quality of service, and only marginally consider power consumption [2], [5]–[7]. However, the temporal nature of user hotspots can be exploited to reduce the network energy expenditure. As an example, consider a small cell network that is provisioned to serve a peak load that is three times the average load. If the peak load is encountered 30% of the time, nearly 50% of the power used by the network is unnecessary.

In this paper, we study the problem of minimizing the energy consumption of dense small cell networks. Use cases include university campuses, sports arenas, or public transportation hubs. In these scenarios, user traffic exhibits time- and space-localized peaks. To reduce network energy consumption, we model the problem as a selective activation problem. Using selective activation, only a reduced set of femtocells are activated at a given time. The remaining cells are put to sleep to conserve power and are activated only when needed as traffic increases. We motivate the use of selective activation, rather than power-management, through measurements of real femtocell energy usage, which show that femtocell energy use is not dependent on load. Therefore, reducing femtocell energy usage is directly translated to reducing the number of active femtocells.

We formalize the selective activation problem as a non-linear optimization problem, which takes into account detailed aspects of LTE networks such as interference and resource block allocation. We show that the problem is NP-Hard and introduce GREENFEMTO, a distributed algorithm for femtocell

Manuscript received October 20, 2015; revised March 17, 2016 and June 19, 2016; accepted July 16, 2016. Date of publication July 27, 2016; date of current version October 7, 2016. The associate editor coordinating the review of this paper and approving it for publication was M. Li.

M. Lin and T. F. La Porta are with the Department of Computer Science and Engineering, Pennsylvania State University, University Park, PA 16802 USA (e-mail: molin@cse.psu.edu; tlp@cse.psu.edu).

S. Silvestri is with the Department of Computer Science, Missouri University of Science and Technology, Rolla, MO 65409 USA (e-mail: silvestris@mst.edu).

N. Bartolini is with the Department of Computer Science, Sapienza University of Rome, Rome 00185, Italy (e-mail: bartolini@di.uniroma1.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2016.2594784

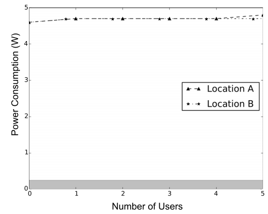


Fig. 1. Femtocell power consumption in W, with 0 to 5 users. The shaded grey area at the bottom indicates the maximum transmission power of the femtocell radio.

selective activation. The algorithm does not assume cell synchronization and dynamically adapts the active set of cells on the basis of the user location and movement. We prove that GreenFemto converges to a stable solution and that the solution is Pareto optimal.

We evaluate the performance of GreenFemto against a previously proposed femtocell power-saving algorithm using a realistic LTE simulator. Results show that GreenFemto outperforms the previous approach and reduces the energy consumption of the network by up to 55%, and comes within 15% of a globally optimal solution.

A preliminary version of this paper appeared in [15].

In summary, our contributions are the following:

- We perform real measurements of the energy consumption of femtocells to motivate the use of selective activation;
- We study the problem of reducing the energy consumption of densely deployed femtocell networks. We formulate this as a non-linear optimization problem and show that it is NP-Hard;
- We introduce GreenFemto, a distributed algorithm for femtocell selective activation. We prove that GreenFemto converges to a Pareto optimal solution;
- We compare the performance of GreenFemto through simulations, showing that it successfully reduces energy consumption and outperforms previous solutions.

II. Femtocell Power Measurements

In this section we analyze the power usage of commercially available femtocells to motivate the use of selective activation. We find that femtocells consume a fixed amount of power, independent of the number and position of the users they are serving, which corroborates results presented in [16]. Although femtocell radio transmission power varies with the number and location of users being served, femtocell radios have a maximum transmit power in the range of 100-250 mW [16], leading to low variation in power consumption that is dominated by fixed-power components such as the processor, amplifier, and memory. Since femtocell power use remains almost constant with respect to load and user position, the best way to control power consumption is to limit the number of active femtocells, regardless of the load on the network. Therefore, we conclude that selective activation is the most effective way of reducing a femtocell network's energy consumption.

We record the femtocell's power usage with zero to five active users at two locations: location A, 0.5 meters away in the same room, with an average received signal strength indicator of −59 dBm, and location B, at the edge of the femtocell's coverage radius, with an average received signal strength indicator of −91 dBm, approximately 15 meters away, through multiple interior walls. The recorded values are shown in Figure 1, where each point is the average of five measurements. The shaded grey region at the bottom of Figure 1 indicates the maximum transmission power of the femtocell radio. As femtocell load increases, the femtocell's power usage remains stable. Based on the lack of correlation between the femtocell's transmit power and its overall power consumption, and taking power supply inefficiency into account, we conclude that femtocells consume a fixed amount of power when activated, regardless of the number of users being served.

Since a femtocell's power consumption arises largely from being powered on, we focus our design at the broad level of minimizing the number of femtocells that are active, rather than managing the power consumption of individual femtocells or users.

III. Network Model and Problem Formulation

In this section we define our network model and problem. We consider an LTE data network with femtocells. The channel bandwidth is divided into subchannels, each of which are further time-divided [17]. Each time division on a subchannel is known as a *resource block* (RB). User transmissions are scheduled across RBs on the base station to which the user is associated. Base stations regularly transmit reference signals (RS) that are used to measure signal quality.

Consider an area of interest $\mathcal{L}$, defined as a set of points $p \in \mathcal{L}$. We define a set $F = \{f_1, \ldots, f_N\}$ of N femtocells deployed over $\mathcal{L}$. We consider two types of coverage: *area coverage* and *user coverage*. Area coverage is defined as coverage of a set of points regardless of users, and is used to detect users. User coverage is defined as the ability to serve a user at a minimum transmission rate, and is used for user service. We define two distinct thresholds for area and user coverage: t^a and t^u. A location or user is *covered* if the SINR of the RS at that point is above t^a or t^u, respectively.

A femtocell f_i's *coverage range*, R_i^a or R_i^u, is defined as the largest contiguous set of locations at which f_i's RS can be received with SINR above t^a or t^u, respectively. As $t^u > t^a$, it holds $||R_i^a|| > ||R_i^u||$. We assume each femtocell is aware of its location and coverage range.[1] A location is redundantly covered if it is covered by at least two femtocells. Femtocells communicate locally using the network that they use for backhaul.

Users are defined as a set $U = \{u_1, \ldots, u_M\}$. Users are served by femtocells that have enough free RBs to schedule the user for transmission. A user u is rejected if no femtocell can schedule her. To emphasize the effect of user rejections, we formalize our algorithms in a network model

[1]In a centrally managed and deployed network, femtocell coverage ranges can be estimated at deployment.

without macrocells, however, this is not a requirement for our algorithms. Nevertheless, since femtocell networks are often deployed to improve signal quality in areas with poor or absent macrocell coverage, such as subways, this is a reasonable assumption.

The selective activation problem is to find a minimal set of *active femtocells*, $F^* \subseteq F$, that provides full area and user coverage. This implies that all locations in $\mathcal{L}$ have an SINR above t^a, and all users have an SINR above t^u. The assignment of users in U to femtocells in F^* must be feasible in terms of femtocell capacity and must satisfy user requirements in terms of minimum acceptable rate. If not all users can be served at their rate targets, users are rejected one by one.

IV. Interference-Aware Femtocell Activation Optimization

In this section we give a theoretical formulation of the problem, described in Section III, to determine a subset of the available femtocells, so as to serve a given set of users at a minimum transmission and provide area coverage while minimizing energy consumption.

An *activation policy* determines which of the N available femtocells should be activated to serve the given set of M users and can be expressed by means of an *activation vector* $\bar{X} \triangleq \{x_1, x_2, \ldots, x_N\}$, where the value of the variable x_i is 1 if the femtocell f_i is activated, while it is 0 if f_i is disabled.

The objective of our problem is to find an activation policy that serves all users in U with the minimum number of active femtocells; that is, an activation vector $\bar{X}$ that minimizes the value of $\sum_{i=1}^{N} x_i$ which is the objective function of the hereby called **MinActivation** problem of Equation (1), that we describe in the following.[2]

$$
\begin{aligned}
&\min \sum_{i=1}^{N} x_i\\
&\text{subject to: } \sum_{j=1}^{M} y_{ij} \le x_i \cdot b_{\mathtt{MAX}} \quad \forall i \text{ s.t. } f_i \in F && (a)\\
&\qquad \sum_{i=1}^{N} z_{ij} = 1 \quad \forall j \text{ s.t. } u_j \in U && (b)\\
&\qquad y_{ij} \ge z_{ij} \cdot b_{ij}(\bar{X}) \quad \forall i, j \text{ s.t. } f_i \in F, u_j \in U && (c)\\
&\qquad y_{ij} \le z_{ij} \cdot b_{\mathtt{MAX}} \quad \forall i, j \text{ s.t. } f_i \in F, u_j \in U && (d)\\
&\qquad \sum_{i=1}^{N} x_i \cdot p_{il} \ge 1 \quad \forall l \in \mathcal{I} && (e)\\
&\qquad x_i, z_{ij} \in \{0, 1\};\ y_{ij} \in \mathbb{N} \quad \forall i, j \text{ s.t. } f_i \in F, u_j \in U
\end{aligned} \tag{1}
$$

To formulate the constraints of the **MinActivation** problem we introduce the binary variables $z_{ij} \in \{0, 1\}$, whose value denotes whether cell $f_i \in F$ serves the user $u_j \in U$, in which case $z_{ij} = 1$, while $z_{ij} = 0$ otherwise. We also denote with $y_{ij} \in \mathbb{N}$ the number of resource blocks provided by cell f_i to user u_j.

Femtocells have a capacity limitation, as the amount of resource blocks available for each cell is limited to $b_{\mathtt{MAX}}$. We introduce the constraint of Equation 1(a), which states that each cell providing resource blocks to any user must be active, and that each active cell can assign up to $b_{\mathtt{MAX}}$ users.

The variable z_{ij} is used in Equation 1 (b) to let a user be served by one and only one cell.

Given a user u_j, being served by the cell f_i, and given the activation vector of the network $\bar{X}$, the rate requirement r for the user u_j can be defined as a function $b_{ij}(\cdot)$, that returns the minimum number of resource blocks $b_{ij}(\bar{X})$ that user u_j needs from cell f_i to operate with rate higher than or equal to r, given that user j's SINR is above t^u. The function b_{ij} depends on system-wide SINR calculations, therefore it is non-linear and there is no known closed form expression in terms of the elements of $\bar{X}$. Therefore, the user quality requirement implies that $y_{ij} \ge b_{ij}(\bar{X})$ when the cell f_i is selected to serve user u_j, which can be expressed with a unique constraint as in Equation 1(c).

Equation 1(d) determines an upper bound on the number of resource blocks that a cell f_i can allocate to a user u_j, which must be zero if the cell f_i is not the one that serves user u_j ($z_{ij} = 0$). If, otherwise, the cell f_i is the one assigned to user u_j ($z_{ij} = 1$), such a constraint is a trivial upper bound, as the constraint of Equation 1(a) already imposes that the number of resource blocks provided by the unique cell serving a user is lower than or equal to $b_{\mathtt{MAX}}$.

Finally, in order to let cells detect users in their coverage range and enable activation decisions, we impose an additional constraint on area coverage. Each cell f_i is able to detect the presence of users in their area coverage range, R_i^a. Note that the area coverage range is larger than the user coverage range, R_i^u, and cell f_i may be unable to serve a user at distance R_i^a. Nevertheless, if every point of the area of interest is within the coverage range of an active cell, the network is able to sense any incoming users and activate additional femtocells as needed.

We model coverage ranges as open intervals, that is, femtocell f_i does not cover the border of its area R_i^a. We define an *intersection point* as a point at which the borders of two coverage ranges intersect, or a point at which a coverage range intersects with the border of the area $\mathcal{L}$. We denote by $\mathcal{I}$ the set of all intersection points in $\mathcal{L}$. In the following theorem, we extend a result from [18] that shows that the coverage of the points in $\mathcal{I}$ implies complete coverage of $\mathcal{L}$ to the case of general non-convex shaped coverage areas.

Theorem 1: Given a set of femtocells F, deployed over an area $\mathcal{L}$ and generating a set of intersection points $\mathcal{I} \subset \mathcal{L}$, $\mathcal{I} \neq \emptyset$, if a subset $\hat{F} \subseteq F$ covers all points in $\mathcal{I}$, then $\hat{F}$ completely covers $\mathcal{L}$.

Proof: The set of coverage ranges partitions the area $\mathcal{L}$ into disjoint coverage patches $H_1, \ldots, H_m$, where patches are bounded by the borders of coverage ranges or by the borders of $\mathcal{L}$.

[2]Notice that, in order to keep the formulation simple, we formulate the constraints of the problem under the assumption that there are enough femtocells to serve all the users. As we use the optimal solution of this problem as a reference for performance comparisons, in the experiments we do not assume this and we let the network drop user requests.

The border of a coverage patch H_i contains intersection points everywhere the border of two coverage ranges, or the border of a coverage range and the border of $\mathcal{L}$, intersect.

We proceed by contradiction. Let $p \in \mathcal{L}$ be a point not covered by the femtocells in $\hat{F}$ and let H_p be the coverage patch to which it belongs. H_p always exists since the coverage patches partition $\mathcal{L}$.

All points in H_p have the same coverage degree, by definition of coverage patch, thus H_p is also not covered. Since we define coverage ranges to be open intervals, an intersection point on the border of H_p, generated by the intersection of the borders of two coverage ranges R_i and R_j, is not covered by the femtocell f_i and f_j. As a result, the intersection points on the border of H_p are also uncovered.

This is a contradiction of the assumption that all intersection points are covered by the cells in $\hat{F}$, therefore the uncovered point p does not exist. ■

$\mathcal{I}$ includes all the intersection points generated by the coverage regions of all the femtocells in F. According to Theorem IV, if all points of $\mathcal{I}$ are within the coverage range of at least an active cell, all of $\mathcal{L}$ is covered, and the active cells are able to detect the presence of a user in any point of the area.

Therefore, we consider formulate the area coverage constraint as follows. Let $l \in \mathcal{I}$ be a point in the region of interest. Let p_{il} be a binary constant coefficient which is equal to 1 if the point l is in the coverage range of femtocell f_i. Then, in order to have complete area coverage, the activation set $\bar{X}$ must satisfy Equation 1(e), which requires that each intersection point of $\mathcal{I}$ be in the coverage range of at least one active cell.

The following theorem proves that **MinActivation** is NP-Hard by showing that Set Cover is a special case of **MinActivation**.

*Theorem 2: **MinActivation** is NP-Hard.*

Proof: We show that Set Cover is a sub-problem of our problem. Let us consider a general instance of Set Cover. We define Ω as the set of all elements, and $S = \{S_1, \ldots, S_n\}$ is a family of n subsets $S_i \subseteq \Omega$ for $i = 1, \ldots, n$ and $\cup_{i=1,\ldots,n} S_i = \Omega$. According to Set Cover we should find the smallest subset of S such that all elements in Ω are covered.

We can create an instance of **MinActivation** from a general instance of Set Cover by the following procedure:

1) Add a point in $\mathcal{I}$ for each element of the universe Ω.
2) Add a femtocell f_i in F for each set S_i in S. f_i can cover the points in $\mathcal{I}$ corresponding to the elements in S_i.

No users are needed in the reduction, so $U = \emptyset$.

A solution to **MinActivation** finds an activation vector $\bar{X}$ that identifies a minimum set of active femtocells needed to cover all points in $\mathcal{I}$. This corresponds to the smallest subset of S that covers all elements in Ω. Therefore, Set Cover is a special case of **MinActivation** and thus our problem is NP-Hard. ■

V. THE GREENFEMTO ALGORITHM

In this section we introduce GREENFEMTO, a distributed algorithm that finds a locally Pareto optimal solution to **MinActivation**. The algorithm consists of two parts that address two different aspects of user management: user detection and user reassignment.

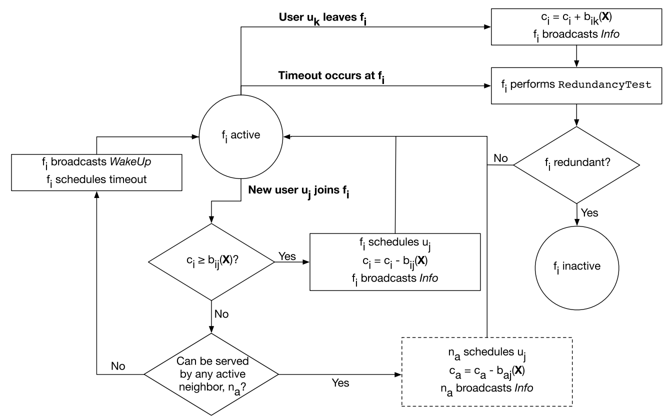


Fig. 2. Femtocell state diagram for user events.

A. Overview

A femtocell status can be either *active* or *inactive*.

Femtocells can be kept active either because they are needed to serve users, or because they are needed for area coverage. Active femtocells provide area coverage by sending beacon messages to announce their presence to the users. Therefore the energy consumption of active cells remains important even when there are no users.

Inactive femtocells do not provide coverage or serve users, but do not consume energy. We assume that femtocells can be turned on remotely using Wake-on-LAN [19]. Femtocells in inactive mode can be contacted by neighbor cells to be woken up upon need and can therefore safely put to sleep their radio for the whole duration of the inactive state.

For each femtocell $f_i \in F$ we define $U_i \subseteq U$, the set of users f_i can serve. We say that a femtocell f_j is a *neighbor* of a femtocell f_i, if their coverage ranges overlap, that is $R_i^u \cap R_j^u \neq \emptyset$. We refer to the set of all neighbors, active or inactive, of f_i as B_i. The set of active neighbors is denoted by B_i^*, and inactive neighbors by B_i^o. c_i is the residual capacity of the femtocell f_i, that is, the number of free RBs it can schedule. We assume that each femtocell f_i knows the coverage ranges, R_j^u and R_j^a, of each of its neighbors.

GREENFEMTO keeps a minimal set of *sentinel* femtocells active for area coverage. This set is found by disabling redundant femtocells, as identified using the procedure in Section V-D. Users are detected by sentinel femtocells, and either served directly by the sentinel femtocell or served by a newly awoken femtocell. After femtocells are awoken, *user reassignment* and a *redundancy test* are performed to minimize the number of active femtocells. These procedures are described in detail in the following sections.

The overall flow of the algorithm can be seen in the femtocell state diagrams in Figures 2 and 3. Figure 2 shows the actions taken by femtocells when users join or leave the network, or when timeouts occur. As evidenced by this figure the algorithm adapts the current solution with local decisions when new users join or leave. As the algorithm is adaptive, it is not necessary to re-run the algorithm multiple times to follow the user dynamics. Figure 3 shows the actions taken when femtocells receive messages, as when an inactive cell is activated,

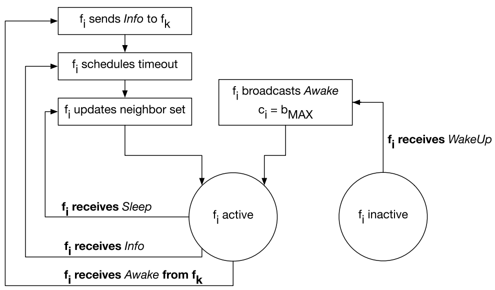


Fig. 3. Femtocell state diagram for message events.

or when an active cell receives other event related messages from neighboring cells.

B. User Detection and Femtocell Activation

When a new user u_j joins the network, at least one femtocell detects the user, since femtocells provide area coverage. This femtocell is referred to as the *sentinel femtocell*, f_s. Furthermore, we assume that u_j notifies f_s of the active set of cells $F^{u_j} = \{f_1^{u_j}, \ldots, f_k^{u_j}\}$ that can serve it. Based on our model assumptions, the cells in F^{u_j} are neighbors of each other.

If f_s is able to serve the user, that is, $c_s \geq b_{sj}(\bar{X})$, it decreases its residual capacity c_s and broadcasts an `Info` message with its new capacity. If u_j cannot be served, f_s checks whether there are currently active cells that can serve u_j, i.e. if $\exists f_i \in F^{u_j}$ s.t. $c_i \geq b_{ij}(\bar{X})$. In this case, f_s does not serve the user and the user selects another femtocell to attach to from F^{u_j}. Otherwise, if all the cells in F^{u_j} have insufficient residual capacity, f_s broadcasts a `WakeUp` message to its inactive neighbors in B_s^o. Also in this case, f_s does not serve u_j, and u_j will try to attach to the newly activated cells.

Inactive neighbors that receive a `WakeUp` message become active, set their capacities to c^{max}, and broadcast an `Awake` message. Neighboring cells that receive the `Awake` message reply with an `Info` message with their capacity to inform the newly activated cells of the current state of the network.

Cells receiving `Awake` messages schedule a timeout that triggers a redundancy test. As explained in Section V-C and V-D, this test determines if a femtocell can be turned off. In order to prevent conflicting decisions of multiple cells concurrently performing the test, timeouts are scheduled in a random time instant in the interval $(0, \tau_{max}]$.[3] As the newly awakened cells receive `Info` messages, they also schedule timeouts. If a new user is detected by an awakened femtocell f_a during its timeout interval, the user detection algorithm runs as normal, but the timeout for f_a is *not* reset. When the timeout occurs, the scheduled user reassignment takes place.

A pseudocode description of the algorithm, as executed in the sentinel femtocell, is in Algorithm 1. The algorithm has complexity $O(|B_s|)$, where we recall B_s is the set of neighboring femtocell of f_s.

[3]Deterministic approaches could also be employed, which may however introduce additional overhead and delays. The efficient techniques proposed to determine cell redundancy make such a probabilistic approach a viable and effective solution.

Algorithm 1 GreenFemto User Detection

Input: a new user, u_j, attaching to the network
$\bar{X}$, the current femtocell activation vector

```
user u_j is detected by sentinel femtocell f_s;
if c_s ≥ b_sj(X̄) then
    c_s −= b_sj(X̄);
    Serve user u_j;
    return
F^{u_j} := user u_j's active cell list;
if ∃ f_i ∈ F^{u_j} s.t. c_i ≥ b_ij(X̄) then
    Reject u_j;
    /* u_j will attach to a cell in F^{u_j}
       with enough capacity                    */
    return
/* at this point, none of the currently
   active femtocells can serve u_j             */
B_s^o := femtocell f_s's neighbors;
wake up neighbors in B_s^o;
/* u_j will attach to a cell in B_s^o with
   enough capacity, if available               */
schedule redundancy test at a random instant in [0, τ_max];
```

C. User Reassignment

When a timeout occurs for a femtocell f_i, it determines whether it can become inactive by performing a *redundancy test*, described in detail in the next section. A femtocell f_i is redundant if its currently active neighbors B_i^* can both cover f_i's area coverage range, R_i^a, and serve all of f_i's users, U_i. The test begins with each user $u_i \in U_i$ sending f_i the active set of femtocells to which it can connect. If every user responds with at least two femtocells, f_i can determine whether there is enough residual capacity in its neighbors to support a reassignment.

A *user reassignment* is a function $\mathcal{R} : U_i \rightarrow B_i^*$ such that for each user $u \in U_i$, $\mathcal{R}(u)$ is the neighboring cell of f_i to which u_j is reassigned. In Section V-D we present a method for performing a redundancy test and determine a reassignment. If f_i is redundant, it sends a `Sleep` message, which notifies its neighbors of the reassignment and reserves space on the new cell for incoming users, transfers its users to its neighbors according to the reassignment, and finally turns itself off.

When a femtocell f_j receives a `Sleep` message and reassignment $\mathcal{R}(\cdot)$, it updates its stored state of f_i and reserves resources for incoming users according to $\mathcal{R}(\cdot)$. Furthermore, to prevent race conditions, if a timeout for a redundancy test for f_j was active, it is reset. Femtocell f_i then initiates a handoff for the users it is transferring to f_j. This ensures that users will not be handed off to femtocells that no longer have sufficient capacity. The reassignment handoff must occur within a time, T_R, after which users are accepted according to the standard method.[4]

[4]The timeout is introduced to take into account possible changes to the network, such as user movement, which may prevent the application of the reassignment. However, since the reassignment can be performed at a small time scale relative to user movement, these changes are unlikely.

Algorithm 2 The RedundancyTest Algorithm

Input: femtocell f_i on which to perform the redundancy test
Result: $\mathcal{T}_f$:= true if f_i is redundant, false otherwise

begin coverage test
  $\mathcal{T}_f^c$:= true;
  B_i^* := femtocell f_i's active neighbors;
  $\mathcal{I}_i$:= femtocell f_i's intersection points with neighbors $f_n \in B_i^*$;
  foreach $p_i \in \mathcal{I}_i$ **do**
    if *p_i not covered by a femto $f_n \in B_i^*$* **then**
      $\mathcal{T}_f^c$:= false;
begin reassignment test
  $\mathcal{T}_f^r$:= true;
  perform generalized assignment of users U_i to neighboring femtocells B_i^*;
  if *no feasible solution* **then**
    $\mathcal{T}_f^r$:= false;
return $\mathcal{T}_f \rightarrow \mathcal{T}_f^c \wedge \mathcal{T}_f^r$

D. Redundancy Test

The full redundancy test algorithm is listed in Algorithm 2. Redundancy tests take place when timeouts occur in newly awakened femtocells. A femtocell f_i is redundant if it can be turned off without creating an area coverage hole, and if all the users served by f_i can be reassigned to the other active femtocells in F^*. In the following section we describe how these tests can be performed.

1) Coverage Test: Theorem 1 can be used to determine whether a femtocell is redundant in terms of coverage, as stated by the following corollary.

Corollary 3: A femtocell f_i, with area coverage range R_i^a, is redundant in terms of coverage if the intersection points $\mathcal{I}_i$ in R_i^a are covered by its set of active neighbors B_i^*.

Note that this test can be performed locally provided that femtocells are aware of the intersection points that lie in their coverage range, as we assume in this paper. The test can be performed in $O(|\mathcal{I}_i| \times |B_i^*|)$, since every intersection point needs to be covered by at least an active neighbor of f_i.

2) User Reassignment: User reassignment can be modeled as a generalized assignment problem [20] with unit costs, where all the users $u_j \in U_i$ on f_i, with respective RB requirements $b_{ij}(\bar{X})\ \forall u_j \in U_i$, must be assigned to $|B_i^*|$ neighboring femtocells, with respective residual capacities $c_k, \forall f_k \in B_i^*$. This is an NP-complete problem, therefore there is no exact polynomial-time algorithm. However, there are both efficient, bounded approximation algorithms, and exact algorithms [20] that perform well on small problem sizes. In our experiments, the number of active neighbors $|B_i^*|$ is always less than ten, and the problem can be solved exactly.

E. User Departures and Handoffs

When a user u_j served by femtocell f_i leaves the network, f_i increases its capacity and broadcasts an `Info` message to its neighbors, to alert them of this change. If f_i is not serving any users and is not needed for area coverage, it broadcasts a `Sleep` message and enters the inactive state.

If a mobile user is leaving the coverage area of its attached femtocell, as detected by decreasing signal quality reports, and there are no active femtocells with available capacity to which the user can be handed off,[5] the current femtocell broadcasts a `WakeUp` message to its sleeping neighbors. This awakens those neighbors, and allows them to be the target of a handoff from the current femtocell. Specifically, the `WakeUp` message is triggered when the current user's SINR drops below a threshold $t^{handoff}$. $t^{handoff}$ is larger than t^u, so a handoff `WakeUp` is triggered before the user loses coverage.

VI. ALGORITHM PROPERTIES

In this section we prove that GREENFEMTO converges. We assume that GREENFEMTO is run with an exact generalized assignment algorithm.

GREENFEMTO reconfigures the network as long as users move, join or leave the network. In the following we prove that the algorithm converges to a stable configuration after the last dynamic event, provided that no new event occurs during the execution of the algorithm.

Recall that an *activation policy* state is a vector $\bar{X} = \{x_1, \ldots, x_N\}$, where $x_i = 1$ if f_i is active and 0 if it is inactive. In a network with n femtocells and m users, the space of possible activation vectors $\mathbb{X}$ is finite and its size is bounded by $|\mathbb{X}| < 2^N$.

Theorem 4: The algorithm GREENFEMTO *converges to a stable configuration.*

Proof: We define the function $g : \mathbb{X} \rightarrow \mathbb{N}^+$, as $g(S) = \sum_{i=1}^{N} x_i$. The function g is trivially lower bounded by 0.

Now consider the three possible network events: a user joining, leaving, or moving. Events can occur in any order, but we assume that events themselves are atomic. If a femtocell is scheduled for a redundancy test, any further redundancy test timeouts at that femtocell are ignored, therefore events will be sequentialized by the length of each femtocell's initial non-overlapping timeout. Since any combination of events leads to a sequence of redundancy tests, it suffices to show that all events lead to redundancy tests, and redundancy tests cannot change a state S to a state S' such that $g(S) < g(S')$, and if $g(S) = g(S')$, $S = S'$.

Consider a user u_j joining the network. If u_j can be served by the sentinel femtocell f_s, or another currently active femtocell, the algorithm halts and no changes to the network state are made. If u_j cannot be served by any active femtocell, f_s wakes up its inactive neighbors B_s^o, and $g(S)$ increases by $|B_s^o|$. Each femtocell $f_k \in B_s^o$ schedules a timeout for a redundancy test.

Consider a user u_j leaving the network. When u_j leaves the network, one of two cases are possible; redundancy tests are performed in both. If u_j is not the last user on a femtocell f_i, f_i broadcasts an `Info` message. If u_j is the last user on a femtocell f_i, f_i broadcasts a `Sleep` message and disables itself.

[5]Recall that each femtocell is aware of the residual capacity of its active neighbors.

Neighboring femtocells receiving either message start a timeout for a redundancy test.

Consider a user u_j moving from femtocell f_i to femtocell f_k. The actions of f_i are identical to when a user leaves the network. If f_k has sufficient capacity to serve u_j, u_j is handed over to f_k, and no more state changes occur. If f_k does not have sufficient capacity to serve u_j, neighboring femtocells B_k^o are woken up. The remainder of the proof is identical to when a user joins the network.

Finally, consider a femtocell f_i performing a redundancy test, with initial state S, and final state S'. f_i checks if it is needed for satisfaction of the area coverage constraint and if its users can be reassigned to neighboring femtocells. If one of the two tests fails, f_i remains active, and $g(S) = g(S')$ and $S' = S$. If the tests are successful, f_i is disabled and the new state S' is such that $g(S) > g(S')$. Since the network state space is finite and the function g is lower bounded, GREENFEMTO eventually converges to a stable state where no more state transitions are possible. ■

We now prove that the final configuration to which GREENFEMTO converges is locally Pareto-optimal, defined as follows:

Definition 5 (Local Pareto-Optimality): A network state S is locally Pareto-optimal if any of the two following conditions holds:

1) *For all active femtocells $f_i \in F^*$, deactivating f_i violates the area coverage constraint*
2) *For all active femtocells $f_i \in F^*$, there does not exist a reassignment of all the users of f_i to its active neighbors*

A locally Pareto-optimal configuration cannot be unilaterally improved by a femtocell; that is, the energy consumption of the network cannot be reduced by any active femtocell based only on its knowledge and coordination with its immediate neighbors. A global reconfiguration may still reduce the number of active femtocells, but this requires global information on the network state which is not available to individual femtocells.

Theorem 6: GREENFEMTO *converges to a locally Pareto-optimal network state.*

Proof: We proceed by contradiction. Let us assume that the final network state S_f is not locally Pareto optimal, i.e. there exists at least one active femtocell f_o such that turning off f_o does not violate the area coverage constraint, and there exists a reassignment of all users of f_o to its active neighbors.

f_o performs a redundancy test if any of three events occurs:

1) f_o detects a new user
2) a user assigned to f_o leaves the network
3) f_o receives an `Info`, `Awake`, or `Sleep` message from one of its neighbors

Consider, T, the most recent redundancy test performed by f_o. Since f_o is active in S_f, T was not successful, either because turning off f_o would violate the area coverage constraint, or because f_o was unable to find a reassignment of all of its users. Therefore, immediately after T, the network state, S_T, is locally Pareto optimal. Since no further redundancy tests occur between T and S_f, we can assume that none of the three events listed above occurred, since if one of the events had occurred, a test T' would have occurred between T and S_f, which violates the assumption that T is the most recent test performed by f_o. Since none of the events occurred, no changes to the local Pareto optimality of the network could have occurred. As a result, f_o cannot be disabled, and S_f is locally Pareto optimal, contradicting the hypothesis. ■

The above theorem ensures that, if there are no significant changes in the network, a user can only experience a finite number of handoffs, therefore preventing indefinite ping-pong effects between cells.

TABLE I

SIMULATION PARAMETERS

Parameter	Value	Description
$\|\mathcal{L}\|$	$100\ m \times 100\ m$	Area of interest
n_{TX}	2	Number of transmit antennas
n_{RX}	2	Number of receive antennas
f	2140 Mhz	System frequency
b	20 Mhz	System bandwidth
p_M^{tx}	40 dBm	Macrocell transmission power
p_F^{tx}	24 dBm	Femtocell transmission power
μ_s	0 dB	Shadow fading mean
σ_s	10 dB	Shadow fading standard deviation
n_r	9 dB	Receiver noise figure
T_k	-174 dBm/Hz	Thermal noise density
t^a	19 dB (SNR)	Area coverage threshold
t^u	-7 dB (SINR)	User coverage threshold
$t^{handoff}$	-6 dB (SINR)	Handoff threshold
N_{rb}	100	Resource blocks per base station
λ	3 sec^{-1}	Arrival rate of users in mobile scenario
μ	3 sec^{-1}	Call duration of user in mobile scenario

VII. SIMULATION RESULTS

In this section we describe our simulations environment and present the results.

A. Simulation Details

We use a Matlab LTE system-level simulator from the Vienna University of Technology [21]. A full description of the simulator is available in the reference. The simulator models an LTE system with multiple eNodeBs and user equipments (UEs), using three radio link quality factors: large-scale pathloss, shadow fading, and small-scale link-level fading. Parameters used in the simulation are listed in Table I.

We consider a dense femtocell deployment with no underlying macrocell coverage. Femtocells use unidirectional antennas in a one-sector configuration. Since the femtocells are densely deployed, we assume that they are connected to the same local network, through which they can communicate using either a local management server or directly in a peer-topeer fashion. Users are uniformly randomly distributed, with minimum rate targets.

B. Results

We compare three algorithms: GREENFEMTO, GLOBAL, a computed global reassignment optimization based on GreenFemto, and IDLE, which is a modification of an algorithm for femtocell power saving described in [14]. We present results from three different experimental scenarios. We begin with results on stationary users, with the number of users increasing between experimental runs, to highlight the performance of the algorithms as the number of users changes. Then we present results with a fixed number of users, as the number of femtocells increases, to show the performance of the algorithm as the per-femtocell load varies. Finally, we present a realistic scenario with mobile users that join and leave the network at the end of the section.

GLOBAL uses the Gurobi [22] solver to solve the **MinActivation** problem in Section III, given the activation vector $\bar{X}$ found by GREENFEMTO and the functions $b_{ij}(\cdot)$ for all femtocells f_i and users u_j as input. It then iterates on this solution until convergence.[6] In the case that the problem is infeasible due to user coverage or femtocell capacity, the users that cannot be served are rejected one-by-one and the problem is reattempted without them until a solution is found. GLOBAL prioritizes serving users over reducing the number of active femtocells: users are only rejected if there is no femtocell that can serve them.

GLOBAL is used to evaluate the femtocell activation and user assignment that GREENFEMTO finds. It finds the optimal assignment of users to femtocells given a set of active femtocells, but it does not find the solution to the larger problem of determining which set of femtocells should be active. By using the solution that GREENFEMTO finds as input, we can determine whether there is a global reassignment of users to femtocells that uses fewer femtocells, improves SINR, or rejects fewer users. The performance difference between GLOBAL and GREENFEMTO highlights the impact of limiting GREENFEMTO to local reassignments.

The IDLE algorithm partially disables a femtocell's electronics when no users are present, leaving only a low-power user detection mode active. When a user enters the network, it is detected by all of the femtocells in range. The user connects to the femtocell that can serve it with the highest SINR, while the remaining femtocells return to the low-power user-detection mode. In IDLE, each femtocell is independent and does not communicate with other femtocells on the network. We modify the IDLE algorithm by requiring that area coverage is provided, as in GREENFEMTO, by finding a minimum covering set of femtocells using Gurobi.

1) Stationary Users, Increasing Number of Users, Fixed Number of Femtocells: Figure 4 shows the number of active femtocells with stationary users that are randomly and uniformly distributed, with 50 total femtocells. In this experiment the varying number of users is obtained by successive evaluations of the network under different levels of user density. All graphs show the average of 10 runs, with error bars indicating the 95% confidence interval. At low load, the number of femtocells activated by GREENFEMTO is dominated by area coverage. The difference between GREENFEMTO and GLOBAL is small, since most of the femtocells activated by GREENFEMTO are also needed for area coverage. IDLE activates three times the number of femtocells activated by GREENFEMTO.

As the number of users increases, the number of femtocells activated by GREENFEMTO increases linearly, while the number of femtocells activated by IDLE increases by the inverse square. IDLE turns on the best femtocell for a user, which in most cases is the nearest femtocell. Therefore, IDLE's performance is dominated by the number of femtocells that are the nearest femtocell to a user, which is determined by the user and femtocell distribution. The gap between the global optimal and GREENFEMTO's solution increases slightly due to the additional degrees of freedom in the ILP problem made available by the activation of more femtocells to cover users. With more femtocells, the probability of a superior global reassignment increases, and the local reassignment limitation of GREENFEMTO hurts its performance more.

Figure 5 shows the average user wideband SINR with 100 femtocells total. This SINR measurement is an average of the user SINR measured across the entire channel bandwidth; due to small-scale fading, individual subchannels may have SINRs above or below the wideband SINR. The lowest user SINR threshold is -7 dB, and all three algorithms find femtocell activations that achieve this target on average. Since IDLE activates the best femtocell for each user, it reaches a high average SINR. The difference between GREENFEMTO and GLOBAL can be explained by the path-dependence and local reassignment limitations inherent in the design of GREENFEMTO. GREENFEMTO focuses on reducing femtocell

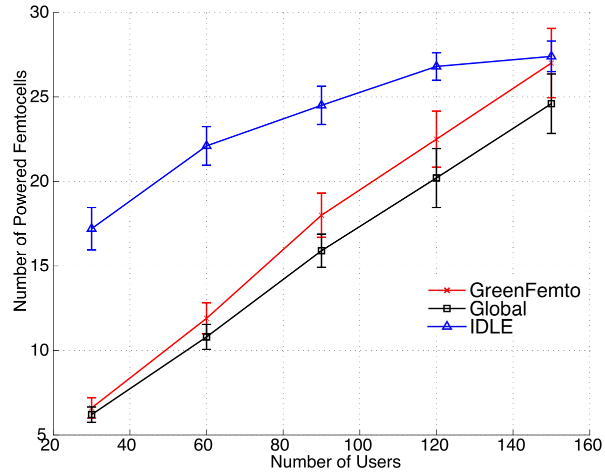


Fig. 4. Number of powered femtocells, 50 femtocells, increasing user scenario.

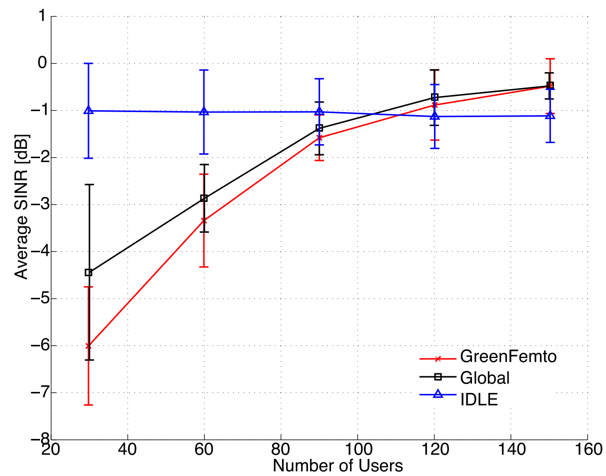


Fig. 5. Average SINR, 50 femtocells, increasing user scenario.

[6]Since **MinActivation** can only activate as many or fewer femtocells as were active in the input activation vector, it is clear to see that it converges.

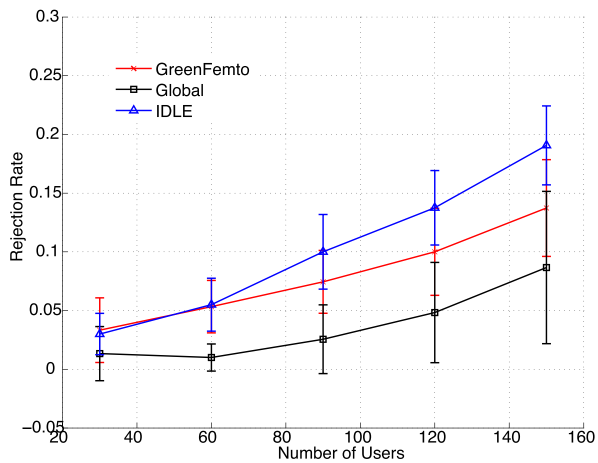


Fig. 6. Number of user rejections, 50 femtocells, increasing user scenario.

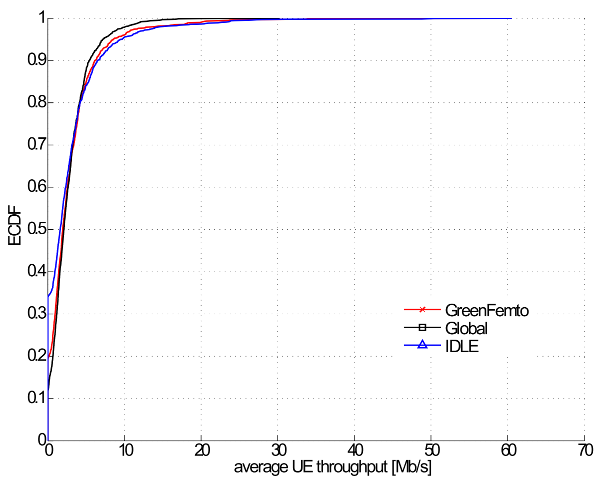


Fig. 7. Throughput Empirical Cumulative Density Function (ECDF), 50 femtocells, increasing user scenario.

power consumption, therefore GREENFEMTO will not reassign a users to improve their SINR.

Figure 6 shows the user rejection rate. IDLE rejects a large number of users because it activates a large number of femtocells when new users enter the network, which increases interference. GLOBAL's performance indicates that for approximately 50% of the users that GREENFEMTO rejects, there exists a way to serve them given the current active set of femtocells, however, it is not possible for GREENFEMTO to find these solutions. For example, this could occur if a new user can only be served by one femtocell, which is full of users that could be served on other femtocells. Since GREENFEMTO does not preempt users, the new user is rejected, while an assignment is possible using GLOBAL.

Figure 7 shows the throughput Empirical Cumulative Density Function (ECDF) for users across all ten runs. IDLE performs poorly in terms of throughput, with 35% of users achieving 0 throughput, due to rejection, insufficient resources, or poor subchannel SINR. Despite using more femtocells on average, IDLE performs only slightly better than GREENFEMTO, with only the top 5% attaining higher throughput than GREENFEMTO. The top 10% of users in GREENFEMTO and IDLE reach the same throughput of 6 Mbps.

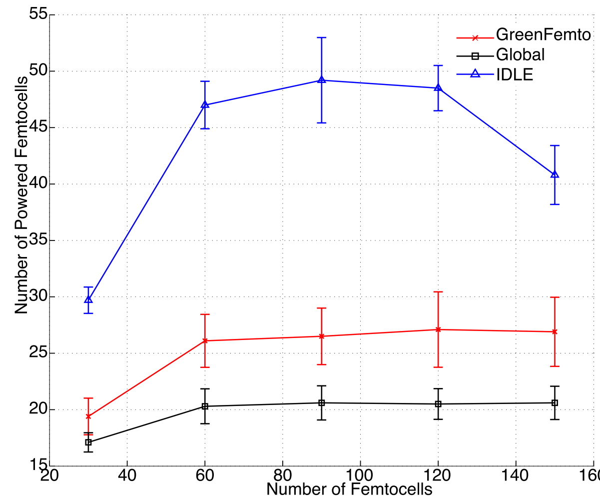


Fig. 8. Number of powered femtocells, 100 users, increasing femtocell scenario.

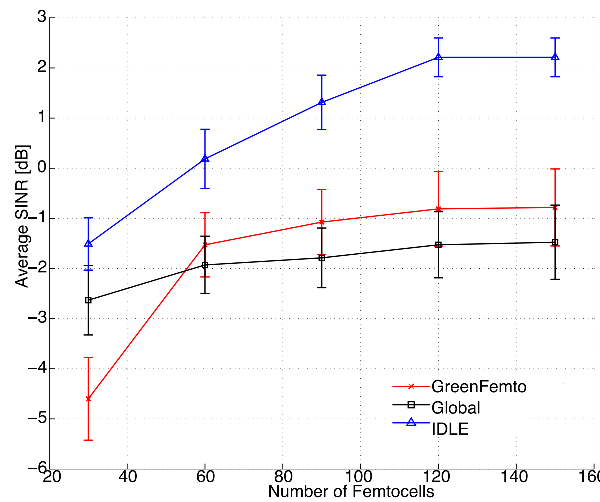


Fig. 9. Average SINR, 100 users, increasing femtocell scenario.

2) Stationary Users, Fixed Number of Users, Increasing Number of Femtocells: Figure 8 shows the number of active femtocells as the total number of femtocells varies, with 100 users. Up to 90 femtocells, the number of active femtocells increases with the total number of femtocells, because not all users can be served, as can be seen in the rejection graph in Figure 10. IDLE scales poorly as the number of femtocells increases, with the number of user rejections *increasing* with the total number of femtocells. This increase is due to the naive way that IDLE wakes up femtocells to serve users, with all femtocells that detect the presence of a user waking up, rather than only a sentinel femtocell's neighbors. The larger number of woken up femtocells cause significant co-channel interference, which leads to high user rejections.

Figure 9 shows the average SINR as the total number of femtocells increases. As the number of femtocells increases, the distance between femtocells and users decreases on average. IDLE directly benefits from this increase, because it turns on each user's best femtocell. GREENFEMTO performs poorly when the relative load is high, finding solutions with average SINRs below GLOBAL, despite using more femtocells. However, as the number of femtocells increases, GREENFEMTO's

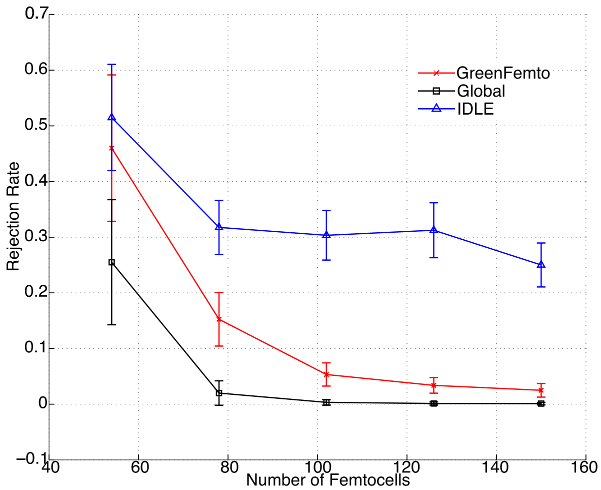


Fig. 10. Number of user rejections, 100 users, increasing femtocell scenario.

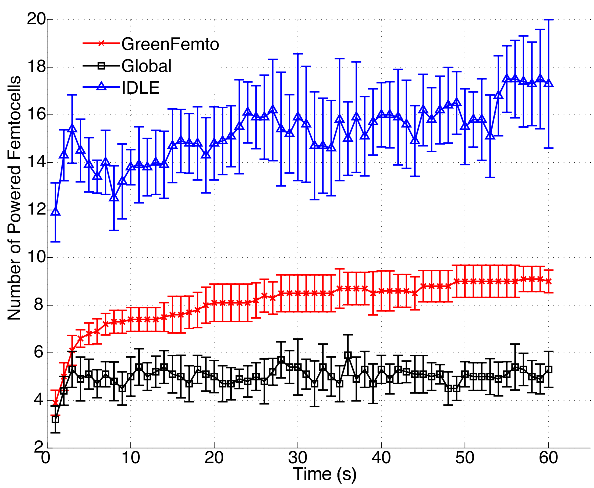


Fig. 11. Number of powered femtocells, 50 mobile users.

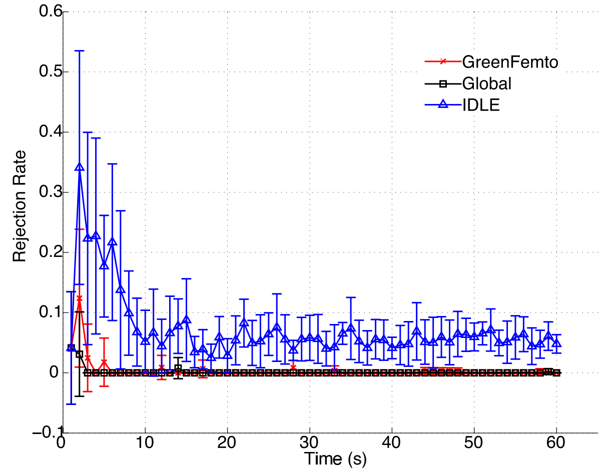


Fig. 12. Rejection rate, 50 mobile users.

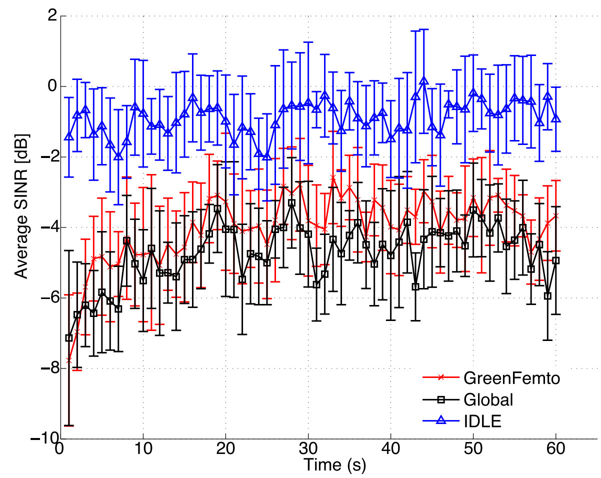


Fig. 13. Mean SINR, 50 mobile users.

solutions improve. GREENFEMTO's sensitivity to the total number of femtocells is due to the local search it performs for alternative femtocells. When there are a small number of femtocells, the local search performs poorly.

3) Mobile Users, Fixed Number of Femtocells: We perform experiments to investigate the performance of the algorithms with mobile users. We considered 50 mobile users that initially join the network. We model the interaction of the users with the network as follows. Users periodically connect to the network performing a call. Both the inter-arrival time between calls and the call duration follow an exponential distribution with parameter $\lambda = 3$ sec^{-1} and $\mu = 3$ sec^{-1}, respectively. Therefore, the total number of connected users varies over time. Additionally, user move at 3 m/s using a random waypoint mobility model. This scenario allows the algorithms' behavior in a realistic setting to be analyzed.

Figure 11 shows the number of active femtocells over time. The number of active femtocells varies significantly with the number of users in the system when using the IDLE algorithm. GREENFEMTO exhibits much lower variation in the number of active femtocells; with around 9 femtocells active for the majority of the simulation time, which is half the number that IDLE uses. Despite using fewer femtocells, GREENFEMTO rejects far fewer uses than IDLE, as seen in Figure 12. Figure 13 shows the mean SINR in the mobile user scenario. Under the IDLE algorithm, mean SINR is higher, but this comes at the expense of a higher rejection rate. Overall GREENFEMTO finds solutions that use significantly fewer femtocells than IDLE, and within 15% of an optimal reassignment.

VIII. RELATED WORK

Reducing the energy consumption of femtocell networks has received significant attention in the recent literature.

Several papers design power control schemes for femtocell networks [5]–[7]. In particular, [5] proposes a distributed algorithm based on game theory to adjust the transmission power of femtocells. The authors of [6] introduce a distributed algorithm to jointly optimize the power consumption of femtocells and the scheduling order for serving users. The work in [7] studies a Pareto optimal power control and scheduling algorithm which aims at improving the spectral efficiency. The above mentioned papers mainly focus on adjusting the transmission power and do not consider the possibility of turning off femtocells, thus achieving less energy saving.

There has been work that focuses on problems loosely related to ours. The authors of [23] consider the problem of

TABLE II
REVIEW OF OTHER APPROACHES TO FEMTOCELL SELECTIVE ACTIVATION

	This paper	[31]	[32]	[33]	[34]
(a) Centralized	No	Yes	No	Yes	Yes
(b) Dynamic	Yes	Yes	Yes	No	No
(c) Reaction Time	Seconds	Minutes	Seconds	N/A	N/A
(d) Metric	Local users	Global users	Local pilot SINR	Global rate	Global SINR
(e) Femto-specific	Yes	No	Yes	Yes	Yes

user scheduling in self-organizing femtocell networks. The approach does not focus on network power consumption while it looks for the femtocell channel settings and user scheduling to reduce interference and provide better quality of service to the users. In [24] the authors propose an analytical model to characterize the power consumption of macrocell, microcell, picocell and femtocell based networks. The authors of [25] study the tradeoffs between network energy consumption and terminal energy consumption. The tradeoffs between user perceived quality of service and network energy consumption are instead studied in [26]. In [27] the authors investigate solutions for reducing the number and size of active macrocell to reduce the network energy consumption. Finally, in [28] the authors study heterogeneous networks with cognitive radio capabilities in order to reduce energy consumption by exploiting spectrum sharing.

The problem of minimizing the energy consumption of a two-tier network composed of a macrocell and several femtocells has been addressed in [8]–[11]. The works [8], [9], and [11] design centralized algorithms, while in [10] a hierarchical reinforcement learning approach is proposed. On the one hand, centralized algorithms do not scale in dense and dynamic scenarios such as the one considered in this paper. On the other hand, these works aim at optimizing the energy consumption of a network composed by a macrocell and several femtocells. In this paper we focus on the more challenging scenario, with less degree of freedom, in which the femtocell network does not belong to the service provider, but to a different institution which aims at reducing its own energy expenses.

The problem of reducing the energy consumption of two-tier networks through selective activation has been considered in [12]–[14] and [30]–[34]. Dhillon *et al.* [29] describe heterogeneous networks where small cells use energy harvesting to supply power. Although this paper broadly explores similar concepts to ours, the restrictions of energy harvesting and their analytical model result in different design decisions. The authors of [12] propose a low complexity sleeping mechanism which makes use of user traffic prediction. This work is complementary to ours, as traffic prediction can be used to further improve the performance of GREENFEMTO. The approach proposed in [13] and [14] is limited in scope to single femtocells, and is used in this paper as a comparison against GREENFEMTO. Combes *et al.* [30] describe an algorithm that uses a centralized controller to reduce the energy consumption of macrocell networks; however, it does not specifically address femtocells.

Dudnikova *et al.* [31] describe a coverage-focused distributed algorithm that adjusts femtocell pilot signals and uses sleep mode to disable unused femtocells. Due to their emphasis on coverage, they provide no throughput results, which made their algorithm poorly suited for a comparison with ours. Jin *et al.* [32] present a centralized algorithm that optimizes the expression global throughput divided by global energy consumption of the network by disabling femtocells. However, it is not a reactive algorithm and must be re-run from scratch when the network changes. Ebrahim and Alsusa [33] present a centralized algorithm that iterates over a fixed context of users and femtocells to create an interference map from which to make decisions about which femtocells to disable. However, the interference map is not updated as femtocells are disabled, leading to suboptimal behavior.

Given the breadth of work that explores similar, but not identical concepts to this paper, we summarize some of the key differences from the papers that are most similar to ours in Table II. The table describes whether the provided algorithm is (a) centralized or distributed, if the algorithm is (b) adaptive to changing network conditions and user demand, (c) the time scale of its responsiveness to changes, (d) the key aspect that triggers the algorithm decisions, and (e) whether the designed algorithm addresses the specific features of femtocell networks.

IX. CONCLUSION AND FUTURE WORK

This paper introduced GREENFEMTO, a novel distributed algorithm that finds a solution to **MinActivation**, a femtocell selective activation problem. Selective activation has not been extensively studied in femtocell networks. However, due to the nature of femtocell power consumption, we find that it is a promising approach to reducing the energy consumption of these networks. Our simulations show that selective activation can be used in these networks to reduce total energy consumption, and improve outage probability. We find that selective activation algorithms, if designed correctly, can improve network performance with low network churn. Our solution to **MinActivation** makes only local changes. Potential future work can be done on a solution that makes global changes. Future solutions, currently under investigation, concern the use of power control and resource block scheduling, to reduce interference and improve quality of service.

REFERENCES

[1] *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2010–2017 White Paper*, Cisco, San Jose, CA, USA, Feb. 2013.

[2] Aruba. *High-Density Wireless Networks for Auditoriums Validated Reference Design*, accessed on Aug. 2016. [Online]. Available: http://community.arubanetworks.com/aruba/attachments/aruba/ArubaVRDs/21/1/HighDensity%20Wireless%20Networks%20for%20Auditoriums.pdf

[3] Northern Towers Industries Inc. *PCT Series*, accessed on Aug. 2016. [Online]. Available: http://www.northerntowers.ca/

[4] Cisco. *Wireless LAN Design Guide for High Density Client Environments in Higher Education*, accessed on Aug. 2016. [Online]. Available: http://www.scribd.com/doc/62035348/High-Density-Client-Deployment-Guide-for-Cisco-Wireless-LANs

[5] L. B. Le, D. Niyato, E. Hossain, D. I. Kim, and D. T. Hoang, "QoS-aware and energy-efficient resource management in OFDMA femtocells," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 180–194, Jan. 2013.

[6] G. Cao, D. Yang, and X. Zhang, "A distributed algorithm combining power control and scheduling for femtocell networks," in *Proc. IEEE WCNC*, Apr. 2012, pp. 2282–2287.

[7] H. Burchardt, S. Sinanovic, G. Auer, and H. Haas, "Pareto optimal SINR scheduling for femto-cell deployment in wireless networks," in *Proc. IEEE VTC*, Sep. 2012, pp. 1–5.

[8] Y.-L. Chung, "Green dynamic configuration in two-tier LTE femtocell networks," in *Proc. IEEE ICOIN*, Jan. 2013, pp. 551–555.

[9] Z. Pan and S. Shimamoto, "Cell sizing based energy optimization in joint macro-femto deployments via sleep activation," in *Proc. IEEE WCNC*, Apr. 2013, pp. 4765–4770.

[10] X. Chen, H. Zhang, T. Chen, and M. Lasanen, "Improving energy efficiency in green femtocell networks: A hierarchical reinforcement learning framework," in *Proc. IEEE ICC*, Jun. 2013, pp. 2241–2245.

[11] L. Saker, S. E. Elayoubi, R. Combes, and T. Chahed, "Optimal control of wake up mechanisms of femtocells in heterogeneous networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 664–672, Apr. 2012.

[12] G. Wang, C. Guo, S. Wang, and C. Feng, "A traffic prediction based sleeping mechanism with low complexity in femtocell networks," in *Proc. IEEE ICC*, Jun. 2013, pp. 560–565.

[13] I. Ashraf, F. Boccardi, and L. Ho, "Power savings in small cell deployments via sleep mode techniques," in *Proc. IEEE PIMRC*, Sep. 2010, pp. 307–311.

[14] I. Ashraf, L. T. W. Ho, and H. Claussen, "Improving energy efficiency of femtocell base stations via user activity detection," in *Proc. IEEE WCNC*, Apr. 2010, pp. 1–5.

[15] M. Lin, S. Silvestri, N. Bartolini, and T. La Porta, "Energy-efficient selective activation in femtocell networks," in *Proc. IEEE MASS*, Oct. 2015, pp. 361–369.

[16] Bureau veritas consumer products services, "FCC TEST REPORT (PART 22)," Tech. Rep. RF130726C18, Taipei, Taiwan.

[17] *Evolved Universal Terrestrial Radio Access (E-UTRA); LTE Physical Layer; General Description*, document TS 36.201, 3GPP, Mar. 2010. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/36201.htm

[18] G. Xing, X. Wang, Y. Zhang, C. Lu, R. Pless, and C. Gill, "Integrated coverage and connectivity configuration for energy conservation in sensor networks," *ACM Trans. Sensor Netw.*, vol. 1, no. 1, pp. 36–72, 2005.

[19] Lieberman Software Corporation, "White Paper: Wake on LAN Technology," Los Angeles, CA, USA. [Online]. Available: http://feishare.com/attachments/article/201/Wake_On_LAN.pdf

[20] S. Martello and P. Toth, *Knapsack Problems: Algorithms and Computer Implementations*. New York, NY, USA: Wiley, 1990, ch. 7.

[21] S. Schwarz, J. C. Ikuno, M. Simko, M. Taranetz, Q. Wang, and M. Rupp, "Pushing the limits of LTE: A survey on research enhancing the standard," *IEEE Access*, vol. 1, pp. 51–62, 2013.

[22] *Gurobi*, accessed on Aug. 2016. [Online]. Available: http://gurobi.com

[23] H.-L. Chao, S.-H. Wu, Y.-H. Huang, and S.-C. Li, "Cooperative spectrum sharing and scheduling in self-organizing femtocell networks," in *Proc. IEEE ICC*, Jun. 2014, pp. 3246–3251.

[24] A. Mukherjee, S. Bhattacherjee, S. Pal, and D. De, "Femtocell based green power consumption methods for mobile network," *Comput. Netw.*, vol. 57, no. 1, pp. 162–178, 2013.

[25] Q. Chen, G. Yu, Y. Jiang, and A. Huang, "Tradeoff between network energy consumption and terminal energy consumption via small cell power control," in *Proc. IEEE IWCMC*, Jul. 2013, pp. 573–578.

[26] Z. Huang, H. Xia, Z. Zeng, and Y. Liu, "Optimization of energy efficiency for OFDMA femtocell networks based on effective capacity," in *Proc. IEEE VTC*, Sep. 2012, pp. 1–5.

[27] C. Khirallah and J. S. Thompson, "Energy efficiency of heterogeneous networks in LTE-advanced," *J. Signal Process. Syst.*, vol. 69, no. 1, pp. 105–113, 2012.

[28] R. Xie, F. R. Yu, and H. Ji, "Energy-efficient spectrum sharing and power allocation in cognitive radio femtocell networks," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 1665–1673.

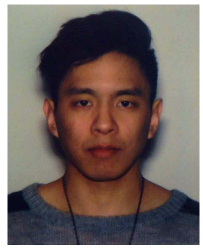
[29] H. S. Dhillon, Y. Li, P. Nuggehalli, Z. Pi, and J. G. Andrews, "Fundamentals of heterogeneous cellular networks with energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2782–2797, May 2014.

[30] R. Combes, S. E. Elayoubi, A. Ali, L. Saker, and T. Chahed, "Optimal online control for sleep mode in green base stations," *Comput. Netw.*, vol. 78, pp. 140–151, Feb. 2015.

[31] A. Dudnikova, D. Panno, and A. Mastrosimone, "Measurement-based coverage function for green femtocell networks," *Comput. Netw.*, vol. 83, pp. 45–58, Jun. 2015.

[32] Y. Jin, L. Qiu, and X. Liang, "Small cells on/off control and load balancing for green dense heterogeneous networks," in *Proc. IEEE WCNC*, Mar. 2015, pp. 1530–1535.

[33] A. Ebrahim and E. Alsusa, "Interference minimization through sleep mode based resource allocation for future femtocell networks," in *Proc. IEEE ICC*, Jun. 2015, pp. 1679–1684.



Michael Lin received the B.S. degree (Hons.), the M.S. degree, and the Ph.D. degree from Pennsylvania State University in 2006, 2008, and 2015, respectively. He was a Research Intern with Sprint Advanced Technology Labs, Burlingame, CA, USA, and Alcatel-Lucent Bell Labs, Murray Hill, NJ, USA. He is currently with Fugue, Inc., Washington, DC, USA. His research interests are in heterogeneous wireless networks and energy-aware networking.

Simone Silvestri received the degree (Hons.) and the Ph.D. degree in computer science from the Sapienza University of Rome, Rome, Italy. He is currently an Assistant Professor with the Computer Science Department, Missouri University of Science and Technology. His research interests lie in the area of network management, hybrid wireless sensor networks, interdependent cyber-physical systems, and smart grid security. He serves as Program Committee Member of several international conferences.

Novella Bartolini received the degree (Hons.) and the Ph.D. degree in computer engineering from the University of Rome, Rome, Italy, in 1997 and 2001, respectively. She was a Researcher with the Fondazione Ugo Bordoni in 1997, a Visiting Scholar with the University of Texas at Dallas from 1999 to 2000, and a Research Assistant with the University of Rome Tor Vergata from 2000 to 2002. She has been an Associate Professor with the Sapienza University of Rome and a Visiting Professor with Penn State University, since 2014. Her research interests lie in the area of wireless networks and network management. She was a Program Chair and a Program Committee Member of several international conferences. She has served on the Editorial Board of the *Computer Networks* (Elsevier) and ACM/Springer *Wireless Networks*.

Thomas F. La Porta (S'85–M'87–SM'99–F'02) received the B.S.E.E. and M.S.E.E. degrees from The Cooper Union, New York, NY, USA, and the Ph.D. degree in electrical engineering from Columbia University, New York, NY, USA. He has been with Bell Laboratories since 1986, where he was the Director of the Mobile Networking Research Department where he led various projects in wireless and mobile networking. He is currently the William E. Leonhard Chair Professor with the Department of Computer Science and Engineering, Pennsylvania State University. He joined Penn State in 2002. He is the Director of the Institute for Networking and Security Research. His research interests include mobility management, signaling and control for wireless networks, mobile data and sensor systems, and network security. He was the Founding Editor-in-Chief of the IEEE TRANSACTIONS ON MOBILE COMPUTING. He served as the Editor-in-Chief of the *IEEE Personal Communications Magazine* for three years. He is an Bell Labs Fellow. He received the Bell Labs Distinguished Technical Staff Award in 1996, and the Eta Kappa Nu Outstanding Young Electrical Engineer Award in 1996. He received the Thomas Alva Edison Patent Awards in 2005 and 2009, respectively.