This document is a preprint of: A. Fehske, H. Klessig, J. Voigt and G. Fettweis, "Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-Organizing Radio Networks," in IEEE Transactions on Vehicular Technology (TVT), vol. 62, no. 5, Jun 2013. DOI:10.1109/TVT.2013.2252474



# Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-Organizing Radio Networks

Albrecht J. Fehske, *Member, IEEE*, Henrik Klessig, *Member, IEEE*, Jens Voigt, and Gerhard P. Fettweis, *Fellow, IEEE*

***Abstract*—Network operators expect a coordinated handling of parameter changes submitted to the operating network's configuration management entity by closed-loop self-organizing network (SON) techniques. For this reason, a major research goal for emerging SON technologies is to achieve coordinated results out of a plethora of independently or even concurrently running use-case implementations. In this paper, we extend current frameworks to compute desirable user associations by an interference model that explicitly takes base-station loads into account. With the aid of this model, we are able to make considerably more accurate estimations and predictions of cell loads compared with established methods. Based on the ability to predict cell loads, we derive algorithms that jointly adapt user-association policies and antenna-tilt settings for multiple cells. We demonstrate by detailed numerical evaluations of realistic networks that these algorithms can be applied to capacity and coverage optimization, mobility load balancing, and cell outage compensation use cases. As a result, rather than performing any heading or tailing coordination, the joint technique inherently comprises all three use cases, making their coordination redundant. For all scenarios studied, the joint optimization of tilts and user association improves quality of service in terms of the fifth percentile of user throughput compared with state-of-the-art techniques. The proposed models and techniques can be straightforwardly extended to other physical and soft parameters.**

***Index Terms*—Capacity and coverage optimization (CCO), cell outage compensation (COC), closed-loop optimization, interference modeling, Long-Term Evolution (LTE), mobility load balancing (MLB), self-optimization, self-organizing networks (SONs), system modeling, wireless network planning and optimization.**

## I. Introduction

THE TERM self-organizing networks (SONs) identifies the next-generation network management technology for planning, optimization, and healing of wireless cellular networks. Although this technology is mainly under discussion for the Long-Term Evolution (LTE) generation of wireless networks within the Third-Generation Partnership Program (3GPP) consortium, the ideas behind SONs will be also adapted for legacy cellular network technologies. The standardization status of SONs with regard to 3GPP release 10 is comprehensively summarized and discussed in [1].

### A. Conceptual Framework for SONs

Pioneering work in the area of SONs was done by the Next Generation Mobile Networks (NGMN) Alliance (see their white papers [2]) and by the European Seventh Framework research project Socrates [3]. The concepts developed within Socrates provide a holistic framework to design SON algorithms and to reveal configuration management (CM) parameter interdependence relations and interactions among different algorithms. Multiple processes can be aggregated to the so-called use cases, which may be independent or may interact since they can operate on common CM parameters. Popular examples of SON use cases for network optimization are mobility load balancing (MLB), capacity and coverage optimization (CCO), and mobility robustness optimization (MRO). Each of these is expected to independently run in a certain deployment area and to address issues related to imbalanced load between cells, coverage holes, low signal-to-interference-and-noise ratios (SINRs), or handover (HO) failures by changing parameters defined in the CM of the cellular network.

### B. SON Use-Case Coordination

These autonomously running SON use-case implementations naturally run into problems of conflicting parameter changes. For that reason, SON coordination is necessary for resolving possible parameter conflicts. The coordination is considered as the most critical challenge to meet, and therefore, coordination mechanisms have to be carefully developed.

Clearly, a key achievement toward the acceptance of a closed-loop implementation of SON use cases (particularly optimization use cases) by network operators is coordinated handling of configuration parameters and their optimization during run time. Consequently, the broad field of coordination of otherwise autonomously running SON use cases enjoys continuing popularity in the wireless community.

The need for use-case coordination was first observed and addressed by the Socrates project. Then, a so-called heading

Manuscript received May 31, 2012; revised October 31, 2012 and April 1, 2013; accepted February 10, 2013. Date of publication March 13, 2013; date of current version June 12, 2013. This work was supported in part by the government of the Free State of Saxony, Germany, within the *Cool Silicon Cluster of Excellence* under Contract 13919/2367 and Contract 14056/2367 and in part by the European Community's Seventh Framework Programme (FP7, 2007–2013) under Grant Agreement 247733 through The EARTH Project. The review of this paper was coordinated by Dr. F. Gunnarsson.

A. J. Fehske, H. Klessig, and G. P. Fettweis are with the Vodafone Chair Mobile Communications Systems, Dresden University of Technology, 01062 Dresden, Germany.

J. Voigt is with Actix R&D, 01067 Dresden, Germany.

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/TVT.2013.2252474

0018-9545/$31.00 © 2013 IEEE

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This document is a preprint of: A. Fehske, H. Klessig, J. Voigt and G. Fettweis, "Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-Organizing Radio Networks," in IEEE Transactions on Vehicular Technology (TVT), vol. 62, no. 5, Jun 2013. DOI:10.1109/TVT.2013.2252474



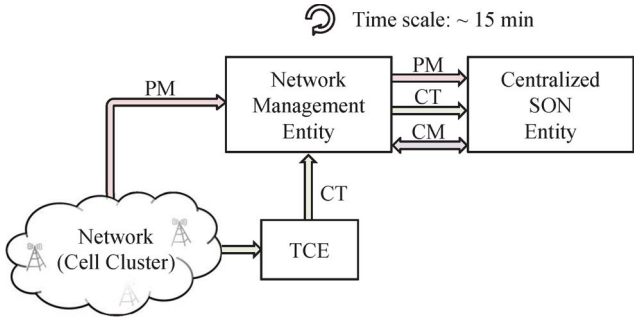


Fig. 1. C-SON system integration.

or tailing coordination of conflicting parameters (separately before or after independently determined parameter changes) was favored. Out of this research project, a general framework for coordination of individual and independent SON use-case implementations was first introduced in [4]. A policy-based approach to SON use-case coordination is given in [5]. This policy-based approach was used for tailing coordination on top of otherwise independently running MLB and MRO use-case implementations in [6].

### C. Integrated SON Approach Developed in this Paper

The approach taken in this paper is different. Instead of having separate and independently running use-case implementations whose outcomes are coordinated afterward (tailing coordination), we address separate use cases into one algorithm and optimize the cellular network toward a joint target. Thus, coordination is inherent in the optimization, avoiding the need for complex policies to coordinate conflicting single use-case implementations.

As just one example, we demonstrate joint handling of the continuously running use cases CCO and MLB for a cluster of cells in an LTE deployment. In addition, we apply the same method to handle an event-triggered use case of cell outage compensation (COC) without any changes. This wide applicability is possible, since the approach presented here targets joint optimization of coverage, user throughput, and load balancing in a cluster of cells.

*1) Scope for Practical Implementation:* SON use cases can be implemented in different architectures, which can be roughly categorized as distributed, hybrid, or centralized. Depending on the architecture, SON use-case implementations are expected to react in different time scales, ranging from milliseconds to hours.

In this regard, the algorithm proposed in this paper caters to a centralized SON (C-SON) architecture, which is attached to an operator's network management entity. Our C-SON method aims to optimize certain CM parameters of a collection of neighboring cells, which we refer to as a cell cluster. Using state-of-the-art hardware and database technology, the time cycle for data exchange between the network management and the C-SON system is typically less than 15 min. Fig. 1 further shows the C-SON system architecture.

*2) Use Cases Addressed:* The proposed method explicitly considers the effect of cell loads on intercell interference and is thus able to predict an optimal network configuration for varying traffic loads or in case of sudden changes in the network configuration, e.g., a base station (BS) power down. Consequently, the proposed algorithm can be used for the aforementioned use cases CCO, MLB, and COC and for compensating other drastic network configuration changes, e.g., for the energy-saving management (ESM) use case.

Since we focus on optimizing the network in time scales of minutes or even hours, we do not consider the MRO use case, which usually deals with HO processes evolving in time scales of milliseconds or seconds.

*3) Multiple-Antenna Systems:* We do not explicitly consider multiple-input–multiple-output (MIMO) specifics, as an extension of CCO toward MIMO deployments is very straightforward. Later in this paper, we shall estimate the achievable data rate based on Shannon's channel capacity in (2), which already includes parameters $a$ and $b$ to account for, e.g., MIMO enhancements of the achievable data rates. Methods to estimate proper values for $a$ and $b$ can be found in the literature. For instance, Voigt *et al.* in [7] suggest that $a$ be estimated as the ratio of Foschini's [8] to Shannon's channel capacity and give guidance on how to calculate a properly normalized $a$ for cross-polarization and antenna array deployments. In addition, in [9], Voigt further separates MIMO gains into power gains due to array types of antennas (reflected by $b$) and into degree-of-freedom (multiplexing) gains (reflected by $a$) and gives guidance on how to estimate $a$ and $b$ in these cases. A mutual-information-based approach can be found in [10].

*4) Uplink:* We will neither include the uplink further than under usual CCO considerations, as capacity limits and load imbalances still mainly occur in the downlink, and, e.g., automated coverage hole detection or random-access channel (RACH) parameter optimization are out of the scope of this paper. In addition, we trust that link budget considerations will be covered outside this algorithm in any comprehensive CCO implementation.

*5) Measurements, Control Parameters, and Constraints:* As stated, we propose a method that combines the MLB and CCO use cases into a joint optimization for the said cluster of cells. The input to the algorithm is measurement data of the network such as performance management (PM) measurement counters of the BSs and measurements of the user equipment (UE), i.e., so-called call traces (CTs), which may be geolocated. These CTs are either provided from the Minimization of Drive Test NGMN/3GPP initiative [11] via a Trace Collection Entity [12] or are proprietary vendor solutions. Further, at least parts of the current CMs of the cell cluster need to be provided to the method from the network management system.

The output of the proposed method is CM parameters, specifically tilt changes for antennas and cell individual power offsets applied in HO algorithms. These parameter changes can be directly submitted to the network's CM.

The downlink maximum transmission power of the cells is not considered in this paper, although it could be an alternative for tilt optimization in a different implementation of the CCO part. In addition, the adjustment of antenna azimuths, beamwidths, or other physical parameters is straightforwardly realizable with the proposed method.

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This document is a preprint of: A. Fehske, H. Klessig, J. Voigt and G. Fettweis, "Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-Organizing Radio Networks," in IEEE Transactions on Vehicular Technology (TVT), vol. 62, no. 5, Jun 2013. DOI:10.1109/TVT.2013.2252474



For the MLB, CCO, and COC use cases, CM parameter settings are supposed to fulfill certain constraints, particularly to assure a minimum pilot or reference signal received power (RSRP) coverage and a maximum cell load for all cells.

*6) Terminology Used:* In addition to the term *cell cluster*, which denotes a collection of neighboring cells, we use the term *user association*, which, in the scope of this paper, comprises the association of UE with a certain cell in active-mode HO and idle-mode cell selection and reselection procedures. The term *cell individual offset* (CIO) is a common identifier for, e.g., the real *Qrxlevminoffset* and *Qqualminoffset* parameters for cell selection [13]; *Qoffset* parameter in cell reselection [13]; and the *Ocs*, *Ocn* or *Ofs*, and *Ofn* parameters for event Ax measurements as preparation for HO procedures [14] in radio resource management (RRM). Such a CIO is thus applied in cell reselection/selection and HO procedures to the received power of the cell's pilot or reference signal with the effect of increasing cells' serving area for the purpose of user association.

Further, the terms *cell* and BS are synonymously used throughout this paper.

### D. System Modeling Techniques

The system model used in this paper considers a dynamic traffic situation where requests arrive to and leave the system over time. Requests are considered as flows, which represent individual data transfers of, e.g., web pages, video, audio, or general data files. Such models are used to model the performance of wireless systems [15]–[18] and are specifically proposed in the context of SONs [19], [20].

The main reason for using such models is to determine the estimates of the average resource utilization of a cell, which we refer to as their *load*. In an orthogonal frequency-division multiplexing (LTE) system, the load thus represents the average fraction of time and frequency resources, i.e., the average number of physical resource blocks, utilized. This load estimation can be compared with real network PM counter measurements (see, e.g., the definition of the "DL PRB Usage for traffic" PM counter in [21]). Load is a desirable quantity to observe, since it determines a variety of measures of the quality of service (QoS) in the network such as delay or throughput.

*1) Dynamic Interference:* Until recently, these flow level techniques did not take the dynamics of intercell interference into account, due to inherent difficulties in solving the underlying queuing-theoretic models (see [22] for details). Typically, a constant interference level is assumed, which is either the maximal possible or some kind of time average. While all BSs in the former case are assumed to perpetually transmit, the latter case assumes that a sufficiently large number of interference scenarios are observed by each data flow. Both assumptions, however, are accurate only in a fully loaded system.

An alternative technique, which we adopt in this paper, has been proposed recently in [23] and independently in [24]. In these load models, intercell interference is simply captured by its time average. As opposed to the latter, the former framework provides an elegant method to numerically compute the cell loads, which we adopt here. The latter framework is applied in [25] for the purpose of load balancing in a heterogeneous network via a CIO given to low-power nodes (small cells). In this application, pure load balancing is considered (i.e., MLB only), and there is no coordination or any other combination with physical base parameter optimization (i.e., for CCO), which we target.

*2) Optimal User Association:* Another line of research, which this paper ties in with, is provided by Kim *et al.* on $\alpha$-optimal user-association policies in [18]. The drawback of the approach in [18] lies, again, in the fact that dynamic intercell interference is not taken into account. In this paper, we extend their theoretical results by precisely this aspect.

From a more practical point of view, their work provides desirable partitioning of a given area into cells, but no CIO results. In addition, it assumes that the UE can take a decision on cell selection based on knowledge of the loads of surrounding BSs. However, in 3GPP, the UE only measure power levels and report them to the BS, where all decisions are taken.

As in real deployments, throughout this paper, traffic intensity is assumed to be nonuniformly distributed over the service region, which extends the applicability of the results to hot-spot scenarios and small cell/heterogeneous (HetNet) deployments.

### E. Paper Outline

This paper is organized as follows: Section II defines the system model, including key performance measures and cell load. Section III states the main theoretical result on the computation of the cell loads and introduces and discusses the optimization problem of interest together with the solution method. Section IV provides simulations to verify the performance of the proposed algorithm and discusses the results. Conclusions are given in Section V. A more detailed description of the implementation of the algorithm is detailed in the Appendix.

This approach was first presented in [26]. We add to our initial presentation by proofing convergence of the iteration used to compute the BS loads and by presenting an additional in-depth analysis of the algorithm's properties.

## II. System Model

This section introduces the mathematical model and assumptions underlying the algorithms developed later in this paper.

### A. Network Layout

Throughout this paper, we consider the downlink of a cellular network consisting of $N$ BSs covering a compact region $\mathcal{R} \subseteq \mathbb{R}^2$. Although BS locations and types can be perfectly arbitrary in principle, we only consider conventional macro BSs in this paper. We assume users to be spatially distributed according to some distribution $\delta(\cdot)$ with $\int_{\mathcal{R}} \delta(u)du = 1$. Users can, in principle, be mobile with certain restrictions, as subsequently explained.

*1) Traffic Model:* Network traffic is modeled on flow level, where flows represent individual data transfers of, e.g., web pages, video, audio, or general data files. We assume that the arrival of flow requests to the network takes place according to a Poisson process with intensity $\lambda$. Flow sizes are assumed

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This document is a preprint of: A. Fehske, H. Klessig, J. Voigt and G. Fettweis, "Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-Organizing Radio Networks," in IEEE Transactions on Vehicular Technology (TVT), vol. 62, no. 5, Jun 2013. DOI:10.1109/TVT.2013.2252474



to be exponentially distributed with common mean $\Omega$. The terms $\lambda$, $\Omega$, and $\delta(u)$ determine traffic intensity distribution $\sigma(u) := \lambda\Omega\delta(u)$ in megabit per second per square kilometer, which we use in the remainder of this paper.

*2) Network Coverage:* For our purposes, we define network coverage based on power received at each location. Let $p_i(u)$ denote the power received from BS $i$ at location $u$. We consider the corresponding coverage region

$$\mathcal{L} := \{u \in \mathcal{R} \,|\, \exists_i : p_i(u) \;\geqslant\; p_{\min}\}$$

where $p_{\min}$ models the minimum receive power required to connect to the network. In practice, $p_{\min}$ is governed by terminals' receiver sensitivity. The degree of network coverage $C$ is then defined as the fraction of users covered, i.e.,

$$C := \int_{\mathcal{L}} \delta(u)\, du.$$

Using standard LTE terminology, we refer to the term $C$ as the RSRP coverage.

We denote the serving area or cell of BS $i$ by $\mathcal{L}_i \subset \mathcal{L}$. The association policy defining $\mathcal{L}_i$ is introduced in Section III. The collection of cells forms a partition of $\mathcal{L}$, which we denote by $\mathcal{P} := \{\mathcal{L}_1, \ldots, \mathcal{L}_N\}$.

### B. Radio Link and Resource Sharing

As previously stated, let $p_i(u)$ denote the power received from BS $i$ at location $u$. The inclusion of all path-loss- and fading-related effects is subsequently discussed. We define the SINR $\gamma_i$ experienced by a data flow at location $u$ with respect to BS $i$ as

$$\gamma_i(u,\eta) = \begin{cases} \frac{p_i(u)}{\sum_{j\neq i} \eta_j p_j(u) + \theta}, & p_i(u) \;\geqslant\; p_{\min} \\ 0, & \text{otherwise} \end{cases} \qquad (1)$$

where $\theta$ denotes noise power. In the given equation, we take into account that a terminal needs to receive a certain minimum signal power $p_{\min}$ to connect to the network. The terms $\eta_j \in [0,1]$ denote the loads of the interfering BSs $j \neq i$. The corresponding achievable data rate is modeled based on the Shannon capacity, i.e.,

$$c_i(u,\eta) = \min\left\{a\, B \log_2\left(1 + b\,\gamma_i(u,\eta)\right),\, c_{\max}\right\} \qquad (2)$$

where $c_{\max}$ denotes the maximum bit rate that is achievable for the system at hand. The purpose of parameters $a$ and $b$ is explained in Section II-B6.

In addition, we require that for any two BSs, the set of locations, where their offered rates are scaled versions of each other, has size zero, i.e., $\int_{\mathcal{L}_{ij}(r)} \delta(u)du = 0$ for all $r > 0$ and all sets $\mathcal{L}_{ij}(r) := \{u \in \mathcal{L} \,|\, c_i(u,\cdot) = r\, c_j(u,\cdot)\}$. This merely technical assumption prevents the algorithm outlined in Theorem 3 from premature convergence.

*1) Fast and Slow Fading:* Here, we presume that serving a data flow takes much longer than the coherence time of a wireless channel, and thus, data flows observe fast fading by its average. Similarly, we presume that shadowing effects happen in a much larger time scale and are constant over the duration of many flows. Consequently, we assume fast- and slow-fading effects to be contained in the location-dependent, but otherwise constant, functions $p_i(\cdot)$.

*2) Antenna Tilts and Antenna Parameters:* Similar to slow-fading effects, the characteristics of the antennas deployed at the BSs in general and the antenna-tilt angles in particular may have a strong effect on the propagation conditions and, thus, on receive power $p_i$. Throughout this paper, we denote the tilt angle of all antennas at BS $i$ by $e_i$ and collect all tilts in vector $e = (e_1, \ldots, e_N)^T$. To keep the notation simple, we omit the dependence of the receive power and all corresponding terms on the tilts, i.e., we write $p_i(\cdot)$ instead $p_i(\cdot, e_i)$.

*3) User Mobility and HOs:* While users can be mobile in principle, we require here that received power $p_i(\cdot)$ does not change during a flow duration. Flow durations are usually rather short periods, e.g., less than a second, and path gains can be assumed to be constant over radii of a few meters [27]. As a consequence, user mobility, although not necessarily zero, is restricted to a few meters per second, i.e., typical pedestrian speed. Such "quasi-stationarity" is fairly realistic, as today, about 80% of all data traffic originates from indoor locations (e.g., [28]).

HO events are commonly triggered by user mobility and the slow-fading process. Since both happen in a much larger time scale than a typical flow duration, here, we do not model HO events explicitly and assume that a flow remains connected to a single serving BS during its entire lifetime.

*4) Effects of Intercell Interference:* Radio link quality is further governed by the interference scenario, i.e., the collection of BSs that are transmitting at any given point in time. In contrast to slow- and fast-fading effects, interference scenarios evolve in the same time scale as the flow dynamics, and as a result, data rates and cell loads of all BSs are strongly interconnected. Accurate modeling of these effects leads to so-called coupled-processor queuing models, which are intractable analytically [22].

To capture the effect of dynamic interference on data rates, we resort to the simpler yet still accurate technique that was independently proposed in [23] and [24]. Rather than modeling the dynamics of interference conditions explicitly, we consider the bitrate as if flows are exposed to *average interference conditions*. According to the underlying M/M/1 PS queuing model, load $\eta_i$ is equivalent to the probability that BS $i$ is transmitting. Consequently, $\sum_{j\neq i} \eta_j p_j(u)$ in (1) denotes the time-averaged interference power.

*5) Sharing of Radio Resources:* Following the approach in [29], we incorporate an *average* packet-scheduling gain into the model via parameters $a$ and $b$, i.e., choosing larger parameters for a more spectrally efficient scheduling mechanism. The main reason for doing so is simplicity. Since fast-scheduling mechanisms explicitly adapt to fast-fading conditions, the approach is justified when flow durations are much longer than the channel coherence time. In this case, each flow experiences the effects of fast fading and fast scheduling only by their averages.

*6) Adapting the Link Model With Parameters $a$ and $b$:* Parameters $a$ and $b$ in (2) are used to further tailor the achievable data rate with a certain SINR $\gamma_i$ and bandwidth $B$ to the

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This document is a preprint of: A. Fehske, H. Klessig, J. Voigt and G. Fettweis, "Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-Organizing Radio Networks," in IEEE Transactions on Vehicular Technology (TVT), vol. 62, no. 5, Jun 2013. DOI:10.1109/TVT.2013.2252474



system under study. The same model is already proposed by Mogensen *et al.* in [29] and used to accurately predict spectral efficiencies of LTE networks based on G-factor distributions. These parameters capture, for instance, the effects of packet scheduling (as previously discussed), MIMO techniques, or system-specific overheads, which individually increase or decrease the average bitrates. In this regard, we can think of products $aB$ and $b\gamma_i$ as the effective bandwidth and effective SINR, respectively.

### C. Average BS Resource Utilization

Given the previous definitions, we define the average resource utilization of BS $i$, i.e., its load, as the integral of load density $\kappa_i(u,\eta) := (\sigma(u))/(c_i(u,\eta))$ over cell area $\mathcal{L}_i$. In this regard, let us define the following function:

$$f_i(\eta) := \min\left\{ \int_{\mathcal{L}_i} \kappa_i(u,\eta)du,\ 1-\epsilon \right\} \tag{3a}$$

with an arbitrarily small $\epsilon > 0$. The operation $\min\{\cdot, 1-\epsilon\}$ is of more technical nature. It becomes necessary when we introduce load-dependent partitions in the following section. Moreover, observe that (3a) only gives an implicit formulation of the cell load: The right-hand side also depends on load vector $\eta$ via the achievable rates $c_i(u,\eta)$. Let $f = (f_1,\ldots,f_N)^T$ denote the vector-valued function with components $f_i$. The load vector of interest is then given as a solution to the system

$$\eta = f(\eta) \tag{3b}$$

i.e., as a fixed point of $f$ in $[0,1)^N$. Note that in (3a), cell areas $\mathcal{L}_i$ are assumed to be fixed, which is an assumption we shall drop in the following section. Further, the load density depends on the antenna tilts via receive power $p_i(\cdot, e)$ (cf. Section II-B2). Thus, interpreting the collection of cell areas $\mathcal{P}$ and antenna tilts $e$ as free variables, the load situation in the network is given as the solution to system $\eta = f(\eta, \mathcal{P}, e)$. In particular, for all tuples $(\mathcal{P}, e)$, system (3) always has a unique fixed point in $[0,1)^N$, and thus, the load is well defined (see [23, Th. 1]).

## III. Joint Optimization of Tilts and User-Association Policies

Based on the definitions in the previous section, here, we introduce the core optimization problem (and several variants of it) and discuss techniques to its solution.

### A. Load as a Proxy for QoS

We observe that the definition of BS loads in (3a) incorporates all system aspects discussed so far, i.e., the division of the region into cells $\mathcal{L}_i$, the location-dependent performance of individual radio links $c_i$, and the actual demand intensity and distribution, which is captured in function $\sigma$.

Moreover, the underlying queuing-theoretic framework suggests that a variety of QoS-related measures are strictly monotonic functions of BS loads $\eta_i$. For instance, in an M/M/1 PS system, the average number of active flows is given by $(\eta_i/1-\eta_i)$. A key measure of QoS is the actual throughput perceived by the users. For the model at hand, the throughput per service request is defined as the size of a flow divided by its sojourn time in the system. An estimate of the time-averaged flow throughput at location $u$ in cell $\mathcal{L}_i$ is given for all locations $u$ by the expression $(1-\eta_i)c_i(u,\eta)$ (see, e.g., [30] and [31] for details), which is likewise monotonic in $\eta_i$. Thus, reducing the load jointly improves all quantiles of the spatial throughput distribution.

### B. Multicell Objective Function

Despite the appealing properties of BS loads, it remains to define a utility function that captures QoS in cases of multiple BSs and, at the same time, lends itself to numerical optimization. Following the work of Mo and Walrand in [32], Kim *et al.* proposed in [18] the use of a parameterized function of the BS loads as an objective function in multicell scenarios, which we also adopt here. Let $\eta_i$ denote the load of BS $i$, as defined in (3a), and let $\alpha \geqslant 0$ denote a nonnegative parameter. We consider the objective function

$$\Phi_\alpha(\eta) = \begin{cases} \sum_{i=1}^N \frac{(1-\eta_i)^{1-\alpha}}{\alpha-1}, & \text{for } \alpha \neq 1 \\ \sum_{i=1}^N -\log(1-\eta_i), & \text{for } \alpha = 1 \end{cases} \tag{4}$$

which is supposed to be minimized. Depending on parameter $\alpha$, function $\Phi_\alpha$ embodies different optimization goals. For $\alpha = 0$, (4) reduces to the sum of the BS loads. Since $\eta_i$ denotes the average resource utilization at BS $i$, $1-\eta_i$ measures the average amount of resources *available* at BS $i$. In this regard, minimizing $\Phi_\alpha$ for $\alpha = 1$ is equivalent to maximizing the geometric mean of the resources available in the network. Further, for $\alpha \to \infty$, minimizing $\Phi_\alpha$ becomes equivalent to minimizing $\max_i \eta_i$, which yields solutions with balanced loads.

Generally, the choice of parameter $\alpha$ in the objective depends on network operators' preferences. For the purposes of this paper, we are particularly concerned with the case $\alpha = 1$ for the following reasons: Objectives $\Phi_\alpha$ for $\alpha \geqslant 1$ to a certain extent prevent overload situations ($\int_{\mathcal{L}_i} \kappa_i(u,\eta)du \geqslant 1$) since they tend to infinity whenever BS loads approach 1. Moreover, we are in favor of minimizing the total system load rather than distributing load equally among cells, which we do not consider as a desirable objective in itself. Taking $\alpha = 1$ satisfies both requirements in the best way possible.

### C. Problem Formulation

Let $e$ and $\mathcal{P}$ denote the vector of antenna-tilt angles and the collection of cell areas, respectively. The optimization we are concerned with in this paper is formulated as

$$\begin{aligned} \underset{\mathcal{P},e}{\text{minimize}} \quad & \Phi_\alpha(\eta) \\ \text{subject to} \quad & \eta = f(\eta,\mathcal{P},e), \quad C \geqslant C_{\min}. \end{aligned} \tag{5}$$

The first condition $\eta = f(\eta,\mathcal{P},e)$ ensures that the objective is evaluated at the solutions to system (3b), where we also consider the dependence of function $f$ on $\mathcal{P}$ and $e$. The second condition ensures that network coverage is satisfactory. Note

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This document is a preprint of: A. Fehske, H. Klessig, J. Voigt and G. Fettweis, "Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-Organizing Radio Networks," in IEEE Transactions on Vehicular Technology (TVT), vol. 62, no. 5, Jun 2013. DOI:10.1109/TVT.2013.2252474



that the second condition does not depend on the particular choice of $\mathcal{P}$.

Surely, problem (5) is considerably complex, due to the implicit formulation of the loads, the rather unorthodox variable $\mathcal{P}$, and the frequently unpredictable effects of antenna tilts on receive power $p_i$ and, thus, the loads themselves. In the following, we discuss the optimization over $\mathcal{P}$ and $e$ separately.

### D. Optimization Over the User-Association Policy

We first focus on minimizing $\Phi_\alpha$ over the collection of cell areas $\mathcal{P}$, assuming tilt vector $e$ to be fixed and such that $C \geqslant C_{\min}$ is satisfied, which yields

$$\begin{aligned} \underset{\mathcal{P}}{\text{minimize}} \quad & \Phi_\alpha(\eta) \\ \text{subject to} \quad & \eta = f(\eta, \mathcal{P}). \end{aligned} \tag{6}$$

For the purposes of this section, consider a general user-association policy, which is defined via functions $q$ taken from the set

$$\mathcal{Q} := \left\{ q : \mathcal{L} \to [0,1]^N \;\middle|\; \forall_{u \in \mathcal{L}} : \sum_{i=1}^{N} q_i(u) = 1 \right\}.$$

The value $q_i(u)$ gives the degree or probability of location $u$ belonging to cell $i$. Consider further the set of loads generated by all possible association policies, i.e.,

$$\mathcal{F} := \left\{ \eta \in (0,1)^N \;\middle|\; \exists_{q \in \mathcal{Q}} \forall_i : \eta_i = \int_{\mathcal{L}} q_i(u) \kappa_i(u,\eta)\, du \right\}. \tag{7}$$

Further, we introduce a key property of the corresponding set $\mathcal{F}$ as follows.

*Property 1 (Full Convertibility):* Let $\kappa_i = (\kappa_1, \ldots, \kappa_N)$ be a vector of load densities defined over some covered region $\mathcal{L}$. Further, let $\mathcal{F}$ denote the set of all possible loads under a generalized association policy, as defined in (7). Set $\mathcal{F}$ is said to have the property of full convertibility if, for all $\eta, \eta' \in \mathcal{F}$, we can find function $\tilde{q} \in \mathcal{Q}$ such that, for all $i$, the following relation holds:

$$\eta_i = \int_{\mathcal{L}} q_i(u) \kappa_i(u,\eta)\, du = \int_{\mathcal{L}} \tilde{q}_i(u) \kappa_i(u,\eta')\, du.$$

Property 1 basically states that loads $\eta \in \mathcal{F}$ need not necessarily be represented as fixed points because the fixed-point representation can be *converted* into a conventional integral, where the integrand does not depend on $\eta$ itself.

*1) Specific User-Association Policies:* To define cell areas with desirable properties, we adopt a user-association policy proposed in [18], which lets users connect to BSs according to the offered data rate and the current load situation. Specifically, location $u$ is associated with BS $i$ if condition $i = s(u,\eta)$ holds, with

$$s(u,\eta) := \underset{j=1,\ldots,N}{\arg\max} (1-\eta_j)^\alpha c_j(u,\eta). \tag{8a}$$

The corresponding cell areas and partition of region $\mathcal{L}$ are given by

$$\mathcal{L}_i(\eta) := \{ u \in \mathcal{L} \,|\, s(u,\eta) = i \} \tag{8b}$$

respectively. Parameter $\alpha$ allows to control the sensitivity of the association rule to load $\eta_i$ of BS $i$. For $\alpha = 0$, user association is purely based on the achievable rate $c_i(u,\eta)$, which depends on the loads of all BSs except BS $i$ [cf. (1)]. For increasing $\alpha$, factor $(1-\eta_i)^\alpha$ forces users to avoid highly loaded BSs even if they provide good achievable rates. In case of ambiguities, users connect to the cell with the lowest index $i$.

*2) Minimizing Properties:* Note that, under the user-association policy in (8a), cell areas $\mathcal{L}_i$ are completely specified by load vector $\eta$. Honoring this fact, let us define the corresponding load functions as

$$\tilde{f}_i(\eta) := \min \left\{ \int_{\mathcal{L}_i(\eta)} \kappa_i(u,\eta)\, du,\ 1-\epsilon \right\} \tag{9}$$

with cell $\mathcal{L}_i(\eta)$ being defined in (8). We state the following result on the optimality of policy (8a) with regard to problem (6).

*Theorem 2:* Let $\mathcal{L}$ denote some coverage region, and let $\kappa$ and $\mathcal{F}$ denote a vector of load densities defined on $\mathcal{L}$ and the corresponding set of loads, respectively. Assume that $\mathcal{F}$ has Property 1. Further, let $\eta^*$ denote a solution to the system

$$\eta = \tilde{f}(\eta) \tag{10}$$

with $\tilde{f} = (\tilde{f}_1, \ldots, \tilde{f}_N)^T$ defined in (9). Then, the corresponding partition $\mathcal{P}(\eta^*) = \{\mathcal{L}_1(\eta^*), \ldots, \mathcal{L}_N(\eta^*)\}$ is a minimizer of problem (6).

The proof of Theorem 2 is provided in Appendix A. Using Property 1, the essential steps follow those in the proof of [18, Th. 1]. Due to the dependence of load density $\kappa$ on the load itself, the structure of set $\mathcal{F}$ in (7) is more intricate compared with that in [18]. In particular, it appears difficult to prove the convexity of $\mathcal{F}$ in our case, even if full convertibility is assumed, so that a simple additional argument is required to prove optimality.

*3) Calculating Cell Areas:* After showing the optimality of the association policy in (8a), we now state a result on how to compute load vector $\eta$ that is generated by this rule.

*Theorem 3:* Consider the load model in (3) where the cell areas are defined according to (8). Further, consider the iteration given by

$$\eta^{(k+1)} := b^{(k)} \tilde{f}\left(\eta^{(k)}\right) + \left(1 - b^{(k)}\right) \eta^{(k)} \tag{11}$$

where $\tilde{f} = (\tilde{f}_1, \ldots, \tilde{f}_N)^T$ is defined in (9), and $b^{(k)} \in [0,1]$ denotes some properly selected step size. Let $\{\eta^{(k)}\}$ denote the sequence generated by iteration (11), then $\{\eta^{(k)}\}$ converges to a fixed point of (10).

Assuming full convertibility and, thus, Theorem 2 to hold, the proof is equivalent to the proof of [18, Th. 1]. In case full convertibility does not apply, the same proof applies, and Theorem 3 still holds; however, we cannot guarantee uniqueness of a fixed point of (10). In this case, it may be possible that

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This document is a preprint of: A. Fehske, H. Klessig, J. Voigt and G. Fettweis, "Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-Organizing Radio Networks," in IEEE Transactions on Vehicular Technology (TVT), vol. 62, no. 5, Jun 2013. DOI:10.1109/TVT.2013.2252474



iteration (11) converges to a local optimum or may have subsequences, which individually converge to separate fixed points with the same function value. The existence of at least one such fixed point, however, is secured. In this regard, condition $\min\{\cdot, 1-\epsilon\}$ in (9) is needed, since for $\eta = 1$, the mapping $\tilde{f}$, which is due to the policy in (8), is not continuous.

### E. Introducing Additional Load Constraints

As stated in Section III-B, using parameter $\alpha \geqslant 1$ in objective function $\Phi_\alpha$ in problem (5) inherently leads to solutions with $\eta_i < 1$, for all $i = 1, \ldots, N$, since $\Phi_\alpha$ grows without bound otherwise. Frequently, however, it is desirable to introduce additional load constraints in the form of $\eta < \hat{\eta}$ for some constraint vector $\hat{\eta} \in (0,1)^N$. The following sections discuss possible ways on how to integrate load constraints into the optimization and how to deal with infeasible constraints.

*1) Integrating Load Constraints Into the Objective Function:* A straightforward approach to deal with additional constraints on the load is to replace all "1" in objective function $\Phi_\alpha$ with $\hat{\eta}_i$, i.e.,

$$\Phi_\alpha(\eta) = \begin{cases} \sum_{i=1}^{N} \frac{(\hat{\eta}_i - \eta_i)^{1-\alpha}}{\alpha - 1}, & \text{for } \alpha \neq 1 \\ \sum_{i=1}^{N} -\log(\hat{\eta}_i - \eta_i), & \text{for } \alpha = 1. \end{cases}$$

Without proof, we state that the corresponding optimal user-association policy can be obtained by equivalently replacing all "1" with $\hat{\eta}_i$ in (8a) and replacing condition $1-\epsilon$ with $\hat{\eta}_i - \epsilon$ in (9).

In this regard, note that operation $\min\{\cdot, \hat{\eta}_i\}$ in (9) renders all solutions with $\eta \geqslant \hat{\eta}$ infeasible, and thus, only strict inequality constraints are meaningful. For obvious reasons, this technique only works for objective functions $\Phi_\alpha$ with parameter $\alpha \geqslant 1$.

*2) Dealing With Infeasible Load Constraints:* Assume that for a given vector of tilts, problem (6) for some $\alpha \geqslant 1$ and some additional load constraints $\eta < \hat{\eta}$ integrated into the objective is infeasible in the sense that there is no feasible partition $\mathcal{P}$, such that constraint $\eta < \hat{\eta}$ is satisfied. Before proceeding with this case, we assert the following property of solutions to the user-association problem (6).

*Assertion 4 (Monotonicity of the Solution):* For a given vector of tilts $e$, let $\eta^*(\alpha)$ be a solution of problem (6) with parameter $\alpha$. Then, there exists $\hat{\alpha}$ such that for $0 \leqslant \alpha, \alpha' \leqslant \hat{\alpha}$, the following relation holds:

$$\alpha \geqslant \alpha' \Rightarrow \max_i \eta_i^*(\alpha) \leqslant \max_i \eta_i^*(\alpha').$$

Although we are not able to provide a proof for the given statement in this paper, we find it to hold in numerical evaluations for a certain range of parameter $\alpha$. Fig. 2 shows one example, where we can observe a monotonic decreasing maximum cell load in the given cluster of cells for increasing parameter $\alpha$ at least for a certain interval.

Assertion 4 suggests that if the optimization in (5) is infeasible for some parameter $\alpha'$, it may be feasible for other parameters $\alpha$ with $\alpha > \alpha'$. To mitigate overload situations as much as possible and in light of Assertion 4, we propose to perform a simple search for a suitable parameter $\alpha$. Let $e$ be a vector of tilts, and let $\eta^*(\alpha, e)$ denote the solution to problem (6) for a given parameter $\alpha$ without any additional load constraints. In case load constraints $\eta < \hat{\eta}$ are given, we select $\alpha$ according to the rule $\alpha = g(e, \hat{\eta})$ with

$$g(e, \hat{\eta}) := \min\{\alpha' \geqslant 1 \,|\, \eta^*(\alpha', e) \leqslant \hat{\eta}\}. \qquad (12)$$

In an actual implementation, the value $\alpha$ needs to be found by repeatedly computing fixed point $\eta = f(\eta, e)$ for different values of $\alpha$.

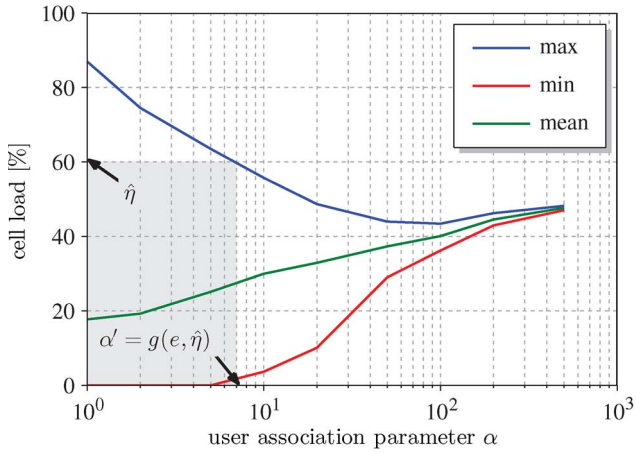


Fig. 2. Impact of user-association parameter $\alpha$ on cell loads for a given tilt configuration.

### F. Optimization Over Tilt Angles

Based on our results in the previous section and assuming that full convertibility holds, we can equivalently reformulate the original optimization in (5) as follows:

$$\begin{aligned} \underset{e}{\text{minimize}} \quad & \Phi_\alpha(\eta) \\ \text{subject to} \quad & \eta = \tilde{f}(\eta, e), \quad C \geqslant C_{\min} \end{aligned} \qquad (13)$$

where we exploit the fact that for any vector of tilts $e$, the load vector minimizing $\Phi_\alpha$ is given as the unique solution of system $\eta = \tilde{f}(\eta, e)$, which is defined in (10).

Unfortunately, in practical systems, no general mathematical model is able to describe the concrete effect of the tilt on the average receive power $p_i(u, e)$ at all locations $u$. Due to complex antenna patterns, varying building types and densities, vegetation, and site-specific characteristics, every propagation scenario is different, and general statements about the existence of globally optimal solutions of problem (13) or methods to compute them cannot be made.

For this reason, we propose resort to heuristic search techniques to optimize over the antenna-tilt angles, such that we achieve at least an improvement over some initially given tilt configuration. In particular, we use the Taxi Cab method, which is a simpler version of Powell's method [33], to search for optima of problem (13). A comprehensive description of the algorithms as used in the numerical evaluations is provided in Appendix B. Here, we summarize the basic steps as follow.

- The algorithm starts with an initial tilt at each BS and a given parameter $\alpha$, which defines objective function $\Phi_\alpha$.

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This document is a preprint of: A. Fehske, H. Klessig, J. Voigt and G. Fettweis, "Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-Organizing Radio Networks," in IEEE Transactions on Vehicular Technology (TVT), vol. 62, no. 5, Jun 2013. DOI:10.1109/TVT.2013.2252474



- The algorithm iterates through the BSs in a certain order (and possibly multiple times). For each BS, a certain set of tilt angles under test is defined based on its current tilt.
- For each BS, the algorithm loops through the set of tilt angles and computes network coverage $C(e)$, load $\eta^* = \tilde{f}(\eta^*, e)$, and objective function value $\Phi_\alpha(\eta^*)$. If the load constraint cannot be fulfilled, other values for parameter $\alpha$ are tested.
- Among all coverage and load-feasible tilts, the tilt minimizing the objective is selected.
- If no feasible tilt can be found, tilt selection is performed among all tilts tested, where priority is given to coverage.

As stated in Section III-D, condition $\eta = \tilde{f}(\eta, e^*)$ yields optimal partition $\mathcal{P}^*$ for any optimizer $e^*$ of problem (13).

### G. Aspects of Practical Implementation

In practical scenarios, the proposed algorithm requires accurate data of average receive power $p_i(\cdot)$ (or propagation conditions) and of traffic demand $\delta(\cdot)$ at all relevant locations. This section briefly discusses how such information may be obtained in practical setups.

*1) RSRP and Location Measurements:* UE locations may be determined by geolocation techniques such as triangulation or by Global-Positioning-System-assisted systems, where the UE reports their coordinates $u$ to their currently associated BSs. In addition, UE shall measure and report long-term received power from several surrounding BSs. In the case of missing receive power data, e.g., for tilt configurations that have not been previously applied during network operation, ray-tracing techniques or drive tests can be utilized to ensure detailed knowledge about the network.

*2) Traffic Measurements:* The mean location-dependent user arrival intensity $\lambda\delta(u)$ and mean flow size $\Omega$ (data amount) may be derived from UE-based measurements. Combining the arrival intensity and mean file sizes yields the spatial distribution of traffic demand $\sigma(u)$. Note that these quantities may highly depend on the time of the day and on the weekday. Long-term assessment of the traffic may be assisted by databases.

*3) Central SON Entity:* Collecting and processing that information requires a central unit, which practically creates a virtual representation of the network and its behavior according to the proposed model and updates it continuously. On the basis of that network representation, several scenarios, e.g., for COC or ESM, can be pretested by this entity, which would not be possible in a distributed SON approach.

*4) Application of Cell Partition:* The adaptation of optimal partition $\mathcal{P}^*$, respectively, of the adjusted user-association policy, can be realized in multiple ways. Since UE at location $u$ shall be permanently associated with a specific BS, this affects initial cell selection, cell reselection, admission control, and HO processes. In the numerical investigations, we transform optimal partition $\mathcal{P}^*$ into standard-compliant CIOs. The corresponding algorithm is detailed in Appendix C.

## IV. Numerical Validation

Here, we present simulation results, utilizing a deployment and propagation model of a real network deployed in a metropolitan area of a large city in North America. We consider three concrete algorithms derived from the model in the previous sections and assess their performance.

TABLE I
Comparison of Characteristics of the Model and Simulations

| | Model | Simulation |
|---|---|---|
| **Traffic** | mean arrival intensity, mean file size per 15x15 m pixel | realizations of Poisson point processes |
| **User association** | optimal partition $\mathcal{P}^*$ | CIOs |
| **User scheduling** | Round Robin | Round Robin |
| **Modulation and coding** | modified Shannon capacity $c_i(u,\eta)$, $a = 0.60$, $b = 0.45$, $c_{\max} = 0.88$ | MCS to achieve 10 % BLER |
| **Link adaptation** | – | CQI reports (every 10 ms + 3 ms round trip time) |
| **Payload overhead** | considered in parameter $a$ | headers up to layer 4, coding, max. 2 retransm., PDCCH and pilots |
| **Cell loads** | calculated acc. $\eta^* = \tilde{f}(\eta^*, e)$ | ratio of occ. and total PRBs |
| **Antenna conf.** | SISO | SISO |

### A. Simulator Details

Simulations are performed by a state-of-the-art system-level simulator for the LTE standard, which implements the full user-plane functionality of LTE release 8 specifications for the PDCP, RLC, MAC, and PHY layers. In addition, most of the corresponding control-plane functions are included in the simulations. The underlying channel model is of the Kronecker type for MIMO scenarios. A variety of fading ITU profiles and user speeds can be chosen. The link layer performance is independently simulated and is applied within the system simulations by mutual-information-based exponential SNR mapping. Table I provides the main differences between the network model used by the algorithm and the settings and procedures implemented in the simulation tool.

### B. Simulation Scenario

The network consists of 18 sites, which are sectorized to 57 cells in total. The BSs are deployed within a region of about $2 \times 2$ km$^2$ and exhibit intersite distances between 300 and 500 m. The area covered is densely populated in most parts. We also consider elevation data of the environment. General parameters of the LTE network are a common bandwidth of 10 MHz at 46-dBm maximum output power and BSs with a single antenna, which are modeled by realistic 3-D patterns. For the proposed algorithm, we consider a minimum RSRP coverage of $C_{\min} = 0.95$, with $p_{\min} = -115$ dBm measured over 10 MHz.

Fig. 3(b) shows the spatially heterogeneous distribution of user arrival intensity $\lambda\delta(u)$, originating from measurements of the corresponding live network at a peak hour. Daily traffic

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This document is a preprint of: A. Fehske, H. Klessig, J. Voigt and G. Fettweis, "Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-Organizing Radio Networks," in IEEE Transactions on Vehicular Technology (TVT), vol. 62, no. 5, Jun 2013. DOI:10.1109/TVT.2013.2252474



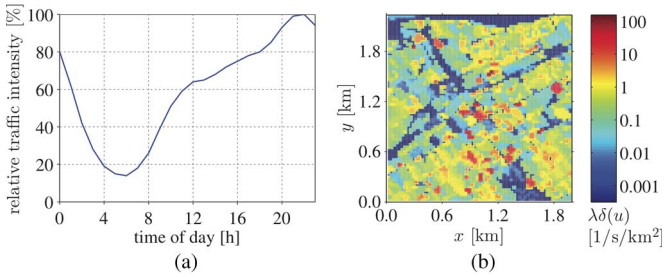


Fig. 3. (a) Traffic intensity throughout a day relative to the peak hour [34] and (b) measured spatial traffic distribution at the peak hour.

variations are taken from [34] [see Fig. 3(a)]. The traffic demand at a peak hour averaged over the scenario region is $\sigma_{\text{peak}} = 20$ Mb/s/km$^2$; the mean flow data size is constant over time and amounts to $\Omega = 200$ kB. As shown, the scenario reveals several hot spots and regions with lower user density in between. Maximum differences in traffic intensity are up to a factor of 6 and $10^5$ in time and space, respectively. For the different use cases, we take a snapshot of the traffic intensity at full hours and simulate at least 180 s of real time for each snapshot.

### C. Algorithms Considered

In the following, we investigate three algorithms.

1) Adaptation of CIOs only (Algorithm **CIO**): This algorithm solves problem (6) with $\alpha = 1$.
2) Joint adaptation of CIOs and tilts (**JOINT1**): This algorithm solves problem (13) with $\alpha = 1$.
3) Joint adaptation of CIOs and tilts with additional load constraint $\eta \leqslant (0.6, \ldots, 0.6)^T$ (**JOINT2**): This algorithm solves problem (13) with load constraints and applies the techniques discussed in Section III-E.

Once computed, optimal partitions $\mathcal{P}^*$ are transformed into CIOs using Algorithm 2, as detailed in Appendix C. Further details are provided in Appendix B.

### D. Reference Simulations

Two use cases are carried out to assess potential performance improvements and gains by joint CIO and tilt adjustment: an event-triggered COC at the peak traffic hour and a continuous CCO/MLB in a 24/7 mode.

For the CCO use case, we consider the following two reference cases:

A) the network in its initial state as configured in reality, i.e., without any adaptation (denoted by **INITIAL**);
B) adaptation of tilts only via a current state-of-the-art CCO optimizer in [35] and [36], which is not load aware (denoted by **TILT**).

For the COC use case, we consider the network in normal operation and optimized by **JOINT1** (denoted by **NORMAL OP**) and the network during outage (denoted by **FAILURE**).

### E. Model Accuracy

For analysis of the model's accuracy, we carry out a snapshot simulation of 60-s duration of the nonoptimized network (**INITIAL**) at 8 P.M. and evaluate the cell loads for five randomly selected cells. The simulated loads are calculated as the mean utilization of physical resource blocks up to the simulation time (see Fig. 4).

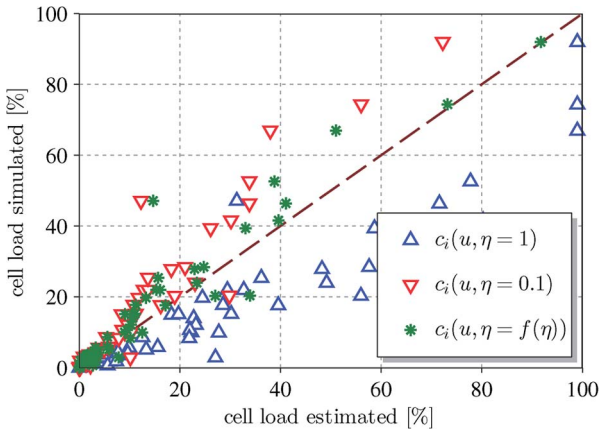


Fig. 4. Cell loads (△) overestimated and (▽) underestimated. (∗) More precise cell load estimation.

Fig. 5(a) shows the convergence of the fixed-point iteration given by (11). It is shown that (11) converges within 13 iteration steps. The comparison of the estimated values of the cell loads, i.e., solution to system (10), and the average resource utilization in Fig. 5(b) reveal the following scenarios.

1) For the given network and simulator configuration, the mean resource utilization already settles after 60 s of simulation.
2) The values obtained by solving system (10) fit the simulation fairly well.

Fig. 4 shows another comparison of the cell loads obtained by simulation and three different flavors of estimated loads, respectively of the fixed point, by considering

1) full interference (△);
2) low interference (▽), i.e., interference intensity that is independent of the BS (which precisely corresponds to the model proposed in [18]);
3) BS-individual interference (loads) (∗) in the computation of the achievable rates $c_i(\cdot, \eta)$.

The accuracy of load estimation highly depends on the interference and rate model that is used. Assuming full interference, one clearly overestimates the cell loads within the network. Assuming too low interference underestimates the cell loads. Highest absolute differences can be observed for highly loaded cells. As shown, the deviation of the estimated values from the simulated values (distance to the purple dashed line) can be significantly reduced by considering BS-individual loads, improving the accuracy of load estimation and affirming the application of the model.

### F. Continuous Use Case: MLB/CCO

We perform snapshot simulations to assess the network's behavior for the traffic load every 2 h during the day. Prior to taking snapshots, all algorithms are triggered to adapt the network to the corresponding traffic situation at that hour of the day. The results are shown in Figs. 6 and 7.

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This document is a preprint of: A. Fehske, H. Klessig, J. Voigt and G. Fettweis, "Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-Organizing Radio Networks," in IEEE Transactions on Vehicular Technology (TVT), vol. 62, no. 5, Jun 2013. DOI:10.1109/TVT.2013.2252474



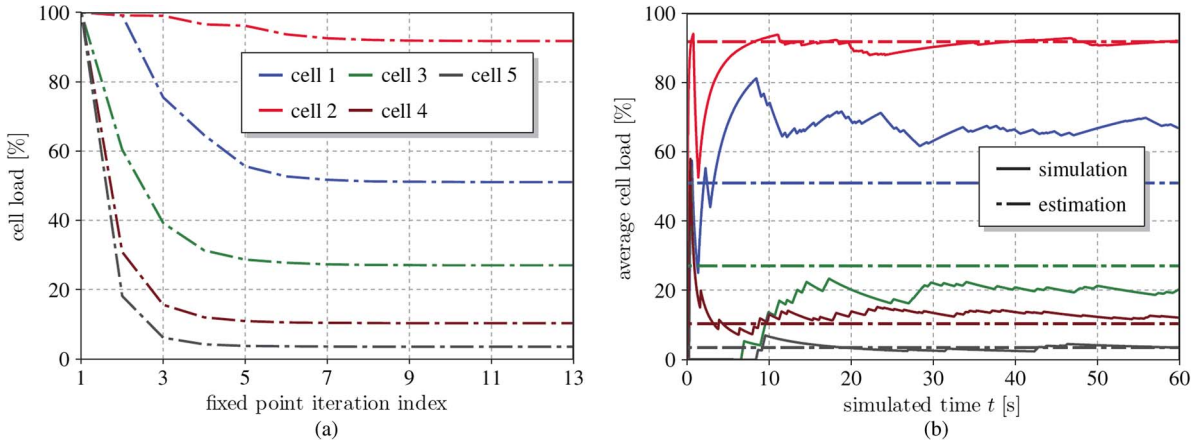


Fig. 5. (a) Convergence of a fixed-point algorithm for randomly selected cells. (b) Comparison of estimated values and resource utilization obtained by simulation and averaged over the interval $[0, t]$.

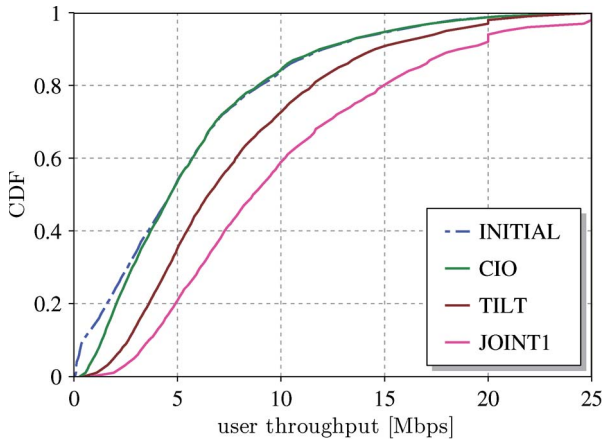


Fig. 6. User throughput cumulative distribution function at 10 P.M.

Applying any of the algorithms presented to the network reduces the maximum cell load in the network. Algorithm **CIO** yields considerable reduction in maximum cell loads not before the load exceeds 50%. Only then that $(1 - \eta_i)$ has considerable influence on the user-association mechanisms. During the peak hour (when loads are rather high), **CIO** improves the user throughput slightly.

In contrast, drastic improvement can be achieved by varying the antenna tilts. The state-of-the-art tilt optimizer **TILT** achieves considerable improvements in user throughput by reducing the interference among the cells (by reducing the cell overlap area given a certain threshold). This yields reduced cell loads due to higher modulation and coding schemes and, thus, to higher user throughput.

An even higher improvement can be achieved by taking the cell load directly into account, which is done by **JOINT1**. According to the objective function, CIOs and antenna tilts are applied, which maximize the sum of logarithms of the BS idle times. Consequently, cell loads are reduced, and highly loaded cells are avoided, which leads to substantial improvements in user throughput, both at 5% and 50%. Interestingly, improvements are also possible during low traffic times and not only at the peak hours.

As stated previously, the presented algorithm indirectly follows the goal of reducing the cell loads. In turn, this means that the average resource utilization is being reduced while serving the traffic offered. Hence, we can observe a certain gain in the average network spectral efficiency throughout the day: **INITIAL**: 0.97 b/s/Hz, **CIO**: 0.93 b/s/Hz (−1.3%), **TILT**: 1.18 b/s/Hz (+21.6%), **JOINT1**: 1.43 b/s/Hz (+47.3%). Algorithm **CIO** reveals slightly reduced spectral efficiency for the sake of improved user throughput at the peak hour, since users are forced to connect to lower loaded cells at the cost of lower rates $c_i$ and, thus, lower spectral efficiency.

### G. Event-Triggered Use Case: COC

We consider the network to be already optimized according to Algorithm **JOINT1**. To create a severe cell outage situation, between 8 P.M. and 11 P.M., we switch off the two cells with the highest load (which also happen to be neighbors). We then observe the capability of different algorithms to compensate the cell outage. Simulation results are presented in Figs. 8 and 9 and are explained as follows.

The cell outage creates very unfavorable load conditions because neighboring BSs have to serve the traffic originally served by the BSs in outage. The maximum resource utilization among the remaining cells increases from around 63% to 100% at the peak hour, i.e., 10 P.M. (see Fig. 8). This overload situation severely degrades the QoS in the network, which is shown by much lower percentiles of the user throughput, as shown in Fig. 9.

The current state-of-the-art network optimizer (denoted by **TILT**) cannot resolve the overload situation in the network (since it is not load aware) and only slightly reduces interference, leading to slightly improved user throughput.

In contrast, optimized CIOs (Algorithm **CIO**) lead to improved user throughput since they ensure that cell loads do not reach 100% (see Section III-E) and avoid congested cells. The maximum CIO being applied amounts to 5.25 dB. However, in this scenario, a 5% user throughput of 1 Mb/s (more often than not a magic number for network operators) cannot be achieved, because the maximum cell load is still rather high with 95%.

Algorithm **JOINT1** (with fixed $\alpha = 1$) is able reduce the maximum cell load to 85%, which leads to an increased 5%

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This document is a preprint of: A. Fehske, H. Klessig, J. Voigt and G. Fettweis, "Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-Organizing Radio Networks," in IEEE Transactions on Vehicular Technology (TVT), vol. 62, no. 5, Jun 2013. DOI:10.1109/TVT.2013.2252474



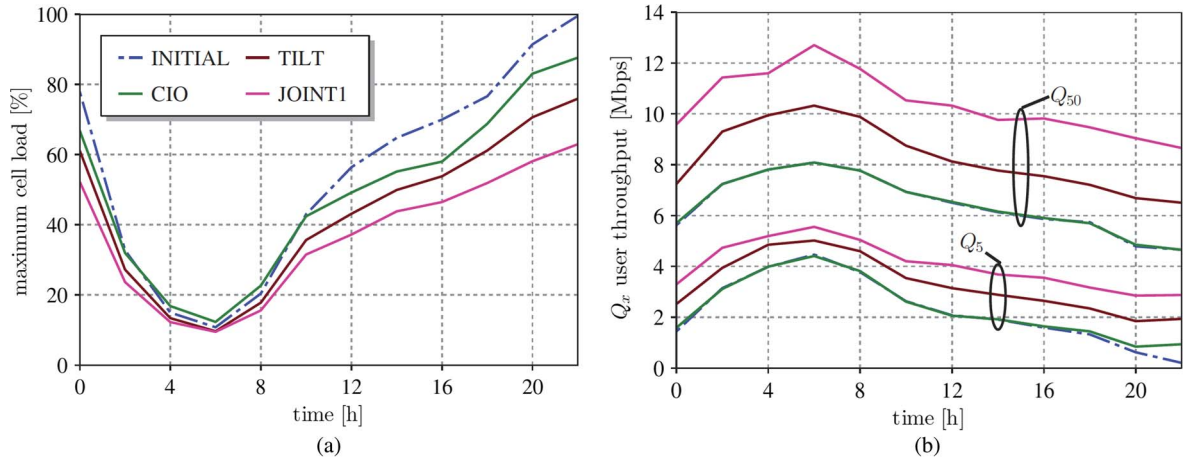


Fig. 7. (a) Maximum cell loads and (b) user throughput 5% and 50% throughout a day. Antenna tilts and CIOs are in their initial configuration or optimized by different algorithms.

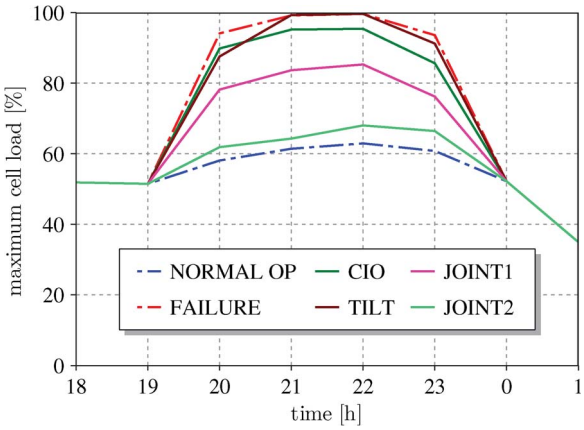


Fig. 8. Maximum cell loads during cell outage.

user throughput of more than 1.4 Mb/s. Thus, the additional "assistance" of tilt adaptation yields substantial improvements in the network's ability to serve the traffic and in providing sufficient QoS. Hence, the maximum CIO being applied is also smaller with 4.25 dB compared with **CIO**.

In both cases, i.e., **CIO** and **JOINT1**, overload is avoided by utilizing $\Phi_1$ as an objective function. One can clearly identify a correlation between the (maximum) cell load and the user throughput (lower quantiles): The higher the load, the lower the throughput, as stated in Section III-A. This leads to the idea of choosing the user-association parameter $\alpha > 1$ and defining a maximum allowed cell load vector of $\hat{\eta} = \{0.6, \ldots, 0.6\}^T$, which is done for **JOINT2**. It can be observed that this goal can almost be achieved (68% at 10 P.M.), almost reaching the value during normal operation (63%). In turn, this yields significant improvement of the 5% user throughput, increased from 0.3 to 2.38 Mb/s close to 2.87 Mb/s (**NORMAL OP**). At 10 P.M. the optimal user-association parameter amounts to $\alpha' = g(e, \hat{\eta}) = 3$, yielding a maximum CIO of 5.75 dB.

## V. Conclusion

In this paper, we have extended current frameworks to compute desirable user associations by an interference model, which explicitly takes BS loads into account. With the aid of these models, we were able to make considerably more accurate estimations and predictions of the cell loads compared with established methods.

Based on the ability to predict the cell loads, we have introduced an algorithm for optimization of user-association policies and antenna-tilt settings. We have shown that the algorithm can be employed on the CCO and MLB SON use cases with no modification.

Instead of performing any external (heading or tailing) coordination of two concurrently running use-case algorithms, it jointly optimizes the aforementioned parameters in a given cluster of cells, which significantly simplifies the process of coordinating different SON use cases. In addition, application of the joint algorithm verifies that it outperforms the optimization of antenna tilts alone (a current state-of-the-art capacity and coverage optimizer) and that it is also better than optimization of user-association policies only, in terms of spectral efficiency and user throughput.

Moreover, the method is also able to compensate for cell outages within the given cluster of cells. Without any use-case-specific changes, the algorithm automatically adapts the antenna tilts and the user-association rules of the remaining cells in the cluster to cope with the overall coverage, user throughput, and cell load thresholds.

The antenna-tilt-setting changes proposed by the algorithm can be directly applied to the configuration management entity of a cellular network. The optimized user-association rule, however, needs to be transformed into possibly air-interface-dependent admission control and HO parameters, i.e., CIOs, to take effect for cell selection, cell reselection, and HO algorithms of the RRM. A corresponding algorithm is provided.

## Appendix A
## Proof of Theorem 2

First, note that problem (6) can equivalently be written as

$$\underset{\eta \in \mathcal{F}}{\text{minimize}} \quad \Phi_\alpha(\eta). \tag{14}$$

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This document is a preprint of: A. Fehske, H. Klessig, J. Voigt and G. Fettweis, "Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-Organizing Radio Networks," in IEEE Transactions on Vehicular Technology (TVT), vol. 62, no. 5, Jun 2013. DOI:10.1109/TVT.2013.2252474



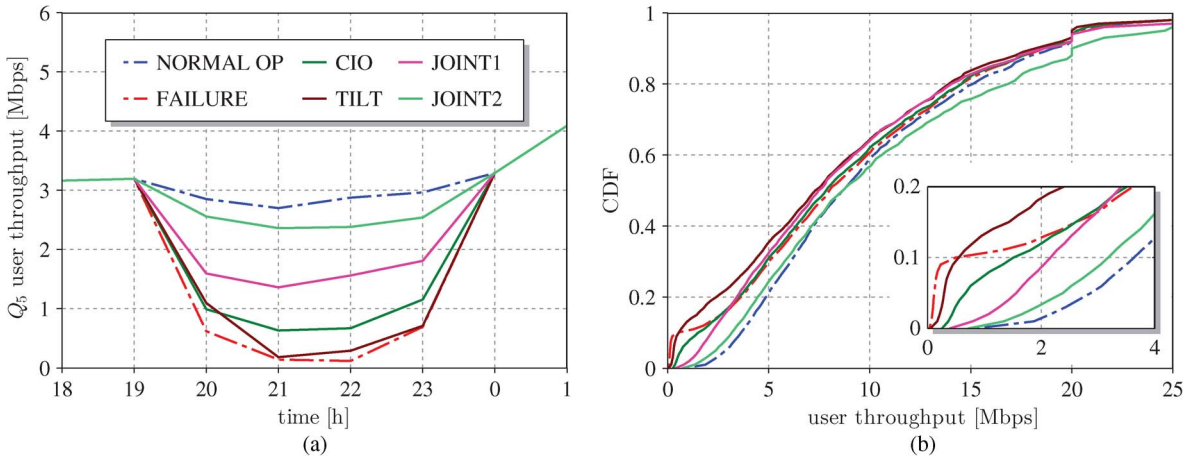


Fig. 9. (a) User throughput 5% during outage and (b) user throughput cumulative distribution function at 10 P.M.

Now, assume that function $\tilde{f}$ has a fixed point, e.g., $\eta^*$, and let $\eta \in \mathcal{F}$ be any other point. Let $q^*$ denote the association policy given by (8a) for $\eta^*$, i.e.,

$$q_i^*(u) = \begin{cases} 1, & \text{if } i = s(u, \eta^*) \\ 0, & \text{otherwise} \end{cases}$$

and let $q \in \mathcal{Q}$ denote any association policy corresponding to $\eta$. Now, we observe the following relation:

$$\begin{aligned} &\nabla\Phi_\alpha(\eta^*)^T(\eta - \eta^*) \\ &= \sum_{i=1}^{N} \frac{\eta_i - \eta_i^*}{(1-\eta_i^*)^\alpha} \\ &= \sum_{i=1}^{N} \frac{\int_{\mathcal{L}} q_i(u)\kappa_i(u,\eta)du - \int_{\mathcal{L}} q_i^*(u)\kappa_i(u,\eta^*)du}{(1-\eta_i^*)^\alpha} \\ &= \sum_{i=1}^{N} \frac{\int_{\mathcal{L}} \tilde{q}_i(u)\kappa_i(u,\eta^*)du - \int_{\mathcal{L}} q_i^*(u)\kappa_i(u,\eta^*)du}{(1-\eta_i^*)^\alpha} \\ &= \int_{\mathcal{L}} \sum_{i=1}^{N} \frac{\tilde{q}_i(u) - q_i^*(u)}{(1-\eta_i^*)^\alpha c_i(u,\eta^*)} \sigma(u) \;\geqslant\; 0. \end{aligned}$$

In the given equation, line two simply follows from $\eta^*, \eta \in \mathcal{F}$. Line three follows from full convertibility, i.e., there exists an association $\tilde{q} \in \mathcal{Q}$ such that each $\eta_i$ can be represented as an integral over $\kappa_i(u, \eta^*)$ instead of $\kappa_i(\cdot, \eta)$. The inequality in the last line follows from the definition of $q^*$, which lets the summation be nonnegative for all $u$.

We further observe that for $\alpha > 0$, function $\Phi_\alpha$ is strictly convex, which implies that the relation $\nabla\Phi_\alpha(\eta^*)^T(\eta - \eta^*) \geqslant 0$, for all $\eta \in \mathcal{F}$, is a necessary and sufficient condition for $\eta^*$ to be the unique optimizer of $\Phi_\alpha$ over set $\mathcal{F}$, that is, if $\mathcal{F}$ is convex. Instead of showing convexity of $\mathcal{F}$ (which appears difficult), consider its convex hull, e.g., $\bar{\mathcal{F}}$, which is a convex set. For $\eta = c\eta' + (1-c)\eta''$, with $\eta', \eta'' \in \mathcal{F}$, we can write

$$\begin{aligned} \nabla\Phi_\alpha(\eta^*)^T(\eta - \eta^*) = c\nabla\Phi_\alpha(\eta^*)^T(\eta' - \eta^*) \\ + (1-c)\nabla\Phi_\alpha(\eta^*)^T(\eta'' - \eta^*) \;\geqslant\; 0. \end{aligned}$$

The given equation follows, since the optimality condition holds for all $\eta', \eta'' \in \mathcal{F}$ and, therefore, for all their convex combinations as well. Since the optimality condition holds for all $\eta$ in $\bar{\mathcal{F}}$, fixed point $\eta^*$ is the sole optimizer of problem (6). For $\alpha = 0$, objective $\Phi_0$ is a linear function of the load vector, and multiple fixed points $\eta^*$ may exist, which yield the same objective function value.

## Appendix B
## Joint Optimization of Tilts and User Association

Algorithm 1 exactly corresponds to Algorithm **JOINT2**, as described in Section IV-C. Algorithm **JOINT1** can be derived from the flowchart by using $\hat{\eta} = 1$ and fixing parameter $\alpha = 1$. Line 14 becomes obsolete.

We start with the initial antenna-tilt setting $e^{(\text{start})}$ and consider the required constraints, i.e., minimum RSRP coverage $C_{\min}$ and the maximum cell load vector $\hat{\eta}$. In addition, we provide the maximum tilt search distance by $e\Delta$. In the following, every BS is touched once in each of $L$ loops. Prior to every loop, the BSs are sorted according their loads $\eta_i$ in descending order (l. 4). Starting to vary the tilts of the BSs with highest loads appears to be beneficial for cell overload resolution and is therefore adopted here.

**Algorithm 1** Joint Optimization of Tilts and User Association

**Input:** $e^{(\text{start})}$, $C_{\min}$, $\hat{\eta}$, $e\Delta \in \mathbb{N}^+$
1: $l := 0$
2: $e := e^{(\text{start})}$
3: **while** $(l < L)$ **do**
4: $\mathcal{B} :=$ sort$(\{1, \ldots, N\})$ acc. $\eta_i$ in descending order
5: **for all** $(i \in \mathcal{B})$ **do**
6: $E_i := \{e_i - e_\Delta, \ldots, e_i - 1^\circ, e_i, e_i + 1^\circ, \ldots, e_i + e_\Delta\}$
7: **for all** $(e_i' \in E_i)$ **do**
8: $e_i := e_i'$
9: calculate $C(e)$, $\alpha = g(e, \hat{\eta})$, and $\eta^* = \tilde{f}(\eta^*, e)$
10: **end for**

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This document is a preprint of: A. Fehske, H. Klessig, J. Voigt and G. Fettweis, "Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-Organizing Radio Networks," in IEEE Transactions on Vehicular Technology (TVT), vol. 62, no. 5, Jun 2013. DOI:10.1109/TVT.2013.2252474



11: $E' := \{e'_i \in E_i | C(e) \geqslant C_{\min}\}$
12: $E'' := \{e'_i \in E_i | C(e) \geqslant C_{\min} \wedge \eta^* \leqslant \hat{\eta}\}$
13: **if** $(E' = \varnothing)$ **then** $e_i := \arg\max_{e'_i \in E_i} C(e)$
14: **else if** $(E'' = \varnothing)$ **then** $e_i := \arg\min_{e'_i \in E'} \sum_{\eta_i^* > \hat{\eta}_i} |\mathcal{L}_i|$
15: **else** $e_i := \arg\max_{e'_i \in E''} \sum_{i=1}^{N} \log(1 - \eta^*)$
16: **end if**
17: **end for**
18: $l := l + 1$
19: **end while**
20: $e^* := e$
21: $\mathcal{P}^* := \mathcal{P}$
22: **return** $\mathcal{P}^*, e^*$

Now, we regard a certain BS $i$ in a certain loop $l$. We vary its antenna tilt $e_i$ in steps of one degree within a maximum range $[e_i - e\Delta, e_i + e\Delta]$ (l. 6). The resulting set of tilt values for BS $i$ is defined as $E_i$.

For each of the tilt settings $e$, where $e \ni e_i = e'_i \in E_i$, RSRP coverage $C(e)$, a proper user-association parameter $\alpha'$ according to Rule (12), and the corresponding cell loads $\eta^*$ according to system (10) are computed (ll. 8–9).

Depending on the network statistics calculated, one of the following actions are performed.

1) If none of the aforementioned tilt settings fulfills the RSRP coverage constraint (cf. l. 11, $E' = \varnothing$), the tilt maximizing the RSRP coverage is applied (l. 13), otherwise.
2) If none of the tilt settings that are feasible in the sense of RSRP coverage fulfills the maximum cell load constraint (cf. l. 12, $E'' = \varnothing$), the tilt minimizing the sum of overloaded cell areas is applied (l. 14) otherwise.
3) The tilt setting that is feasible both in the sense of minimum RSRP coverage and in the sense of the maximum cell load constraint (cf. l. 12, $E'' \neq \varnothing$) and that minimizes $\Phi_1$ is applied.

Note that with this procedure, we ensure that the RSRP coverage constraint has a higher priority than fulfilling the maximum cell load constraint and that, once one of the constraints is fulfilled, it cannot be hurt in the subsequent optimization steps.

Note further that we do not adopt a min–max load policy, if the maximum cell load constraint cannot be fulfilled (no proper $\alpha'$ can be found and none of the tilt settings resolves the overload). Instead, the tilt setting is chosen, which minimizes the sum of the area of overloaded cells (l. 14). Although this is not an obvious choice, it has some practical advantages. Since cell loads tend to decrease for smaller cell areas covering less traffic, it is more likely to resolve overload situations in case multiple neighboring cells exhibit too high loads. Simulations have shown that, applying a min–max load policy, the overloaded cells are more likely to shift the traffic among themselves rather than to outer lower loaded cells. However, the optimality of the proposed policy cannot be necessarily taken for granted.

After the $L$th iteration over all BSs, the antenna-tilt setting applied in the last step is outputted and the corresponding optimal partition $\mathcal{P}^*$ is further transformed into standard-compliant CIOs by Algorithm 2.

**Algorithm 2** Transformation into CIOs

**Input:** $\text{CIO}_{\max}$, $\text{CIO}_\Delta$, $\mathcal{P}^*$
1: $\text{CIO}' := \{-\text{CIO}_{\max}, \ldots, -\text{CIO}_\Delta, 0\,\text{dB}, \text{CIO}_\Delta, \ldots, \text{CIO}_{\max}\}$
2: **for all** $(i \in \mathcal{B})$ **do**
3: **for all** $(j \in \mathcal{B}) \wedge (j > i)$ **do**
4: $\text{CIO}_{i,j} = \arg\min_{\text{CIO}_{i,j} \in \text{CIO}'} A_i$
5: $\text{CIO}_{j,i} := [\text{CIO}_{i,j}]^{-1}$
6: **end for**
7: $\text{CIO}_{i,i} := 1$
8: **end for**
9: $\text{CIO}^* := \text{CIO}$
10: **return** $\text{CIO}^*$

## APPENDIX C
## ALGORITHM: TRANSFORM INTO CIOS

To be compliant with current mobile communications standards, we transform optimal partition $\mathcal{P}^*$ into CIOs. We simplify the HO and cell reselection processes to the condition, where an HO from cell $i$ to cell $j$ is initiated. Since we assume that flow arrivals and service durations happen in much smaller time scales than the users' mobility (quasi-stationary, see Section II-B3), we do not consider hysteresis or time-to-trigger parameters, which are of much more importance, if MRO is considered (not part of this paper).

Let $\text{CIO} := (\text{CIO}_{i,j})^{N \times N}$ be the matrix of CIOs. An HO from cell $i$ to cell $j$ is initiated if

$$p_i(u) \cdot \text{CIO}_{i,j} < p_j(u) \tag{15}$$

where $p_i$ and $p_j$ are the two strongest received signal power at location $u \in \mathcal{L}$. Let

$$v^*(u) := \{i \,|\, u \in \mathcal{L} \in \mathcal{P}^*\} \tag{16}$$

$$v(u, \text{CIO}) := \arg\max_{i,j} \text{CIO}_{i,j} \frac{p_i(u)}{p_j(u)} \tag{17}$$

be the functions that map locations $u$ to BS indexes, according to optimal partition $\mathcal{P}^*$ and to HO user-association rule (15), respectively. Further, let

$$\mathbb{1}_i(x) := \begin{cases} 1, & \text{if } x = i \\ 0, & \text{else.} \end{cases} \tag{18}$$

On that basis, we formulate the determination of CIOs as

$$\text{CIO}^* = \arg\min_{\text{CIO}} \sum_{i \in \mathcal{B}} A_i(\text{CIO}), \text{ where} \tag{19}$$

$$A_i(\text{CIO}) = \int_{u \in \mathcal{L}} \left| \mathbb{1}_i\left(v^*(u)\right) - \mathbb{1}_i\left(v(u, \text{CIO})\right) \right| du \tag{20}$$

denotes the mismatch of cell $i$ according to $\mathcal{P}^*$ and CIO in unit area. Hence, we determine the CIOs, which minimize the sum of areas, where the cells according to $\mathcal{P}^*$ and the cells generated by adjusting CIOs parameters do not intersect, as shown in Fig. 10.

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This document is a preprint of: A. Fehske, H. Klessig, J. Voigt and G. Fettweis, "Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-Organizing Radio Networks," in IEEE Transactions on Vehicular Technology (TVT), vol. 62, no. 5, Jun 2013. DOI:10.1109/TVT.2013.2252474



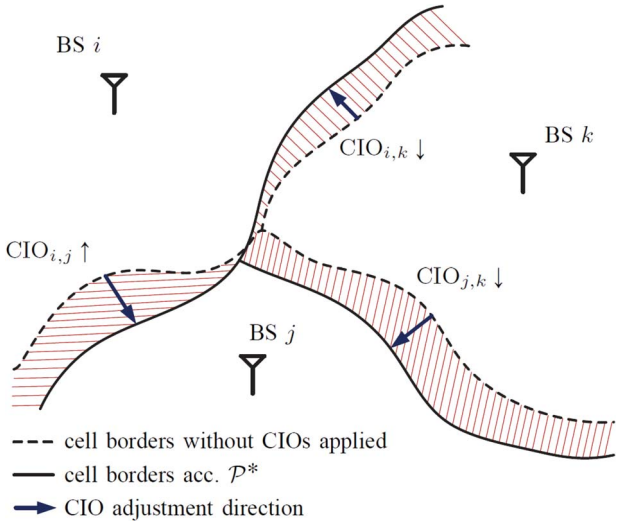


Fig. 10. Minimization of cell overlap.

Equation (19) can be implemented with the help of a directed search similar to that in Algorithm 1. Algorithm 2 shows the implementation, where we define $\mathrm{CIO}_{\max}$ and $\mathrm{CIO}_{\Delta}$ as the maximum CIO value and CIO step size, respectively. For every BS $i$, $\mathrm{CIO}_{i,j}$ is adjusted with respect to all other BSs $j$ and according to the minimization of the mismatch area $A_i$ (l. 4). To ensure a proper user association, the CIOs between each cell pair is reciprocal, i.e., $\mathrm{CIO}_{j,i} = [\mathrm{CIO}_{i,j}]^{-1}$ (l. 5). For completeness, the values in the main diagonal are set to $\mathrm{CIO}_{i,j} = 1$ (l. 7).

## Acknowledgment

The authors would like to thank S. Funck and Dr. C. Isheden of Actix GmbH, Dresden, Germany, for providing and extending the simulator and for providing valuable comments during the internal review process, respectively.

## References

[1] 4G Americas, White Paper Self-optimizing networks: The benefits of SON in LTE 2011, White Paper.
[2] Next Generation Mobile Networks. [Online]. Available: http://www.ngmn.org
[3] SOCRATES web page. [Online]. Available: http://www.fp7-socrates.org
[4] L. Schmelz, M. Amirijoo, A. Eisenblätter, R. Litjens, M. Neuland, and J. Turk, "A coordination framework for self-organisation in LTE networks," in *Proc. IFIP/IEEE Int. Symp. Integr. Netw. Manage.*, Dubin, Ireland, May 2011, pp. 193–200.
[5] T. Bandh, R. Romeikat, H. Sanneck, and H. Tang, "Policy-based coordination and management of self-organizing-network (SON) functions," in *Proc. IFIP/IEEE Int. Symp. Integr. Netw. Manage.*, Dubin, Ireland, May 2011, pp. 827–840.
[6] A. Lobinger, S. Stefanskiy, T. Jansen, and I. Balan, "Coordinating handover parameter optimization and load balancing in LTE self-optimizing networks," in *Proc. IEEE VTC Spring*, Budapest, Hungary, 2011, pp. 1–5.
[7] J. Voigt, R. Fritzsche, and J. Schüler, "Optimal antenna typeSelection in a real SU-MIMO network planning scenario," in *Proc. IEEE VTC Fall*, Anchorage, AK, USA, 2009, pp. 1–5.
[8] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.
[9] J. Voigt, "Next generation cellular network planning and optimization," in *Coordinated Multi-Point in Wireless Communications*, P. Marsch and G. Fettweis, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2011, pp. 432–443.
[10] J. Salo, P. Suvikunnas, H. M. El-Sallabi, and P. Vainikainen, "Some insights into MIMO mutual information: The high SINR case," *IEEE Trans. Wireless Commun.*, vol. 5, no. 11, pp. 2997–3001, Nov. 2006.
[11] Radio measurement collection for minimization of drive tests (MDT), Third-Generation Partnership Project, Sophia-Antipolis, France, 3GPP TS 37.320. [Online]. Available: www.3gpp.org
[12] Subscriber and equipment trace; Trace concepts and requirements, Third-Generation Partnership Project, Sophia-Antipolis, France, 3GPP TS 32.421. [Online]. Available: www.3gpp.org
[13] User equipment (UE) procedures in idle mode, Third-Generation Partnership Project, Sophia-Antipolis, France, 3GPP TS 36.304. [Online]. Available: www.3gpp.org
[14] Radio resource control (RRC) protocol specification, Third-Generation Partnership Project, Sophia-Antipolis, France, 3GPP TS 36.311. [Online]. Available: www.3gpp.org
[15] T. Bonald, S. Borst, and A. Proutiere, "Inter-cell scheduling in wireless data networks," in *Proc. 11th Eur. Wireless Conf.—Next Generation Wireless Mobile Communications Services*, 2005, pp. 1–7.
[16] T. Bonald, S. Borst, and A. Proutière, "Inter-cell coordination in wireless data networks," *Eur. Trans. Telecommun.*, vol. 17, no. 3, pp. 303–312, May/Jun. 2006.
[17] T. Bonald and N. Hegde, "Capacity gains of some frequency reuse schemes in OFDMA networks," in *Proc. IEEE GLOBECOM*, Nov. 2009, pp. 1–6.
[18] H. Kim, G. de Veciana, X. Yang, and M. Venkatachalam, "Distributed $\alpha$-optimal user association and cell load balancing in wireless networks," *IEEE/ACM Trans. Netw.*, vol. 20, no. 1, pp. 177–190, Feb. 2012.
[19] R. Combes, Z. Altman, and E. Altman, "Scheduling gain for frequency-selective Rayleigh-fading channels with application to self-organizing packet scheduling," *Perform. Eval.*, vol. 68, no. 8, pp. 690–709, Aug. 2011.
[20] R. Combes, Z. Altman, and E. Altman, "Self-organization in wireless networks: A flow-level perspective," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 2946–2950.
[21] Performance management (PM); Performance measurements, Third-Generation Partnership Project, Sophia-Antipolis, France, 3GPP TS 32.425. [Online]. Available: www.3gpp.org
[22] T. Bonald, S. Borst, N. Hegde, and A. Proutiére, "Wireless data performance in multi-cell scenarios," in *Proc. Joint Int. Conf. Meas. Model. Comput. Syst. SIGMETRICS*, Jun. 2004, vol. 32, pp. 378–380.
[23] A. J. Fehske and G. P. Fettweis, "Aggregation of variables in load models for cellular data networks," in *Proc. ICC*, 2012, pp. 5102–5107.
[24] I. Siomina and D. Yuan, "Analysis of cell load coupling for LTE network planning and optimization," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2287–2297, Jun. 2012.
[25] I. Siomina and D. Yuan, "Load balancing in heterogeneous LTE: Range optimization via cell offset and load-coupling characterization," in *Proc. IEEE ICC*, Ottawa, ON, Canada, Jun. 10–15, 2012, pp. 1357–1361.
[26] H. Klessig, A. Fehske, J. Voigt, and G. Fettweis, "Improving coverage and load conditions through joint adaptation of antenna tilts and cell selection rules in mobile networks," in *Proc. 2nd IWSON/9th ISWCS*, Paris, France, Aug. 2012, pp. 21–25.
[27] Z. Wang, E. Tameh, and A. Nix, "Joint shadowing process in urban peer-to-peer radio channels," *IEEE Trans. Veh. Technol.*, vol. 57, no. 1, pp. 52–64, Jan. 2008.
[28] Internet Business Solutions Group, Cisco System, Inc., Connected Life Market Watch, 2011.
[29] P. Mogensen, W. Na, I. Z. Kovacs, F. Frederiksen, A. Pokhariyal, K. I. Pedersen, T. Kolding, K. Hugl, and M. Kuusela, "LTE capacity compared to the Shannon bound," in *Proc. IEEE 65th VTC-Spring*, Apr. 2007, pp. 1234–1238.
[30] S. B. Fred, T. Bonald, A. Proutiere, G. Régnié, and J. W. Roberts, "Statistical bandwidth sharing: A study of congestion on flow level," *SIGCOMM—Proc. Conf. Appl., Technol., Architect., Protocols Comput. Commun.*, vol. 31, no. 4, pp. 111–122, Aug. 2001.
[31] A. Kherani and A. Kumar, "Stochastic models for throughput analysis of randomly arriving elastic flows in the Internet," in *Proc. 21st Annu. Joint Conf. IEEE Comput. Commun. Soc.*, 2002, vol. 2, pp. 1014–1023.
[32] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Netw.*, vol. 8, no. 5, pp. 556–567, Oct. 2000.
[33] M. J. D. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *Comput. J.*, vol. 7, no. 2, pp. 155–162, 1964.
[34] G. Auer, O. Blume, V. Giannini, I. Godor, M. A. Imran, Y. Jading, E. Katranaras, M. Olsson, D. Sabella, P. Skillermark, and W. Wajda, "Earth project D2.3—Energy efficiency analysis of the reference systems, areas of improvements and target breakdown," 2011.

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This document is a preprint of: A. Fehske, H. Klessig, J. Voigt and G. Fettweis, "Concurrent Load-Aware Adjustment of User Association and Antenna Tilts in Self-Organizing Radio Networks," in IEEE Transactions on Vehicular Technology (TVT), vol. 62, no. 5, Jun 2013. DOI:10.1109/TVT.2013.2252474



[35] M. Berg, J. Deissner, J. -E. Dietert, J. Hübner, D. Hunold, J. Plehn, P. Schneider, P. Steinkönig, and J. Voigt, "Method and device for optimising cellular wireless communication networks," Eur. Patent EP1 559 289, Aug. 3, 2005.

[36] M. Berg, J. Deissner, J.-E. Dietert, J. Hübner, D. Hunold, J. Plehn, P. Schneider, P. Steinkönig, and J. Voigt, "Method and device for optimising cellular wireless communication networks," U.S. Patent 7 768 968, Aug. 3, 2010.

**Albrecht J. Fehske** (M'12) received the Dipl.Ing. degree from Dresden University of Technology, Dresden, Germany, in 2007. He is currently working toward the Ph.D. degree with the Vodafone Chair Mobile Communications Systems, Dresden University of Technology.

During his studies, he spent one year with the Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, VA, USA, where he worked with the Mobile and Portable Radio Research Group and was involved in research related to machine learning algorithms and signal classification for cognitive radio systems. In 2006, he was with Vodafone Group R&D, Newbury, U.K., where he looked at scheduling algorithms for streaming traffic in high-speed downlink packet access. His current research focuses on energy efficiency modeling and optimization in cellular networks.

**Henrik Klessig** (M'12) received the Diploma from Dresden University of Technology, Dresden, Germany, in 2012.

During his studies, he was with Alcatel-Lucent, Bell Labs, Germany, for a six-month internship, where he was engaged in Long-Term Evolution base station power modeling and research on integrated energy-efficient hardware and resource management solutions for wireless base stations. Since May 2012, he has been with the Vodafone Chair Mobile Communications Systems, Dresden University of Technology. His research interests include cellular heterogeneous networks, self-organizing networks, and energy efficiency optimization.

**Jens Voigt** received the Diploma and the Ph.D. degree from Dresden University of Technology, Dresden, Germany, in 1995 and 2001, respectively.

In 2000, he cofounded Radioplan GmbH (a Dresden-based specialist in cellular network simulation and automatic cell optimization solutions), which was acquired by Actix, Dresden, in 2006. Since 2006, he has been a Principal Research Engineer and a Project Manager in various research and development projects with Actix R&D. He has coauthored more than 30 publications and holds multiple international patent families. With over 15 years of experience in radio-frequency engineering, cellular network technology, optimization algorithm research, and product specification, his research interests include various aspects of self-organizing network (SON) technology, cellular network simulations, and ray-tracing-based wireless channel simulations.

Dr. Voigt is actively involved in several activities of the Next Generation Mobile Networks Alliance regarding SON technology.

**Gerhard P. Fettweis** (F'09) received the Dipl.Ing. and Ph.D. degrees under the guidance of Prof. H. Meyr from Aachen University of Technology (RWTH), Aachen, Germany, in 1986 and 1990, respectively.

From 1990 to 1991, he was a Visiting Scientist with the IBM Almaden Research Center, San Jose, CA, USA. From 1991 to 1994, he was with TCSI Inc., Berkeley, CA, USA, where he was responsible for signal processor development. Since September 1994, he has been the Vodafone Chair with Dresden University of Technology, Dresden, Germany.

Dr. Fettweis has been an Elected Member of the IEEE Solid-State Circuits Society Board (Administrative Committee) since 1999. He was elected as an IEEE Distinguished Lecturer for the Solid-State Circuits Society from 2009 to 2010 and for the IEEE Vehicular Technology Society from 2011 to 2013. He also serves on several supervisory boards and on advisory committees of companies and research institutes. He initiated the Leading-Edge Cluster *Cool Silicon* in 2009 and led it until 2010. Since 2011, he has been leading the Collaborative Research Center *Highly Adaptive Energy-Efficient Computing* at Dresden University of Technology as its Speaker. As the Chairman of Dresden University of Technology's focus area Information Technology and Microelectronics, he coordinates the proposal for a Cluster of Excellence within the German Excellence Initiative *Center for Advancing Electronics Dresden*, addressing different paths to advance electronics beyond current/upcoming limitations of complementary metal–oxide–semiconductor technology.

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.