# Optimizing Small-Cell Range in Heterogeneous and Load-Coupled LTE Networks

Iana Siomina and Di Yuan

**Linköping University Post Print**



N.B.: When citing this work, cite the original article.

# Optimizing Small Cell Range in Heterogeneous and Load-Coupled LTE Networks

Iana Siomina and Di Yuan *Member, IEEE*

### Abstract

We address small cell range optimization in heterogeneous LTE networks with the performance target of maximizing the scaling-up factor of traffic demand that can be served, while accounting for load coupling between cells. We prove the problem's complexity, and develop a solution approach consisting in two complementary modules. We also demonstrate how to gauge the deviation from global optimality. Performance results show the effectiveness of the approach and highlight the benefit of range optimization.

**Index Terms**: Cell range, heterogeneous networks, load coupling, long term evolution, optimization, small cells.

## I. INTRODUCTION

A significant step in Long Term Evolution (LTE) is the deployment of heterogeneous networks (HetNets) integrating macro cells (MCs) with small cells (SCs) [1], [2]. In HetNets, SCs offload MCs and provide enhanced capacity to traffic hotspots. We model and computationally solve the optimization problem of setting SC range via cell-specific offsets, with the performance target of maximizing the scaling-up factor of traffic demand that can be met. The use of offset does not physically change transmit power. Rather, offset provides a virtual amplification; namely, if offset is used by a cell, the user equipments (UEs) will add the cell's offset value to the received reference signal strength, and take the sum as the preference score in the procedure of cell selection. As results, offset acts as a bias to the physical signal strength, and enable SCs to be selected by UEs even if the strongest signal originates from some MC.

We refer to [3]–[5], respectively, for throughput evaluation, interference mitigation, and analysis of outage probability with the presence of biased cell selection with offsets. Our study differs from these works in several aspects. First, we consider cell-individual offsets and use them as optimization variables. Second, our system model includes the decision of SC deployment. Third, we perform range optimization subject to the non-linear LTE load-coupling model [6]–[8]. Here,

a cell's load represents the utilization level of the time-frequency resource units in orthogonal frequency division multiple access (OFDMA). Load coupling refers to the dependency relation between the cell load levels due to mutual interference, where a cell's load dictates the interference it generates to other cells. In HetNets with shared resource, load coupling occurs among the MCs, SCs, as well as between the two cell groups. The load-coupling model was proposed in [6], [7]. Dynamic simulations in [6] show the model is sufficiently accurate for network-level performance characterization. In [8], we provided a theoretical analysis of the model. The model has been used in the context of offloading [9], network planning [10], and load balancing [11].

The current paper presents original contributions in relation to [7], [8]. In [7], the analysis is limited to uniform cell load – an assumption that is not realistic, in particular for HetNets. The study in [8] focused on theoretical properties of load coupling for fixed network configuration, that is, the work is confined to given UE-cell association, without any network optimization nor performance objective. Moreover, whereas for given network configuration the load-coupling model submits to convex analysis in [7], [8], range assignment for capacity optimization in the current paper calls for novel combinatorial optimization techniques. Hence analysis and development of optimization methods for range optimization in the forthcoming sections represent original contributions in respect of [7], [8]. In comparison to [11], the current work considers a new performance objective and provides tractability analysis, and presents original solution methodologies. Specifically, we present the following contributions. We formulate the range optimization problem and prove its NP-hardness. Then, we develop a solution approach consisting in two complementary modules, via concave Perron-Frobenius theory and integer linear programming (ILP), respectively. Combining the modules delivers a sequence of monotonically improving solutions. Next, we demonstrate how to use ILP for gauging the optimality gap. We present performance results for a representative HetNet scenario, showing the effectiveness of the approach and the benefit of range optimization.

## II. SYSTEM MODEL

In this section we formalize the system model. For clarity, a summary of the basic notation is provided in Table I. We denote by $M$ and $N$ the numbers of MCs and SCs, respectively. The set of MCs is denoted by $\mathcal{M} = \{1, \ldots, M\}$. To avoid any ambiguity of cell indices, the SCs are numbered from $M + 1$ and upwards, hence the SC set is $\mathcal{N} = \{M+1, \ldots, M+N\}$. The union of the two cell types is denoted by $\mathcal{C} = \mathcal{M} \cup \mathcal{N}$. The set of UEs is denoted by $\mathcal{J}$. In general, an element of $\mathcal{J}$ may be a UE or a demand point aggregating multiple UEs. Notation $p_i$, $g_{ij}$, and $\sigma^2$ are used to denote, respectively, the power spectral density per resource unit of cell $i$, the total gain between the antenna of cell $i$ and UE $j$, and the noise power. The baseline demand value of

UE $j$ is denoted by $d_j$, to which the maximization of demand scaling applies.

We use $\mathcal{S}$ to represent the set of candidate SC range offsets. The values are in linear scale, acting as multipliers to received signal strength. We let $\mathcal{S}$ include offset zero, which, if selected for an SC, implies that no UE will be served by the SC and hence in effect this SC is not deployed. The deployment budget, i.e., the maximum number of SCs to be deployed, is denoted by $n$, with $0 < n \leq N$. For UE $j$, we refer to the MC giving the strongest signal as the default MC, and denote it by $m_j$. Formally, $m_j = \text{argmax}_{i \in \mathcal{M}} p_i g_{ij}$. Binary variables $y_{ir}, i \in \mathcal{N}, r \in \mathcal{S}$ are used to indicate whether SC $i$ uses range $r$. If $y_{ir} = 1$, then the received signal of UE $j$, biased by range $r$, is $p_i g_{ij} r$. In the sequel, the cell giving the strongest signal to UE $j$, induced by range assignment $\boldsymbol{y}$, is denoted by $c(j, \boldsymbol{y})$. Note that $c(j, \boldsymbol{y})$ equals the default MC, if no SC gives a stronger signal including the effect of its range.

We use $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_{M+N})^T$ to denote the load vector, where $\rho_i$ is the proportion of resource units consumed by cell $i$ for serving its UEs. The upper limit is denoted by $\boldsymbol{\rho}^{max}$, with $\boldsymbol{\rho}^{max} \leq \boldsymbol{1}$. Resource unit refers to the basic time-domain unit for scheduling in OFDMA. For LTE, it corresponds to 12 sub-carriers and 1 ms time span (i.e., a sub-frame with two consecutive time slots). For a UE $j$ of $i$, the signal-to-interference-and-noise ratio (SINR) is $\text{SINR}_{ij} = \frac{p_i g_{ij}}{\sum_{k \in \mathcal{C} \setminus \{i\}} p_k g_{kj} \rho_k + \sigma^2}$, in which the other cells' loads have the effect of interference scaling. On one resource unit, the rate is $B \log_2(1 + \text{SINR}_{ij})$, with $B$ being the bandwidth. Thus, to satisfy demand $d_j$ of UE $j$, cell $i$ has to utilize $\frac{d_j}{B \log_2(1 + \text{SINR}_{ij})}$ resource units. Taking the sum of these terms over cell $i$'s UEs, and then dividing the sum by $K$, where $K$ denotes the total number of resource units of the spectrum and the time in question, represent the proportion of resource consumption and hence gives the load in cell $i$. Without loss of generality, we normalize such that $KB = 1$. Thus the cell load levels, $\rho_i, i \in \mathcal{C}$, are coupled via the following non-linear equation system [6]–[8], [10].

$$\rho_i = f_i(\boldsymbol{\rho}, \boldsymbol{y}, \boldsymbol{d}) = \sum_{j \in \mathcal{J}: c(j, \boldsymbol{y}) = i} \frac{d_j}{\log_2(1 + \frac{p_i g_{ij}}{\sum_{k \in \mathcal{N} \setminus \{i\}} p_k g_{kj} \rho_k + \sigma^2})}, \quad i \in \mathcal{C}. \tag{1}$$

The demand to be served is not a constant bit rate in network operation. In network planning and dimensioning studies, to which the scope of the current paper belong, the maximum capacity is of a particular interest. One way of modelling the network capacity aspect in network planning is to maximize the amount of traffic that can be served, instead of looking at a constant amount of demand. This motivates the performance objective of scaling up the demand of all UEs with a common scaling factor and maximizing its value. This optimization problem is formalized in (2). The objective function is stated in (2a). Here, $\kappa$ is the auxiliary variable representing the demand

scaling factor, to be maximized by finding the optimal values of range offsets (vector $\boldsymbol{y}$) and cell load levels (vector $\boldsymbol{\rho}$). Equation (2d) is a compact notation of (1) with scaling factor $\kappa$ included. The other two constraint sets ensure, respectively, that exactly one range is selected for each SC and the total number of deployed SCs is bounded by $n$.

$$\max_{0 \leq \boldsymbol{\rho} \leq \boldsymbol{\rho}^{max}, \ \boldsymbol{y} \in \mathbb{B}^{N \times |\mathcal{S}|}} \kappa \tag{2a}$$

$$\sum_{r \in \mathcal{S}} y_{ir} = 1, i \in \mathcal{N} \tag{2b}$$

$$\sum_{i \in \mathcal{N}} \sum_{r \in \mathcal{S} : r \neq 0} y_{ir} \leq n \tag{2c}$$

$$\boldsymbol{\rho} = \boldsymbol{f}(\boldsymbol{\rho}, \boldsymbol{y}, \kappa \boldsymbol{d}) \tag{2d}$$

The optimization problem in (2) exhibits a combinatorial structure and appears to be non-convex. This does not imply its intractability, as it is known that some discrete and combinatorial problems (e.g., bipartite matching) are tractable, i.e., solved to global optimality in polynomial time. Thus the tractability result is of significance, both theoretically and for considering solution algorithms. To the best of our knowledge, there is no formal tractability result of HetNet range assignment with a rigorous proof. The following theorem resolves this aspect.

*Theorem 1:* Range optimization with maximum demand scaling, as defined in (2), is NP-hard.

*Proof:* See the appendix. ■

One can observe from the proof in the appendix that the result is not limited to (2a). Indeed, Theorem 1 applies as long as load feasibility has to be determined for given demand.

## III. SOLUTION APPROACH

Our solution approach consists of two modules: maximizing $\kappa$ for given range assignment, and load minimization for given $\kappa$, respectively. The algorithm alternates between the two modules until no positive slack in load can be found. The algorithmic development relies on some forthcoming theoretical insights (Theorems 2-4). Theorem 2 is derived from concave Perron-Frobenius theory. Theorems 3 and 4 represent non-trivial theoretical insights of range assignment, and provide justifications of the algorithmic flow and the bounding scheme, respectively. For succinctness, the proofs are given in compact form as long as clarity is not lost.

### A. Demand Maximization for Fixed Range Assignment

Demand maximization for given range assignment $\boldsymbol{y}$ is formulated below. Load coupling is simplified to $\boldsymbol{\rho} = \boldsymbol{f}(\boldsymbol{\rho}, \kappa \boldsymbol{d})$ here because cell-UE association is implied by $\boldsymbol{y}$. Thus we have

$$\max_{0 \leq \boldsymbol{\rho} \leq \boldsymbol{\rho}^{max}} \kappa \tag{3a}$$

$$\boldsymbol{\rho} = \boldsymbol{f}(\boldsymbol{\rho}, \kappa \boldsymbol{d}) \tag{3b}$$

Denote by $(\kappa^\star, \boldsymbol{\rho}^\star)$ an optimum of (3). As $\kappa$ is a scalar, $\kappa^\star$ can be found by bi-section search in which (3b) is solved for demand $\kappa \boldsymbol{d}$ for each trial value of $\kappa$. For uniform load limit, i.e., $\rho_i^{max} = \rho^{max}, \forall i \in \mathcal{C}$, we show that, via concave Perron-Frobenius theory, $(\kappa^\star, \boldsymbol{\rho}^\star)$ can be computed without bi-section search. Concave Perron-Frobenius theory [13] extends the classical Perron-Frobenius theory in linear algebra to concave mappings in standard cone $\mathbb{R}_+^{M+N}$. Particularly, for concave mappings, the conditional eigenvalue problem of finding a vector in $\mathbb{R}_+^{M+N}$ to form a self-mapping, such that a monotone norm of the vector equals a given constant, submits to normalized fixed-point iterations. We transform (3) into such a conditional eigenvalue problem.

Consider any initial load vector $\boldsymbol{\rho}^0 \in \mathbb{R}_+^{M+N}$, e.g., $\boldsymbol{\rho}^0 = \boldsymbol{f}(0, \boldsymbol{d})$. Given $\boldsymbol{\rho}^k$, a normalized fixed-point iteration computes the next iterate $\boldsymbol{\rho}^{k+1}$ as in (4), where $||\cdot||_\infty$ is the maximum norm.

$$\boldsymbol{\rho}^{k+1} = \frac{\boldsymbol{f}(\rho^{max} \boldsymbol{\rho}^k, \boldsymbol{d})}{||\boldsymbol{\rho}^k||_\infty} \tag{4}$$

*Theorem 2:* The sequence generated by (4) converges to a load vector $\mathring{\boldsymbol{\rho}}$ for any uniform load limit $\rho^{max} > 0$, and, at convergence, $\boldsymbol{\rho}^\star = \rho^{max} \mathring{\boldsymbol{\rho}}$, and $\kappa^\star = \frac{\rho_1^\star}{f_1(\boldsymbol{\rho}^\star, \boldsymbol{d})} = \cdots = \frac{\rho_1^\star}{f_{M+N}(\boldsymbol{\rho}^\star, \boldsymbol{d})}$.

*Proof:* Consider first $\rho^{max} = 1$. By the definition of $||\cdot||_\infty$, $\boldsymbol{\rho} \leq \mathbf{1}$ if and only if $||\boldsymbol{\rho}||_\infty \leq 1$. Suppose $||\boldsymbol{\rho}^\star||_\infty < 1$. Then $f_i(\boldsymbol{\rho}^\star, \kappa^\star \boldsymbol{d}) < 1, \forall i \in \mathcal{C}$, and therefore there exists $\epsilon > 1$ such that $f_i(\boldsymbol{\rho}^\star, \epsilon \kappa^\star \boldsymbol{d}) \leq 1, \forall i \in \mathcal{C}$. By Lemma 1 of [12] (cf. Theorem 5 in [8]), starting from $\mathbf{1}$ and applying regular fixed-point iterations for fixed demand $\epsilon \kappa^\star \boldsymbol{d}$ generates a monotonically decreasing sequence converging to the solution solving $\boldsymbol{\rho} = \boldsymbol{f}(\boldsymbol{\rho}, \epsilon \kappa^\star \boldsymbol{d})$. In other words, there is a load vector supporting demand $\epsilon \kappa^\star \boldsymbol{d}$, contradicting that $\boldsymbol{\rho}^\star$ is optimal. Hence $||\boldsymbol{\rho}^\star||_\infty = 1$, and (3) is equivalent to $\max\{\kappa : ||\boldsymbol{\rho}||_\infty = 1, \boldsymbol{\rho} = \boldsymbol{f}(\boldsymbol{\rho}, \kappa \boldsymbol{d}), \boldsymbol{\rho} \in \mathbb{R}_+^{M+N}\}$. By Theorem 2 of [8], $f_i(\boldsymbol{\rho}, \kappa \boldsymbol{d})$ is strictly concave in $\boldsymbol{\rho}$ for any $\boldsymbol{\rho} \geq 0$, $\forall i \in \mathcal{C}$, meaning that $\boldsymbol{f}(\boldsymbol{\rho}, \kappa \boldsymbol{d})$ is a concave mapping on $\mathbb{R}_+^{M+N} \to \mathbb{R}_+^{M+N}$. As $\boldsymbol{f}$ is linear in $\kappa$, $\frac{1}{\kappa} \boldsymbol{\rho} = \boldsymbol{f}(\boldsymbol{\rho}, \boldsymbol{d})$ is equivalent to (3b). Moreover, $||\cdot||_\infty$ is monotone. The system $\{||\boldsymbol{\rho}||_\infty = 1, \frac{1}{\kappa} \boldsymbol{\rho} = \boldsymbol{f}(\boldsymbol{\rho}, \boldsymbol{d}), \boldsymbol{\rho} \in \mathbb{R}_+^{M+N}\}$ is thus a conditional eigenvalue problem for concave mapping. Applying concave Perron-Frobenius theory (e.g., Theorem 1, [13]), solution $(\kappa^\star, \boldsymbol{\rho}^\star)$ of $\{||\boldsymbol{\rho}||_\infty = 1, \frac{1}{\kappa} \boldsymbol{\rho} = \boldsymbol{f}(\boldsymbol{\rho}, \boldsymbol{d}), \boldsymbol{\rho} \in \mathbb{R}_+^{M+N}\}$ is unique (and hence the max-operator over $\kappa$ is void), and normalized fixed-point iteration (4) converges to $\boldsymbol{\rho}^\star = \mathring{\boldsymbol{\rho}}$. The theorem's second part follows then from that $\frac{1}{\kappa^\star} \boldsymbol{\rho}^\star = \boldsymbol{f}(\boldsymbol{\rho}^\star, \boldsymbol{d})$. For the general cases of $\rho^{max}$, note

that $||\boldsymbol{\rho}||_\infty = 1$ if and only if $||\rho^{max}\boldsymbol{\rho}||_\infty = \rho^{max}$. The theorem follows from setting $\tilde{\boldsymbol{\rho}} = \rho^{max}\boldsymbol{\rho}$ and applying the above proof to $\tilde{\boldsymbol{\rho}}$. ∎

### B. Load Minimization for Given Demand

Intuitively, a given $\kappa$ is not optimal if there is a range assignment supporting $\kappa\boldsymbol{d}$ with slack in all cells' load levels. This is formalized below. We denote by $\kappa^\star(\boldsymbol{y})$ the maximum demand scaling for $\boldsymbol{y}$ (Section III-A), and $\boldsymbol{\rho}(\boldsymbol{y}, \kappa^\star(\boldsymbol{y}))$ the solution of the load-coupling system $\boldsymbol{\rho} = \boldsymbol{f}(\boldsymbol{\rho}, \boldsymbol{y}, \kappa^\star(\boldsymbol{y})\boldsymbol{d})$.

*Theorem 3:* Range assignment $\boldsymbol{y}^\star$ is optimal in (2) if and only if there does not exist any range assignment $\boldsymbol{y}'$ such that $\boldsymbol{\rho}(\boldsymbol{y}', \kappa^\star(\boldsymbol{y}^\star)) < \boldsymbol{\rho}^{max}$.

*Proof:* Suppose there is $\boldsymbol{y}'$ with $\boldsymbol{\rho}(\boldsymbol{y}', \kappa^\star(\boldsymbol{y}^\star)) < \boldsymbol{\rho}^{max}$. Then $\boldsymbol{f}(\boldsymbol{\rho}, \boldsymbol{y}', \kappa^\star(\boldsymbol{y}^\star)\boldsymbol{d}) < \boldsymbol{\rho}^{max}$, implying the existence of $\epsilon > 1$ with $\boldsymbol{f}(\boldsymbol{\rho}, \boldsymbol{y}', \epsilon\kappa^\star(\boldsymbol{y}^\star)\boldsymbol{d}) \leq \boldsymbol{\rho}^{max}$. Starting from $\boldsymbol{\rho}^{max}$ and applying regular fixed-point iterations to $\boldsymbol{\rho}$ with demand $\epsilon\kappa^\star(\boldsymbol{y}^\star)\boldsymbol{d}$ gives a monotonically decreasing and convergent load sequence by Theorem 2 of [8], hence $\boldsymbol{\rho}(\boldsymbol{y}', \epsilon\kappa^\star(\boldsymbol{y}^\star)) \leq \boldsymbol{\rho}^{max}$. This, together with $\epsilon > 1$, contradict that $\boldsymbol{y}^\star$ is optimal. For sufficiency, assume for all $\boldsymbol{y}' \neq \boldsymbol{y}^\star$, $\rho_i(\boldsymbol{y}', \kappa^\star(\boldsymbol{y}^\star)) \geq \rho_i^{max}$. Then for any $\hat{\kappa} > \kappa^\star$, by Lemma 1 of [12], starting from $\boldsymbol{\rho}(\boldsymbol{y}^\star, \hat{\kappa}(\boldsymbol{y}^\star))$ and applying regular fixed-point iterations to $\boldsymbol{\rho}$ with demand $\hat{\kappa}\boldsymbol{d}$ gives a monotonically increasing and convergent sequence to $\boldsymbol{\rho}(\boldsymbol{y}', \hat{\kappa})$, provided that $\boldsymbol{\rho}(\boldsymbol{y}', \hat{\kappa})$ exists at all. Therefore $\boldsymbol{\rho}(\boldsymbol{y}', \hat{\kappa}) \nleq \boldsymbol{\rho}^{max}$ and thus $\hat{\kappa}\boldsymbol{d}$ cannot be supported by $\boldsymbol{y}'$. Consequently $\boldsymbol{y}^\star$ is optimal. ∎

By Theorem 3, for given $\kappa$, range optimization amounts to minimizing the maximum load, or, equivalently, maximizing the minimum slack in load, with demand $\kappa\boldsymbol{d}$. For this non-linear and discrete problem, we apply ILP, enabled by a linear approximation of the load function $\boldsymbol{f}$, targeting close-optimal solutions. Denote by $\bar{\boldsymbol{y}}$ the range assignment in demand maximization (Section III-A), and by $\kappa^\star(\bar{\boldsymbol{y}})$ and $\boldsymbol{\rho}(\bar{\boldsymbol{y}}, \kappa^\star(\bar{\boldsymbol{y}}))$ the corresponding demand scaling and load, respectively. The linear approximation utilizes the partial derivatives of $\boldsymbol{f}$ at point $\boldsymbol{\rho}(\bar{\boldsymbol{y}}, \kappa^\star(\bar{\boldsymbol{y}}))$. To this end, we introduce continuous variable $w_{ij} = \sum_{k\in\mathcal{C}\backslash\{i\}} p_k g_{kj}\rho_k$ for $i \in \mathcal{C}, j \in \mathcal{J}$, to denote the amount of interference if cell $i$ serves UE $j$, and parameter $\bar{h}_{ij}$ for the partial derivative $\partial(\frac{\kappa^\star(\bar{\boldsymbol{y}})d_j}{\log_2(1+\frac{p_i g_{ij}}{w_{ij}+\sigma^2})})/\partial w_{ij}$, taken at $w_{ij} = \sum_{k\in\mathcal{C}\backslash\{i\}} p_k g_{kj}\rho_k(\bar{\boldsymbol{y}}, \kappa^\star(\bar{\boldsymbol{y}}))$, which is the total interference induced by $\boldsymbol{\rho}(\bar{\boldsymbol{y}}, \kappa^\star(\bar{\boldsymbol{y}}))$. By definition, $\bar{h}_{ij}$ represents the slope of the tangent line of the non-linear load function, for the interference induced by $\bar{\boldsymbol{y}}$ and $\kappa^\star(\bar{\boldsymbol{y}})$. Using the tangent line and denoting by $\bar{c}_{ij}$ the linear function's value for zero interference, the linear approximation reads $\bar{h}_{ij}w_{ij} + \bar{c}_{ij}$.

Our ILP formulation uses binary variables $x_{ij}$ to represent UE-cell association, which was denoted by $c(j, \boldsymbol{y})$ earlier in Section II. We use $\mu$ to denote the smallest load slack among all cells, and $\ell_{ij}$ to denote the load of cell $i$ due to serving UE $j$. The other sets of variables that have been introduced earlier are $\boldsymbol{y}, \boldsymbol{\rho}$ with $\rho_i = \sum_{j\in\mathcal{J}} \ell_{ij}$, and $\boldsymbol{w}$.

$$\text{max } \mu \tag{5a}$$

$$(2b) - (2c)$$

$$x_{m_j j} + \sum_{i \in \mathcal{N}} x_{ij} = 1, \ j \in \mathcal{J} \tag{5b}$$

$$x_{ij} \leq 1 - \sum_{a \in \mathcal{S}: a \leq r} y_{ia}, \ j \in \mathcal{J}, i \in \mathcal{N}, r \in \mathcal{S} : p_{m_j} g_{m_j j} > p_i g_{ij} r \tag{5c}$$

$$x_{m_j j} \leq 1 - \sum_{a \in \mathcal{S}: a \geq r} y_{ia}, \ j \in \mathcal{J}, i \in \mathcal{N}, r \in \mathcal{S} : p_{m_j} g_{m_j j} < p_i g_{ij} r \tag{5d}$$

$$x_{ij} \leq 2 - \sum_{a \in \mathcal{S}: a \leq r} y_{ia} - \sum_{a \in \mathcal{S}: p_h g_{hj} a > p_i g_{ij} r} y_{ha}, \ j \in \mathcal{J}, i \in \mathcal{N}, h \in \mathcal{N}, r \in \mathcal{S} : i \neq h \tag{5e}$$

$$\rho_i = \sum_{j \in \mathcal{J}} \ell_{ij}, \ i \in \mathcal{C} \tag{5f}$$

$$\rho_i + \mu \leq \rho_i^{max}, \ i \in \mathcal{C} \tag{5g}$$

$$\ell_{ij} \geq \bar{h}_{ij} w_{ij} + \bar{c}_{ij} x_{ij}, \ j \in \mathcal{J}, i \in \{m_j\} \cup \mathcal{N} \tag{5h}$$

$$w_{ij} + \sum_{k \in \mathcal{C} \setminus \{i\}} p_k g_{kj} (1 - x_{ij}) \geq \sum_{k \in \mathcal{C} \setminus \{i\}} p_k g_{kj} \ell_k, \ j \in \mathcal{J}, i \in \{m_j\} \cup \mathcal{N} \tag{5i}$$

$$\boldsymbol{x} \in \mathbb{B}^{(M+N) \times |\mathcal{J}|}, \boldsymbol{y} \in \mathbb{B}^{N \times |\mathcal{S}|}, \boldsymbol{w}, \boldsymbol{\ell} \in \mathbb{R}_+^{(M+N) \times |\mathcal{J}|} \tag{5j}$$

The first two constraint sets have been presented in Section II. By (5b), every UE is served either by the default MC or one of the SCs. The next three sets of inequalities define the relations between $\boldsymbol{x}$ and $\boldsymbol{y}$. For any UE $j$ and SC $i$ for which the received power, biased by $r$, remains weaker than that of the default MC, $j$ is not served by $i$. The result remains for any range $a < r$, as stated by (5c). Inequalities (5d) follow the same line of construction for conditions under which the default MC is not the serving cell. Inequalities (5e) define cell preference among SCs. Consider any UE $j$, two SCs $i$ and $h$, and range $r$. Suppose SC $i$ uses a range not larger than $r$, and cell $h$ adopts any range $n$ leading to higher preference at UE $j$ in relation to SC $i$ with range $r$. Then the two sums of (5e) are ones, and consequently $x_{ij}$ has to be zero, i.e., SC $i$ does not serve UE $j$. By (5c)–(5e), each UE is constrained to be served by the strongest cell, taking into account SC range offsets. Constraints (5f) and (5g) state total cell load and load limit, respectively. Constraints (5h) formulate the linearized load-coupling function. Finally, (5i) defines interference.

Solving (5) leads to a range assignment that is guaranteed to improve demand scaling, if the corresponding slack is positive. By construction, our approach yields a monotonically improving

sequence. The sub-optimality lies in the approximation of $\boldsymbol{f}$ – there may exist range assignment that does have positive slack in respect of $\boldsymbol{f}$, but not for the linear function in (5h).

### C. Relaxation and Bounding

Combining the schemes in Sections III-A and III-B provides a demand scaling factor, henceforth denoted by $\bar{\kappa}$, that is attainable but not necessarily globally optimal. We adapt the ILP scheme to gauge how much room is left for improvement in respect of global optimum. In the sequel, we use $\kappa^\star$, $\boldsymbol{y}^\star$, and $\boldsymbol{\rho}^\star$ to refer to the entities in (2) at global optimum. Denote by $\check{\boldsymbol{\rho}}$ a load vector satisfying $\check{\boldsymbol{\rho}} \leq \boldsymbol{\rho}(\boldsymbol{y}, \bar{\kappa}\boldsymbol{d}), \forall \boldsymbol{y} \in \mathbb{B}^{N \times |\mathcal{S}|}$, i.e., the load for serving demand $\bar{\kappa}\boldsymbol{d}$ is at least $\check{\boldsymbol{\rho}}$ irrespective range assignment. For any UE $j$ and each of the candidate serving cells $i \in \mathcal{N} \cup \{m_j\}$, we define the linear load function $\check{h}_{ij}w_{ij} + \check{c}_{ij}$, obtained via evaluating the non-linear load function with demand $\bar{\kappa}d_j$, for interference induced by $\check{\boldsymbol{\rho}}$ and $\boldsymbol{\rho}^{max}$. That is, $\check{h}_{ij} = (\frac{\bar{\kappa}d_j}{\log_2(1+\frac{p_i g_{ij}}{w_{ij}^{max}+\sigma^2})} - \frac{\bar{\kappa}d_j}{\log_2(1+\frac{p_i g_{ij}}{\check{w}_{ij}+\sigma^2})})/(w_{ij}^{max} - \check{w}_{ij})$, where $\check{w}_{ij} = \sum_{k \neq i} p_k g_{kj}\check{\rho}_k$ and $w_{ij}^{max} = \sum_{k \neq i} p_k g_{kj}\rho_k^{max}$, and $\check{c}_{ij} = \frac{\bar{\kappa}d_j}{\log_2(1+\frac{p_i g_{ij}}{w_{ij}^{max}+\sigma^2})} - \check{h}_{ij}w_{ij}^{max}$. The linear load function leads to the following ILP.

$$\check{\mu}^\star = \max \ \mu$$

$$(2b) - (2c), (5b) - (5g), (5i) - (5j)$$

$$\ell_{ij} \geq \check{h}_{ij}w_{ij} + \check{c}_{ij}x_{ij}, \ j \in \mathcal{J}, i \in \{m_j\} \cup \mathcal{N} \tag{6a}$$

*Theorem 4:* For any $\tilde{\mu} \geq \check{\mu}^\star$, $\kappa^\star \leq \bar{\kappa} \cdot \min_{i \in \mathcal{C}} \rho_i^{max}/(\rho_i^{max} - \tilde{\mu})$.

*Proof:* Consider $\boldsymbol{y}^\star$ and denote its induced load by $\boldsymbol{\rho}(\boldsymbol{y}^\star, \bar{\kappa})$, evaluated with $\boldsymbol{f}$ for demand $\bar{\kappa}\boldsymbol{d}$. By definition of $\check{\boldsymbol{\rho}}$, $\check{\boldsymbol{\rho}} \leq \boldsymbol{\rho}(\boldsymbol{y}^\star, \bar{\kappa})$ and thus $\check{\boldsymbol{h}}$ is a lower approximation of $\boldsymbol{f}$ at $\boldsymbol{\rho}(\boldsymbol{y}^\star, \bar{\kappa})$. Hence $\rho_i^{max} - \rho_i(\boldsymbol{y}^\star, \bar{\kappa}) \leq \check{\mu}^\star, \forall i \in \mathcal{C}$. For $\boldsymbol{y}^\star$ and $\kappa^\star$, at least one cell attains the load limit (see the proof of Theorem 2), implying zero as the minimum slack among the cells with load function $\boldsymbol{f}$. Because $\bar{\kappa} \leq \kappa^\star$, $\rho_i^{max} - \rho_i(\boldsymbol{y}^\star, \bar{\kappa}) \geq 0$, and thus $\check{\mu}^\star \geq 0$. Consider scaling up the demand from $\bar{\kappa}\boldsymbol{d}$ to $\kappa'\boldsymbol{d}$, where $\kappa' = \bar{\kappa}\min_{i \in \mathcal{C}} \frac{\rho_i^{max}}{\rho_i^{max}-\check{\mu}^\star}$, and let $i'$ be the cell for which the minimum is attained. By the form of (1), for cell $i'$, the resulting load with $\boldsymbol{f}$ for range assignment $\boldsymbol{y}^\star$, disregarding any load increase in other cells, is at least $\rho_i^{max}$, because $\rho_i^{max} - \rho_i(\boldsymbol{y}^\star, \bar{\kappa}) \leq \check{\mu}^\star, \forall i \in \mathcal{C}$. The line of argument is clearly valid for any $\tilde{\mu} \geq \check{\mu}^\star$, and the theorem follows. ∎

For an SC, zero load occurs if the SC is not deployed. For the MCs, we take a two-step approach to compute better-than-zero values for $\check{\boldsymbol{\rho}}$. First, we observe that typically some of the UEs are served by MCs, even if all SCs use the maximum range. For these UEs and demand $\bar{\kappa}\boldsymbol{d}$, solving the corresponding equation system of (1) yields load values $\check{\rho}_i, i \in \mathcal{M}$, that, by construction,

bound the minimum load of MCs from below. To strengthen $\check{\rho}_i, i \in \mathcal{M}$, we adopt the ILP below to individually and sequentially maximize the load slack of MCs, starting with the linear load function obtained from $\check{\rho}_i, i \in \mathcal{M}$. Here $\mu_i$ a variable representing the load slack of $i \in \mathcal{M}$.

$$\mu_i^\star = \max \ \mu_i$$
$$(2b) - (2c), (5b) - (5f), (5i) - (5j), (6a)$$
$$\rho_k + \mu_k \leq \rho_k^{max}, \ k \in \mathcal{C} \tag{7a}$$

The above ILP amounts to maximizing the load slack of an individual MC $i$, for the linear load function derived from valid lower bound on MC load. Therefore the load of MC $i$, irrespective of range assignment, is at least $\rho_i^{max} - \check{\mu}_i$. The ILP is solved for each MC. Figure 1 illustrates a typical case of load approximation via setting $\check{\rho}$ by the two-step approach, and compare the result to that from the trivial bound $\check{\rho} = 0$. One can observe that the former is significantly tighter.

## IV. PERFORMANCE RESULTS

We consider a 3GPP LTE HetNet specification in [14]. There are three macro-sites with inter-site distance of 500 meters. The sites are sectored, giving a total of nine MCs with directional antennas. There are eighteen candidate SCs with omni-directional antennas. The total number of UEs is 270, with 30 UEs per MC area. Further parameter specifications are provided in Table II, with an illustration given in Figure 2. For convenience, logarithmic scale is used here for range offset. Offset $-\infty$, corresponding to $0$ in linear scale, represents the choice of not deploying an SC. In order to see the impact of augmenting the candidate offsets on maximum demand scaling, we first consider candidate offset set $\mathcal{S}_1 = \{-\infty, 0, 9\}$dB, and then augment it to $\mathcal{S}_2 = \{-\infty, 0, 3, 6, 9\}$dB. Ten instances of the deployment scenario have been used to obtain performance results.

The main flow of the algorithmic approach is run in MATLAB, that is used to implement as well the normalized fixed-point algorithm in Section III-A. Load minimization in Section III-B is implemented using solver package IBM CPLEX [15]. The tests are performed on a Linux server with 2.6 GHz CPU and 8 GB RAM. Figure 3 displays the results of demand scaling, including the average values and the intervals over the instances. With ten SCs (i.e., $n = 10$), SC range optimization enables to accommodate $60\%$ and $70\%$ more traffic in average, with $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively, in comparison to the baseline value. By deploying all candidate SCs (i.e., $n = 18$). the improvement numbers grow further to $67\%$ and $82\%$, respectively. Thus the effect of deploying range-optimized SCs on network capacity is very significant. From Figure 3(a), expanding the candidate range offsets from $\mathcal{S}_1$ to $\mathcal{S}_2$ yields more improvement than deploying

more SCs, demonstrating the importance of utilizing SC range instead of simply increasing the number of SCs. Indeed, the marginal benefit of allowing more SCs to be deployed tends to decline over $n$. Comparing Figure 3(a) to Figure 3(b), the optimization approach performs very well; in fact, the average gap to the upper bound and hence the worst-case deviation from global optimum is less than $5\%$. Hence there is very little room for further improvement in optimization.

For comparison, we consider two basic settings: $0$ dB (default) and $9$ dB (maximum). The average demand scaling factors by letting all 18 SCs use these two offsets are $1.6$ and $1.3$, respectively. These values are outperformed by deploying only 10 SCs with optimized range assignment (see Figure 3(a)). If all 18 SCs are deployed and their ranges are optimized with $\mathcal{S}_2$, the improvement factors over the basic settings are approximately $16\%$ and $40\%$, respectively.

Figure 4 compares, for a representative instance, cell load distributions for the default range, maximum range, and optimized range, with offset set $\mathcal{S}_2$ and the baseline demand $d_j = 400$ Kbps, $\forall j$. It is apparent that optimized range assignment leads to considerably more balanced load levels.

In HetNets, SC ranges are semi-static and hence they target performance in an average sense over a time period. Indeed, frequently changing cell range generates negative impacts and hence impractical for real networks. Examples of impacts are frequent hand-off due to mobility and ping-pong effect, signaling overhead by more dynamic inter-cell coordination, and unstable cell boundaries. Hence solving range assignment is not intended to run on a short-term basis. Rather, the targeted use, which is coherent with the rationale of the load-coupling model [6], [8], is network planning and re-planning, where snapshots based on knowledge of users (e.g., hotspots) are used for performance evaluation. In this context, an optimization time of minutes suffices. The optimization algorithm consists in normalized fixed-point iterations (Section III-A) and solving (5) (Section III-B). Although fixed-point iterations for general standard interference functions [12] may converge sub-linearly, it is not the computational bottleneck, because most of the effort lies in solving the non-convex problem (5). However, a time limit can be imposed to bound the computational effort, and the approach remains applicable. From an optimization standpoint, scalability is mainly related to the number of candidate range offsets, because the optimization decisions do not involve MCs, and, once the range offsets are set, what remains becomes simple. Indeed, the hardness in Theorem 1 originates from range selection. To examine scalability, we consider the solution time in respect of $|\mathcal{S}|$. For $|\mathcal{S}| = 3$ and $|\mathcal{S}| = 5$, the times are approximately one and three minutes, respectively. Allowing ten candidate offsets in interval $[0, 9]$ with 1 dB spacing, for which the demand scaling is almost identical to that of $|\mathcal{S}| = 5$, the corresponding time is about eleven minutes. In view of the results and given the NP-hardness of the problem, the algorithm scales reasonably well.

## V. CONCLUSIONS

We have investigated SC range optimization in heterogeneous and load-coupled LTE networks. Performance results show substantial improvement over two basic range offset settings, and that the proposed algorithm yields close-to-optimal solutions. Interesting lines for further research include an extensive study of scalability and other performance metrics.

## APPENDIX: PROOF OF THEOREM 1

*Proof:* The recognition version of (2) is to determine whether or not demand $\kappa \boldsymbol{d}$ can be accommodated for given $\kappa$. We construct a polynomial-time reduction from the 3-satisfiability (3-SAT) problem. A 3-SAT instance consists in $b$ boolean variables and $c$ clauses. A literal is a variable or its negation. Each clause is a disjunction of three literals. The task is to determine if there exists a value assignment of the boolean variables, such that all clauses hold true. Consider any irreducible 3-SAT instance in which each variable and its negation together appear at most four times in the clauses. This version of 3-SAT remains NP-complete, see [16]. Let $v_1, \ldots, v_b$ denote the variables and $\bar{v}_1, \ldots, \bar{v}_b$ their negations. We construct an instance of (2) with one MC and $2b$ candidate SCs. Notation $v_i$ and $\bar{v}_i$, $i = 1, \ldots, b$, are reused for the SCs. There are $b + c$ UEs in two groups $\mathcal{J}_b$ and $\mathcal{J}_c$, referred to as boolean UEs and clause UEs, respectively. The $i$th boolean UE is denoted by $u_b^i$. The demand and limit parameters are set as follows: $\kappa = 1$, $\boldsymbol{\rho}^{max} = \boldsymbol{1}$, $n = b$, $d_k = 1$, $\forall k \in \mathcal{J}_b$, and $d_k = 0.05$, $\forall k \in \mathcal{J}_c$. The power and gain values are set to meet the following conditions, with normalized noise $\sigma^2 = 1$. For any SC $v_i$, the received power is $1.0$ for the (up to three) clause UEs in which literal $v_i$ appears in 3-SAT, and the received power is $5.0$ for the boolean UE $u_b^i$. The same line of construction applies to $\bar{v}_i$, $i = 1, \ldots, b$. For the MC, the received power for all clause UEs is $2.0$. The received powers other than the ones specified are negligible. The MC reaches load of $1.0$ for serving some additional UEs (that are not introduced explicitly), for which the received power from any SC is zero. The offset set $\mathcal{S} = \{0, 3\}$ in linear scale.

All the $b + c$ UEs must be served by SCs, as otherwise the MC is overloaded. By construction, any boolean UE $u_i^b$ must be served by either SC $v_i$ or SC $\bar{v}_i$. It then follows from $n = b$ that for any pair of SCs $v_i$ and $\bar{v}_i$, $i = 1, \ldots, b$, one uses range offset 3, and the other uses range offset 0, as otherwise some boolean UE will be unserved. By received power, a clause UE gives preference to the MC. However, if any of the three SCs corresponding to the clause's literals deploys offset 3, cell selection is biased and the MC is no longer the serving cell. Suppose SC $v_i$ is assigned offset 3, implying $\bar{v}_i$ has offset zero and hence zero load. Because a 3-SAT literal appears in at most three clauses, SC $v_i$ will serve, at most, boolean UE $u_b^i$ and three clause UEs. For the former,

there is no interference from $\bar{v}_i$ and thus the required load equals $\frac{1}{\log_2(1+\frac{5}{1})} < 0.39$. For each of the clause UEs, there are at most two interfering SCs corresponding to the other two literals in the clause. For the worst-case load of 1.0 at the two SCs, the load for serving the three clause UEs is at most $\frac{3 \times 0.05}{\log_2(1+\frac{1}{2+2+1})} < 0.58$. Thus any SC, with offset 3, can support its corresponding boolean UE and clause UEs. Therefore, any clause UE is served without causing SC overload, as long as at least one of the three SCs has range offset 3. In conclusion, determining if there is a feasible load solution by range assignment is equivalent to the 3-SAT instance. Hence the recognition version of (2) is NP-complete. Its optimization version is consequently NP-hard. ■

## REFERENCES

[1] X. Chu, D. Lopez-Perez, Y. Yang, and F. Gunnarsson, *Heterogeneous Cellular Networks: Theory, Simulation and Deployment*. Cambridge, UK: Cambridge University Press, 2013.

[2] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3gpp heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, pp. 10–21, Jun. 2011.

[3] A. B. Saleh, Ö. Bulakci, S. Redana, and B. Raaf, "Enhancing lte-advanced relay deployments via biasing in cell selection and handover decision," in *Proc. IEEE PIMRC*, Sep. 2010.

[4] K. Okino, T. Nakayama, C. Yamazaki, H. Sato, and Y. Kusano, "Pico cell range expansion with interference mitigation toward lte-advanced heterogeneous networks," in *Proc. IEEE ICC Workshops*, Jun. 2011.

[5] H.-S. Jo, Y. J. Sang, P. Xia, and J. G. Andrews, "Heterogeneous cellular networks with flexible cell association: a comprehensive downlink SINR analysis," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 3484–3495, Oct. 2012.

[6] A. J. Fehske and G. P. Fettweis, "Aggregation of variables in load models for interference-coupled cellular data networks," in *Proc. IEEE ICC*, Jun. 2012.

[7] I. Siomina, A. Furuskär, and G. Fodor, "A mathematical framework for statistical qos and capacity studies in ofdm networks," in *Proc. IEEE PIMRC*, Sep. 2009.

[8] I. Siomina and D. Yuan, "Analysis of cell load coupling for lte network planning and optimization," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 2287–2297, Jun. 2012.

[9] C.-K. Ho, D. Yuan, and S. Sun, "Data offloading in load coupled networks: solution characterization and convexity analysis," in *Proc. IEEE ICC Workshop on Small Cell Wireless Networks*, Jun. 2012.

[10] K. Majewski and M. Koonert, "Conservative cell load approximation for radio networks with shannon channels and its application to lte network planning," in *Proc. IEEE ANTS*, Dec. 2010.

[11] I. Siomina and D. Yuan, "Load balancing in heterogeneous lte: range optimization via cell offset and load-coupling characterization," in *Proc. IEEE ICC*, Jun. 2012, (Best Paper Award at the IEEE ICC '12 CQRM Symposium).

[12] R. D. Yates, "A framework for uplink power control in cellular radio networks," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 1341–1347, Sep. 1995.

[13] U. Krause, "Concave perron-frobenius theory and applications," *Nonlinear Analysis*, vol. 47, pp. 1457–1466, Aug.. 2001.

[14] 3GPP, "Evolved universal terrestrial radio access (e-utra); further advancements for e-utra physical layer aspects," 3GPP, Tech. Rep. TS 36.814, v.10.0.0, 2010.

[15] "IBM CPLEX Optimizer," 2013. [Online]. Available: http://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/

[16] C. A. Tovey, "A simplified np-complete satisfiability problem," *Discrete Appl. Math.*, vol. 8, pp. 85–89, Jan. 1984.

TABLE I
A SUMMARY OF THE NOTATION.

| Notation | Description |
|---|---|
| $M$ | Number of MCs |
| $N$ | Number of SCs |
| $\mathcal{M} = \{1, \dots, M\}$ | Set of MCs |
| $\mathcal{N} = \{M+1, \dots, M+N\}$ | Set of SCs |
| $\mathcal{C} = \mathcal{M} \cup \mathcal{N}$ | The union set of MCs and SCs |
| $\mathcal{S}$ | Set of candidate range offsets |
| $\mathcal{J}$ | Set of UEs |
| $p_i$ | Power spectral density per resource unit of cell $i$, $i \in \mathcal{C}$ |
| $g_{ij}$ | Total gain between the antenna of cell $i$ and UE $j$, $i \in \mathcal{C}, j \in \mathcal{J}$ |
| $\sigma^2$ | Noise power |
| $n$ | Maximum number of SCs to be deployed |
| $m_j = \operatorname{argmax}_{i \in \mathcal{M}} p_i g_{ij}$ | The default MC of UE $j$, $j \in \mathcal{J}$ |
| $y_{ir}$ | Binary optimization variable indicating whether or not SC $i$ uses range $r$, $i \in \mathcal{N}$, $r \in \mathcal{S}$ |
| $\boldsymbol{y} \in \mathbb{B}^{N \times |\mathcal{S}|}$ | Binary matrix of all $y$-variables |
| $c(j, \boldsymbol{y})$ | The serving cell of UE $j \in \mathcal{J}$, induced by range assignment $\boldsymbol{y}$ |
| $\boldsymbol{\rho} = (\rho_1, \dots, \rho_{M+N})^T \in \mathbb{R}_+^{M+N}$ | Cell load vector |
| $\boldsymbol{\rho}^{max} = (\rho_1^{max}, \dots, \rho_{M+N}^{max})^T \in \mathbb{R}_+^{M+N}$ | Vector of upper limit of cell load levels |
| $B$ | Bandwidth of one resource unit |
| $K$ | Total amount of resource units of the available spectrum and the time for which UE demand is specified |
| $f_i(\boldsymbol{\rho}, \boldsymbol{y}, \boldsymbol{d}) \in \mathbb{R}_+^{M+N} \times \mathbb{B}^{N \times |\mathcal{S}|} \times \mathbb{R}_+^{|\mathcal{J}|} \to \mathbb{R}_+$ | The load of cell $i$ as a function of the cell load levels, the range assignment, and UE demand |
| $\boldsymbol{f} = (f_1, \dots, f_{M+N})^T \in \mathbb{R}_+^{M+N}$ | Vector of load functions |
| $\kappa$ | Demand scaling factor |
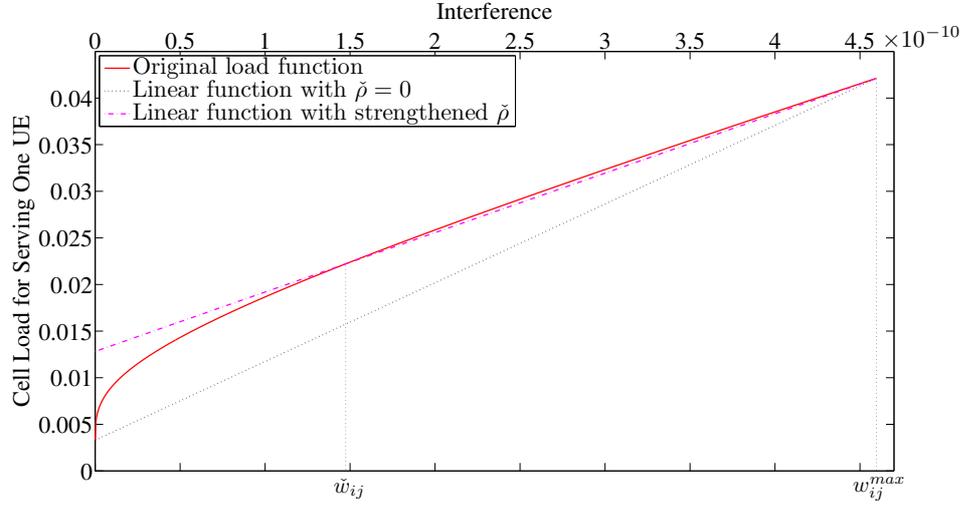
Fig. 1.    An illustration of linear approximation $\check{h}_{ij}w_{ij} + \check{c}_{ij}$ for some UE $j$ and a candidate serving cell $i$.
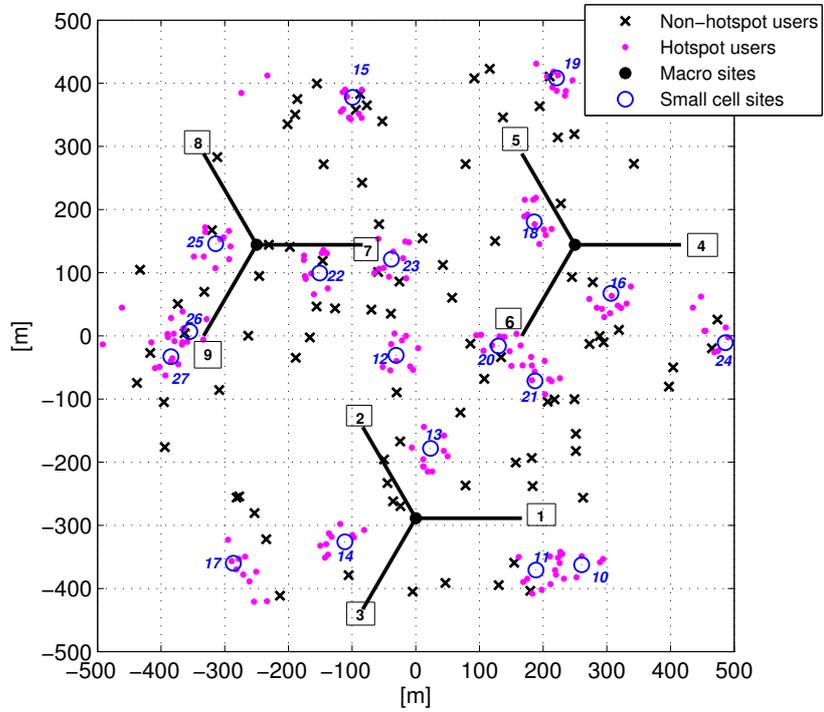


Fig. 2.    The HetNet scenario used for performance evaluation. (MC IDs: 1–9, SC IDs: 10–27.)

TABLE II
PARAMETERS IN THE HETNET SCENARIO.

| Parameter | Value |
|---|---|
| Number of three-sector MC sites (with wrap-around) | 3 |
| MC inter-site distance | 500 m |
| Number of MCs | 9 |
| Number of SCs | 18 |
| Carrier frequency | 2 GHz |
| Bandwidth | 10 MHz |
| Propagation | 3GPP Model 1 [14] (Okumura-Hata, urban, 8 dB standard deviation for shadow fading) |
| MC power | 46 dBm |
| SC power | 30 dBm |
| Noise level | -112 dBm |
| Total number of UEs | 270 |
| Number of generated UEs per MC area (with no SC) | 30 |
| Number of hotspots per MC area | 2 |
| Number of generated UEs per hotspot area | 20 |
| Baseline demand $d$ | 400 Kbps |



(a) $\bar{\kappa}$ from the solution algorithm.

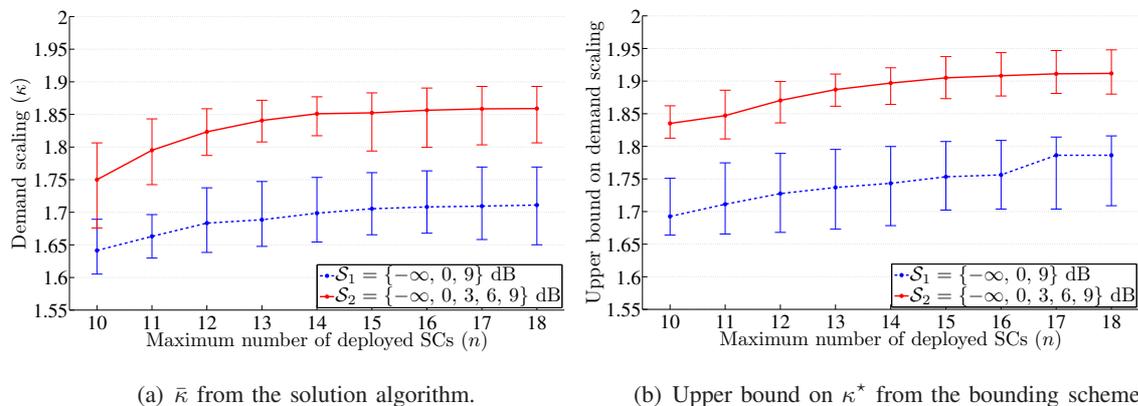(b) Upper bound on $\kappa^\star$ from the bounding scheme.

Fig. 3. Performance results of SC deployment and range optimization. (Average demand scaling with only MCs: 1.03)
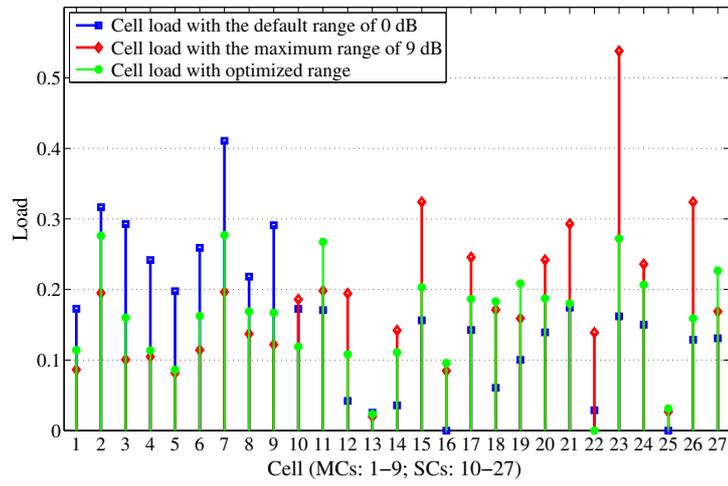
Fig. 4. A comparison of load distribution for the baseline demand.