Mobile Netw Appl
DOI 10.1007/s11036-011-0293-7

Self Organized Network Management Functions for Energy Efficient Cellular Urban Infrastructures

Konstantinos Samdanis · Tarik Taleb · Dirk Kutscher · Marcus Brunner

Abstract Energy efficiency is a significant requirement for the design and management of mobile networks and has recently gained substantial attention from both network operators and the research community. The general concept of energy saving management aims to match the capacity offered by operators to the actual demand at given times and geographic areas. This paper introduces the notion of energy partition, an association of powered-on and powered-off BSs to deliver network-level energy saving. It then elaborates how such concept is applied to perform energy reconfiguration to flexibly re-act to load variations encouraging none or minimal extra energy consumption. A simulation-based study evaluates the performance of the proposed algorithms under different network topologies and traffic conditions, highlights the benefits and drawbacks, and provides recommendations for deployment scenarios.

Keywords 3GPP · energy saving management (ESM) · SON · network management · LTE

K. Samdanis (✉) · T. Taleb · D. Kutscher · M. Brunner
NEC Europe Ltd., Kurfürsten-Anlage 36,
69115 Heidelberg, Germany
e-mail: samdanis@neclab.eu

T. Taleb
e-mail: taleb@neclab.eu

D. Kutscher
e-mail: kutscher@neclab.eu

M. Brunner
e-mail: brunner@neclab.eu

1 Introduction

Energy saving in Information and Communication Technology (ICT) is gaining sustainable attention due to environmental and Operational expenditure (OPEX) reasons. ICT is responsible for nearly 2% of CO_2 emissions [1], consuming 60 billion kilowatt hours, an amount which corresponds to nearly 0.5% of the global electricity consumption [2]. Considering the sector of wireless communications the main power expenditures, accounting 80% of the total energy, are associated with the radio access network and with the operations of Base Stations (BSs) [2]. This energy cost is expected to rise dramatically along with the increase in mobile data forecasted to double every year until 2014 [3], matched by technologies such as Long Term Evolution (LTE) that provides more capacity and higher data rates through the deployment of a larger numbers of smaller cells and extended spatial spectrum multiplexing. In principle there are five levels for reducing the energy consumption of mobile radio networks:

- Improve the energy saving via energy efficient management, optimizing network resources given traffic loads, mobility patterns and location information.
- Improve the efficiency of the network design by optimizing the physical placement of resources, i.e. BSs.
- Optimize the material used in BSs and explore new "green" ones that are energy friendly.
- Develop new technologies based on renewable or alternative sources such as solar, wind or biofuels.

– Increase the social awareness and encourage conservation by rewarding or penalizing users that conserve or are responsible for increasing consumption.

Typical configurations of radio access networks for urban environments serve high capacity demand by utilizing dense BS deployments. In dense configurations certain BSs operate with reduced transmission power forming smaller cells in either overlaid formations where coverage is provided by more than one Radio Access Technology (RAT) or single RAT, e.g. pure LTE. This paper addresses energy saving from a management perspective, permitting certain BSs to reconfigure the transmission power and adjust the antenna tilt/pilot based on load and coverage constraints. Such a re-configuration allows selected overlaid BSs to be powered off once coverage compensation is guaranteed to match the traffic demand at all times. In fact peak hour network utilization differs from the average off-peak providing potential for energy saving. Its magnitude depends on the load difference among peak and off peak times as well as on topology specific characteristics.

Powered-on BSs provide a wider coverage compensating on behalf of the powered-off ones forming a type of energy subset or partition as illustrated in Fig. 1. Energy partitions are associations among powered-on and powered-off BSs defined within the coverage limits of the operating site. Since energy conservation is associated with dynamic traffic conditions, an efficient way for management it is via Self-Organized Network (SON) means [4]. Energy SON functions should be activated when the generic traffic demand is relatively low, a decision based on the aggregate load of several BSs of a wider network region as introduced in [5].

The concept of energy partition advances existing methods by introducing coordination among BSs. In essence, such coordination permits better energy solutions since network elements collectively contribute to energy decisions. In addition, maintaining information related to the association of powered-off and powered-on BSs ensures flexibility for further adjustments based on evolving traffic conditions. Such flexibility allows a dynamic energy re-configuration that balances the increased load among existing partitions allowing the optimal BS to operate instead of powering-on additional ones. Each BS entering an energy saving state cause some overhead related to computation and network update, while extra cost is required to re-power it on. It is therefore essential to determine a minimum period during which each BS may profitably be powered-off.

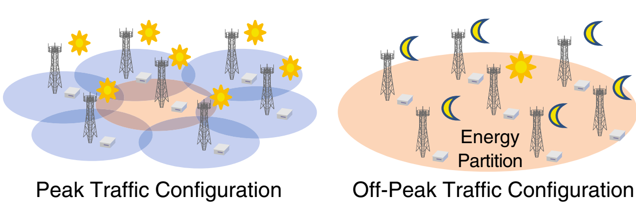


Fig. 1 A simple energy partition example

This paper introduces and evaluates the methods to establish and re-configure energy partitions. Energy partitions may be applied in any type of cellular network. However, the details considered for this study are LTE-specific. The remaining of this paper is organized as follows. Section 2 presents an overview of energy saving management summarizing the related proposals. Section 3 introduces the main concepts and deployment issues, while Section 4 elaborates the algorithms. Section 5 describes the simulation setup and analyzes the results. Finally, Section 6 summarizes our work and provides further research directions.

2 Related work: energy saving management

Managing energy consumption in radio access networks should combine a dynamic operation with network planning. A careful placement of BSs, configuration of transmission power, antenna tilt and other parameters allow for obtaining a minimum number of BSs without compromising coverage. Based on the notion of spectral efficiency per area [6] analyzes energy saving by introducing micro sites to complement conventional ones in various formations. Results show that area power consumption can indeed be reduced without neglecting throughput targets. However, network planning is primarily an off-line activity that is not indented to handle dynamic scenarios. By dynamically identifying redundant capacity in addition to network planning, significantly higher savings can be obtained. Such energy saving potential depends on the specific network layout and traffic distribution.

2.1 Local energy management

Local energy management allows BSs to switch off selected functions, such as radio carriers at off-peak times, based on local measurements. Operator-provided thresholds and policies ensure that coverage and Quality of Service (QoS) requirements are met, and that increasing utilization leads to re-activating appropriate components. A different approach is to operate certain components in energy saving mode. To reduce power consumption in low-traffic periods, [7] proposes energy aware schedulers that queue data in the packet buffer until the varying channel gain becomes sufficiently good. The scheduler also selects the appropriate modulation and coding schemes and controls the power amplifier that is switched off when there is no data to send. Although, local energy management provides adequate savings, decisions on decreasing the energy footprint are based on local observations without considering information from other BSs. Studies such as [1] demonstrate that this would enable much higher savings.

2.2 Overlaid network energy management

When considering network-level energy saving, different scenarios can be identified. In the overlaid network scenario, the redundant coverage in some areas can be realized by local hotspots as illustrated in Fig. 2. A comprehensive overview of the different variants of overlaid energy saving is provided in 3GPP's TR 32.826 [1]. Fundamentally, energy saving is achieved when load conditions allow switching off hotspot cells. Reactivating an overlaid node may be achieved through the macro BS or using a periodic process as described in [8]. Whilst most overlaid energy saving schemes aim to switch off hotspots, it is also possible to take the opposite approach. Green Delay Tolerant Networks (DTN) [9] can achieve energy saving by relaxing coverage and service requirements at off-peak times leveraging disruption-tolerant networking concepts, allowing operators to power-off a number of macro BSs as depicted in Fig. 2c. In overlaid networks, it is relatively easy to identify BSs to power-off without causing radio complications but this scenario, represents only a special case. The generic one is a network with partly redundant coverage, where cells are not fully contained by others.

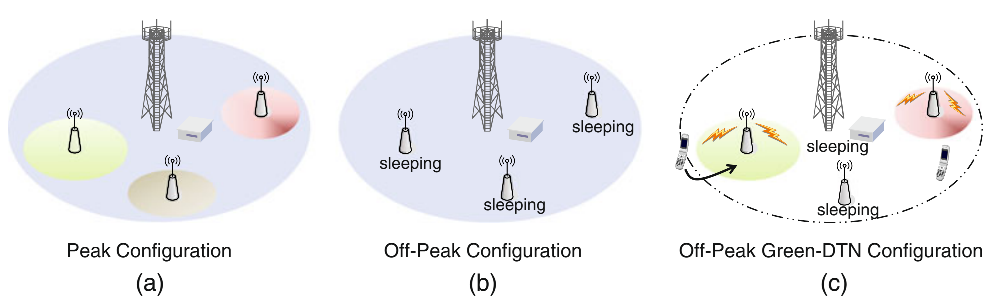
2.3 Non-overlaid network energy management

Generalizing network-level energy saving to non-overlaid scenarios covers a broader range of scenarios, where energy saving requires more than simply switching off specific BSs. If complete coverage should be maintained, powered-on BSs are required to compensate by reconfiguring transmission power and adjusting antenna tilt. Numerical and analytical results verify the potential of the non-overlaid energy saving management in dense network deployments. In particular, the numerical study of [1] provides an estimation of energy saving based on measurements of carrier and air-conditioning expenditures, while the analytical work in [5] focuses on determining the amount of energy conservation considering different cell layouts.

In [10] the concept of cell zooming that blends energy saving and load balancing is described, a study that continues the work of [11]. The proposed centralized algorithm aims to shift traffic towards the highest loaded BSs in a greedy way with the objective to power off as many under-loaded BSs as possible. This algorithm operates on per-user basis, examining whether the process of shifting individual users satisfies load and coverage targets—an activity that raises scalability and precision concerns considering the user information and the sensitivity of user mobility. The distributed algorithm provides energy saving as a handover parameter policy, tuning user handover selection preferences by employing a utility function that associates a higher weight to BSs with higher load concentrating traffic at particular locations allowing lower traffic BSs to "sleep". In a simulation study, the authors compare the two variants, emphasizing the tradeoff between energy saving and coverage, and showing that the centralized algorithm outperforms its distributed counterpart.

A similar distributed energy saving approach based on a combination of a threshold policy with outage detection and SON radio configuration is described in [12]. Once a certain BS is powered off, coverage is

Fig. 2 Overlaid network energy management



established by adjusting the transmission parameters of neighbor sites via supplementary SON functions, while when QoS drops below a limit the reverse process occurs. Another centralized approach introduced in [13] considers a hypothetical static case where a heuristic algorithm with a complete knowledge of stationary user locations aims to associate users to cells solving an assignment problem with network energy saving being its main constraint. Both approaches share the same limitations as cell zooming.

2.4 Energy aware cooperative networks

When considering redundancy in the network infrastructure, it is important to note the duplication of coverage by multiple operators. Cooperation aims at utilizing the combined network resources of different operators and at employing energy saving regimes that work across operator boundaries. An energy conservation management approach, allowing mobile terminals to roam between two different operator networks that overlap the same geographical area, is proposed in [14]. Cooperation among providers is efficient but business burdens may be applied since coverage and infrastructure investments are core operator assets.

3 Energy partition concepts and deployment implications

This section introduces the main concepts beneath our proposal and elaborates the key issues for deploying energy-aware re-configuration including the BS transition process among energy saving states, the handover of multiple users as well as interference control and update of BS neighbor relation.

3.1 Energy partition concepts

Energy partition algorithms are performed once the network enters an off-peak period assuming that BSs are able to enter a stand-by mode with the ability to wake-up when needed. Off-peak periods of low traffic demand ensure stability of the energy saving decisions as demonstrated in [5], which is critical in realizing higher benefits by allowing BSs being profitably into a sleeping mode avoiding temporary local traffic minima that may result in uneconomical solutions. Monitoring mechanisms that decide the triggering time should encounter traffic forecasting methods that consider smoothing to compensate rapid load fluctuations, while hold down timers are also recommended to introduce a minimum spacing among sequential triggers.

3.2 Energy-aware network re-configuration

Energy-aware re-configuration is the most significant feature of the proposed scheme allowing flexibility in balancing the increasing load among existing partitions instead of powering-on additional BSs. Figure 3 illustrates an example where the energy partitions are changing their coverage and even the location of powered-on BS to accommodate load variations keeping energy expenditure constant. In particular, partition A experiences congestion, i.e., load surpasses the predetermined threshold. Based on collected information,

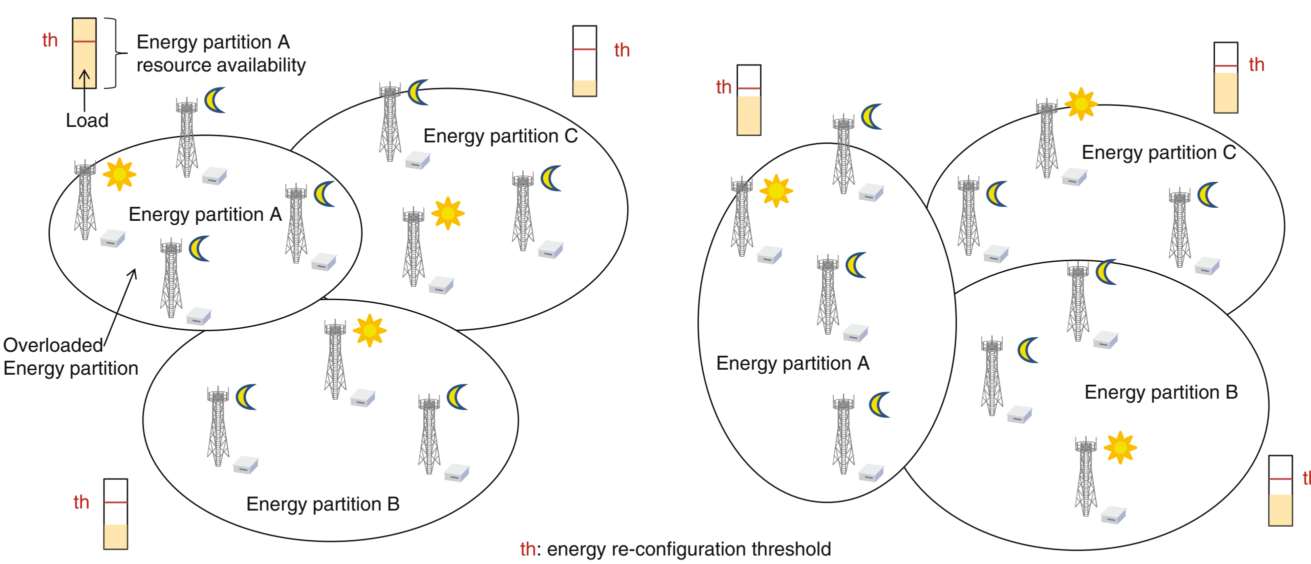


Fig. 3 A simple example of the energy-aware network re-configuration

partition A knows that its neighbors are under-utilized and triggers a re-configuration process to balances the load by re-arranging the energy partitions.

To perform such a re-configuration, methods that collect information need to associate load and coverage with certain powered-off BSs, which is a fundamental difference to the initial energy partition establishment. Powered-on BSs are expected to collect load information regarding powered-off ones using either exact measures or estimations. Exact measures rely on waking up selected BSs at regular intervals to measure the associated load, while estimations use time advance information and UE measurements from adjacent BSs [15]. This paper adopts the estimation option where BSs are able to associate potential hot-spots with the location of powered-off BS inside their energy partition territory.

3.3 BS coverage transition implications

Energy partitions power-off a specified set of BSs allowing certain other BSs to compensate in terms of coverage. Once the energy state of each BS is determined, a transition policy follows with the objective to avoid outage, while ensuring a smooth and cost efficient handover of the affected UEs. The reduction of signaling across the core network is also considered since a higher than expected amount of handovers is likely to occur. The transition process is expected to be performed gradually in a number of steps corresponding to pre-determined power levels. BSs are assumed to be able to switch among such different power levels with BS scheduled to be powered-off decreasing their power until they enter a stand-by mode.

The specified transition policy directs coverage compensating BSs to anticipate increasing their transmission power creating temporal overlapping areas to ensure no outage. Coordination among the involved BSs is essential to manage the transition process effectively. In particular, the coverage compensating BS should increase its transmission power, wait until all camping UEs handover and then signals to the powering-off one to switch into a lower power state. Once such step is completed, feedback may be provided through UEs measurements forming a closed loop control guidance. UEs residing in the overlapping areas handover from the BS that is scheduled to be powered-off towards the coverage compensation one nearly at the same time. To reduce the overhead towards the core network, BSs may handle handover signaling in bulk. It should be noted whilst the bulk signaling handling approach is applicable to any procedure that involves a high number of devices with common fields in the exchanged signaling messages, the following analysis focuses on its benefits in the context of LTE. Indeed, in case of LTE, BSs or eNBs in the 3GPP terminology may hold back the Non-Access Stratum (NAS) signaling messages, i.e. Tracking Area Update (TAU) requests, for the duration of each power step or for a certain number of messages, to proceed with a bulk of NAS message towards the Mobility Management Entities (MMEs).

TAU messages consist of mandatory fields, worth 15 octets and a set of optional fields. Since we concentrate only on UEs that are associated with the same MME, the only parameter which is UE specific, is the M-TMSI (MME Temporary Mobile Subscriber Identities), which identifies a UE at one MME. The other fields, consisting 11 octets out of a total of 15 octets, are common to all UEs associated with the MME. The effort of parsing the parameters of many messages is minimized, decreasing also the time spent for location updates by a large factor, while signaling efficiency is also improved avoiding the processing of multiple messages.

3.4 Interference management and neighbor relations

The process of changing the coverage configuration of neighboring BSs raises concerns regarding interference. Assuming a partially frequency re-use Orthogonal Frequency Division Multiplexing (OFDM) system, we envision two main alternatives to manage interference during transitions. The main idea is to either ensure that UE's frequency blocks are orthogonal or use cooperative transmission.

Ensuring orthogonal frequencies among UEs camping in progressively changing overlapping areas may be a part of a pre-planning activity where a management entity provides hints to neighbor BSs for utilizing certain frequency blocks. Alternatively, coloring algorithms as introduced in [16] may coordinate interference permitting higher resource flexibility. Besides scheduling orthogonal frequency blocks, a different approach is to utilize cooperative communications as described in [17]. Such an approach synchronizes simultaneous transmissions utilizing the same frequency blocks towards UEs ensuring a smooth handover. However, cooperative communications require synchronization procedures in the core and radio access network, which is complex but ensure a better service quality.

When energy partitions or re-configurations are applied, the neighbor relations is not the same and therefore an update is essential. In the context of 3GPP LTE such update needs also to re-establish the X2 interface. The update process may be determined at the same time as the energy arrangement and simply communicated among the coverage compensating BSs. Alternatively, neighbor relations could automatically be verified via the corresponding Automatic Neighbor Relation (ANR) SON function [18].

4 Energy partitions and re-configuration

Once the network enters an off-peak period, energy partition and re-configuration algorithms are employed to determine the specific energy state of individual BSs. Three types of algorithms are considered: (1) centralized, executed at a management entity based on information collected from the complete radio access network, (2) distributed, acting on individual BSs based on neighbor information and (3) local BS-awaking, performed on coverage compensating BSs. Such algorithms rely on load and coverage information, which are exchanged among specific network elements either regularly or on demand depending on the variant. The load indicators considered are those of hardware or UE and radio physical resource block, defined in [4] as (low, medium, high, overloaded). Besides load indicators, load thresholds limiting the shifted load are also introduced as a protection margin.

The wireless cell model employed in this paper assumes that BSs form circular cells employing directional antennas. Each BS is supposed to be capable

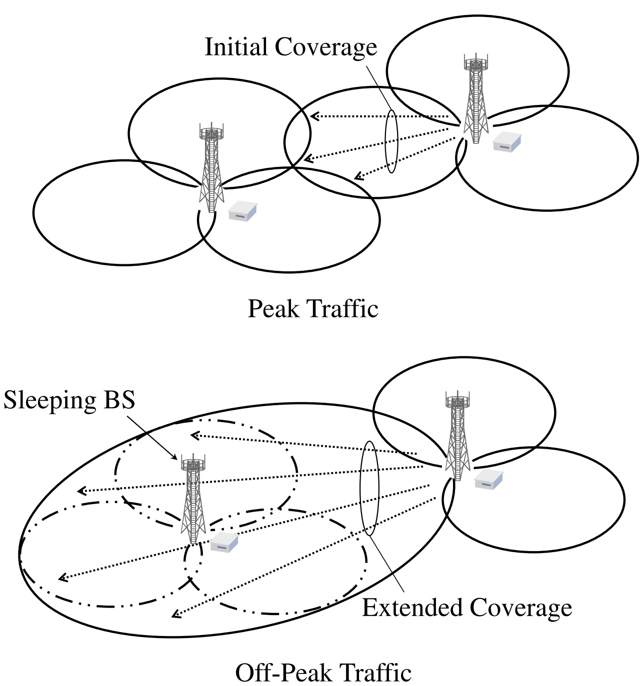


Fig. 4 Cell model and coverage extension

to extend its coverage in a directional manner by re-configuring its transmission power. Outage is avoided by exchanging coverage information among neighbor BSs assuming that they could adjust the antenna tilt/pilot accordingly without encountering any specific location restrictions. A simple cell representation is depicted in Fig. 4 considering a three-sector arrangement providing an overview of the coverage extension.

4.1 Centralized max-load algorithm

Management systems are responsible for executing centralized algorithms, collecting first load information from individual BSs. The proposed centralized algorithm selects the maximum loaded BS and explores its neighbor list starting from the lowest loaded neighbor within its coverage limits. Its purpose is to concentrate the network load into a few moderately loaded BS powering off as many under-loaded BSs as possible. Once the management entity determines the energy formation, it triggers the appropriate changes, which are executed on the corresponding BSs.

Algorithm 1 Centralized Max Load

create shorted Set_{BS} of decreasing order; $S_{\mathrm{EP}} = \emptyset$;
foreach BS $\in Set_{\mathrm{BS}}$
  $p_{\mathrm{on}} = \emptyset$; $S_{\mathrm{off}} = \emptyset$;
  $p_{\mathrm{on}} = \{\mathrm{BS}\}$;
  $S_{\mathrm{BS}} \Leftarrow S_{\mathrm{BS}} \cup p_{\mathrm{on}}$;
  while $\sum_{\mathrm{BS} \in S_{\mathrm{off}}} L_{(\mathrm{UE,R})} < th_L \;\&\; Cv(S_{\mathrm{off}}) < D_{\mathrm{lim}}$
    $\mathrm{BS}_n = min L_{(\mathrm{UE,R})}$;
    break ties based *min* distance;
    $S_{\mathrm{off}} = S_{\mathrm{off}} \cup \mathrm{BS}_n$;
    if $\sum_{\mathrm{BS} \in S_{\mathrm{off}}} L_{(\mathrm{UE,R})} > th_L$
      break;
    else
      $S_{\mathrm{EP}}(p_{\mathrm{on}}) = S_{\mathrm{off}}$;
      update p_{on} neighbor list and $L_{(\mathrm{UE,R})}$;
    end
  end
  $Set_{\mathrm{BS}} \Leftarrow Set_{\mathrm{BS}} \setminus S_{\mathrm{off}}$
end

Algorithm 1 illustrates a pseudo-code version of the centralized algorithm called max load. The algorithm initializes a sorted Set_{BS} of BSs based on maximum load constraints and selects the first one as powered on, p_{on}. For each p_{on}, it keeps track of the set of associated powered off BSs, S_{off} and the energy partitions using the set S_{EP}. As long as the UE and Radio load constraints

$L_{(\mathrm{UE,R})}$ are satisfied and the coverage $Cv(S_{\mathrm{off}})$ is within distance limits D_{lim}, the algorithm centered on the p_{on} continuously selects its minimum load neighbor BS_n, breaking ties based on distance. Otherwise, it selects the next maximum load BS until the Set_{BS} is empty. At each step the algorithm performs the essential updates and upon termination the management entity provides the energy (re-)configuration.

The centralized max load algorithm is deployed using binary heaps for maintaining the neighbor list elements sorted, encountering UE load as the primary constraint and radio resources consumption as a secondary. The computational complexity is $O(1)$ for retrieving the optimal element, $O(logV)$ for updating the binary heap and $O(VlogV)$ for sorting V elements, where V is the number of BSs involved in the process.

4.2 Distributed algorithms

Distributed algorithms rely on regular load and coverage information and on distributed coordination mechanisms that avoid conflict among the decisions of neighbor BSs. The algorithms for establishing energy partitions select a pair and then determine the energy saving state of each BS, while the ones that perform re-configuration iteratively enforce changes among the congested energy partition and its neighbors.

To select a pair, the energy partition algorithm initially activates energy saving on a certain BS, which examines its neighbor list to select the minimum load one breaking ties based on distance. Provided that load and coverage constraints are fulfilled, it proceeds to determine the BS to keep powered-on, otherwise, it terminates on that particular BS. The preference of powered on BSs is towards the ones that seem able to accommodate further energy potential, i.e. towards BSs with a maximum number of neighbor list within the shortest distance. The BS kept powered on performs load and coverage updates and the process of selecting pairs continues until there is no more energy saving improvement. Upon congestion, the problematic partitions, select the least loaded neighbor to perform load balancing by exchanging powered-off BSs. Such process may cause coverage compensating BS to switch-off and awake other BSs instead to change the coverage range and consequently the load balancing potential.

A pseudo-code version of the algorithm is shown in Algorithm 2, with two parts, one for establishing energy partitions and another for re-configuring. In establishing energy partitions, the selected BS, $\mathrm{BS}_{\mathrm{sel}}$ becomes a member of the $pair_{\mathrm{BS}}$ and determines from its neighbor list, n_l, the optimal neighbor BS_n considering load and coverage $Cv(list_n)$, which is added in $pair_{\mathrm{BS}}$. Provided that the UE and radio load $L_{(\mathrm{UE,R})}$ is below the predefined thresholds, the algorithm determines the energy state for each BS, i.e. $p_{\mathrm{on}} - p_{\mathrm{off}}$ and performs the related updates.

Algorithm 2 Distributed Pair-wise Partition Match

if no partitions are established
  $pair_{\mathrm{BS}} = pair_{\mathrm{BS}} \cup BS_{\mathrm{sel}}$;
  while $\sum_{BS \in pair_{\mathrm{BS}}} L_{(\mathrm{UE,R})} < th_L$ & $Cv(n_l) < D_{\mathrm{lim}}$
    $BS_n \in n_l$ with $min L_{(\mathrm{UE,R})}$;
    break ties based *min* distance;
    $pair_{\mathrm{BS}} = pair_{\mathrm{BS}} \cup BS_n$;
    if $\sum_{BS \in pair_{\mathrm{BS}}} L_{(\mathrm{UE,R})} > th_L$
      break;
    end
     $p_{\mathrm{on}} = BS$ with $max\ n_l$; break ties by *min* distance;
     $p_{\mathrm{off}} =$ remaining, other BS;
     p_{on} performs $L_{(\mathrm{UE,R})}$ and $Cv(n_l)$ updates;
   end
  else
    create sorted S_{EP} and $S_{\mathrm{off}} \in EP_c$ based on load;
    while $EP_c L_{(\mathrm{UE,R})} > th_L \& \sum_{S_{\mathrm{EP}}} L_{(\mathrm{UE,R})} < th_L$
      $EP_t \in Set_{\mathrm{EP}}$ with $min\ L_{(\mathrm{UE,R})}$;
      $BS_{\mathrm{ex}} = p_{\mathrm{off}} \in S_{\mathrm{off}}$ with $L_{(\mathrm{UE,R})}$ & $Cv < D_{\mathrm{lim}}$;
      if $(EP_t \cup BS_{\mathrm{ex}}) L_{(\mathrm{UE,R})} < th_L$
        $EP_t = EP_t \cup BS_{\mathrm{ex}}$;
        update EP_c p_{off} association and $L_{(\mathrm{UE,R})}$;
      else
        $S_{\mathrm{off}} = S_{\mathrm{off}} \setminus BS_{\mathrm{ex}}$;
      end
      if $S_{\mathrm{off}} = \emptyset$ {**break**;} **end**
    end
end

Energy re-configuration begins from the congested partition, EP_c provided that at least one neighbor is not overloaded. Initially, the EP_c creates a set S_{EP} of neighbor partitions EP_n in an increasing load order, and a set of its powered-off BS S_{off} in a decreasing load order. The least loaded neighbor partition is then selected as target EP_t, and the most loaded powered-off BS within its coverage area as a candidate to be exchanged $\mathrm{BS}_{\mathrm{ex}}$. As long as the EP_t is not overloaded, the algorithm shifts $\mathrm{BS}_{\mathrm{ex}}$ and performs the appropriate updates. Otherwise $\mathrm{BS}_{\mathrm{ex}}$ is excluded from the S_{off} since its shift is not feasible. The algorithm continues provided that S_{off} is not empty, until a solution is reached or there are no more under-utilized neighbor partitions. If under-utilized neighbor partitions exist but cannot assist the congested one due to coverage limitations, the algorithm powers-off the compensating BS inside the overloaded partition and awakes another one within a closer proximity. Such a process may prove complex though, especially when major changes are introduced due to coverage reasons.

In establishing energy partitions updates are triggered on per pair basis causing excessive overhead. A variant algorithm is introduced to overcome such problem. It selects a powered on BS, which keeps constant forming consecutive pairs with the optimal neighbors. The variants of the distributed algorithm are illustrated in Fig. 5, the first one is referred to as pair-wise match, while the second is dubbed as max neighbor list. The worst-case complexity of distributed algorithms is similar to centralized ones because of the use of binary heaps; the difference is on the scope of operation. Specifically, the complexity of algorithms for establishing initial partitions is $O(NlogN)$, where N denotes the number of BSs, and for re-configuration $O(kMlogk)$, where k is the number of neighbor partitions and M is the BSs inside the overloaded one.

Algorithm 3 BS-Awaking Selection Algorithm

create sorted $\mathrm{Set_{BS}}$ of decreasing load order;
$\mathrm{Set}_{c_{p_{on}}} = \emptyset$;
foreach $\mathrm{p_{off}} \in \mathrm{Set_{BS}}$
  $\mathrm{c}_{p_{on}} = p_{\mathrm{off}}$ max load; $Cv(c_{p_{on}}) = \emptyset$;
  foreach $\mathrm{p_{off}} \in \mathrm{S_{EP}}$
    if $\mathrm{d}(\mathrm{c}_{p_{on}}, \mathrm{p_{off}}) < \mathrm{D_{lim}}$
      $Cv(c_{p_{on}}) = Cv(c_{p_{on}}) \cup \mathrm{p_{off}}$;
    end
  end
  $\mathrm{Set}_{c_{p_{on}}} = \mathrm{Set}_{c_{p_{on}}} \cup Cv(c_{p_{on}})$;
  update sorted $\mathrm{Set}_{c_{p_{on}}}$ of coverage-distance;
  if $Cv(c_{p_{on}})$ contains all $\mathrm{Set_{BS}}$ elements
    break;
  end
end
$\mathrm{p_{on}}$ = highest coverage-distance $\mathrm{c}_{p_{on}} \in \mathrm{Set}_{c_{p_{on}}}$;

4.3 BS-awaking selection algorithm

In case the re-configuration process fails, an essential approach is to awake a BS to accommodate the increased load. Although waking up any BS may improve the network performance balancing the load of the congested partition, poor planning might lead to temporary solutions, which may cause further BS awaking and energy expenditure. The aim is to select a BS within the problematic region taking into account flexibility, considering the potential for further coverage adjustments upon additional load variations. For this reasons the BSs with the highest coverage potential are preferred.

A pseudo-code version of the BS-awake selection algorithm is shown in Algorithm 3. The algorithm begins considering only the powered-off BSs associated with the hotspot, Set_{BS}, which are sorted based on load. Each member of the Set_{BS} is a candidate BS to be powered-on, $c_{p_{on}}$, assessed in terms of load, coverage and distance encountering the complete overloaded partition S_{EP}. As long as the distance between the $c_{p_{on}}$ and p_{off} satisfies the coverage limits, D_{lim}, the p_{off} is stored in a temporary variable $Cv(c_{p_{on}})$ that keeps track of the coverage and distance relations. Distance is considered as an additional energy metric because it indicates the transmission power for providing coverage.

The $Cv(c_{p_{on}})$ information is then used to create a sorted set, $Set_{c_{pon}}$ that lists candidate powered-on BSs based on their flexibility, i.e. coverage and distance. Once the highest load $c_{p_{on}}$ is identified with potential coverage to the entire overloaded partition, the algorithm is completed. The p_{on} solution is the optimal one with highest flexibility. Otherwise, it continues until all $c_{p_{on}}$ are considered and selects as p_{on}, the candidate BS with the highest potential coverage breaking ties based on distance. The BS-awaking selection algorithm has an $O(QWlogW)$ complexity, where Q and W are the number of BS inside the hotspot and overload partition respectively, using binary heaps for maintaining the distance-coverage information.

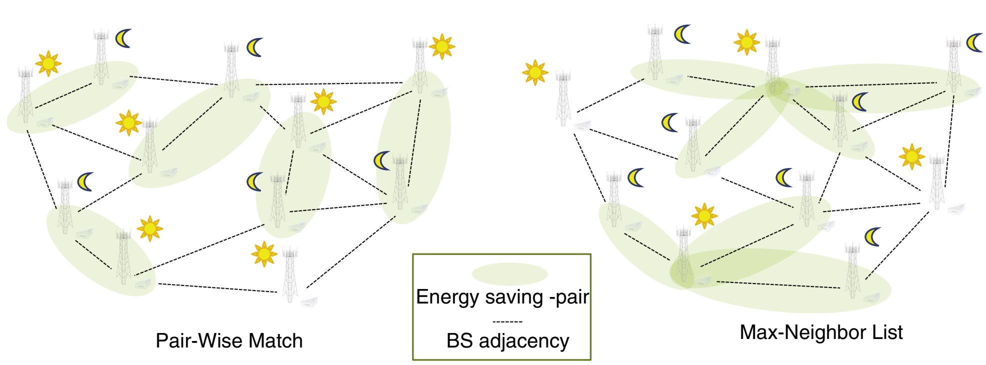


Fig. 5 Cell model and coverage extension

Table 1 Simulation parameters for evaluating energy partition algorithms

BSs (V)	C_{UE} C_{Bw}	$[L_R, H_R]$	$p(\mathrm{BS}_{\mathrm{low}})$	$Load_{\mathrm{lim}}$	p_{act}	[L,H]Bw
[10100]	100, 24 Mb	[0.5, 3]	0.7	20%	0.7	6.54 k, 3 Mb
[10100]	100, 24 Mb	[0.5, 3]	0.4	40%	0.3	6.54 k, 1 Mb

5 Simulation setup and result analysis

Following the description of the energy partition algorithms, this section compares their performance through a simulation based study. The simulation environment and the energy algorithms are implemented in Matlab. The objective is to compare the proposed algorithms using different off peak network configurations. Initially, the energy partition algorithms are compared considering the following scenarios: (i) high user traffic volume per BS and (ii) high bandwidth per session. In the first scenario, more users are introduced with low bandwidth demands, while the second utilizes fewer users with higher per session bandwidth demand. Then the re-configuration algorithms are compared with a more simple energy partition, which maintains a constant configuration regardless of load alternations.

The radio access network is represented as a graph $G(V, E)$ where V and E indicate BSs and adjacency among two neighbor BSs permitting users to handover. The Erdos–Renyi model $G(V, p)$ [19] is employed to generate random topologies considering different degrees of adjacency with $p = 0.1$. Such a model is used to study the effect of adjacency as a function of network size. Cells are circular with radius uniformly distributed between $[L_R, H_R]$ with wider coverage potential three times higher the maximum radius. Cell dimensions are assumed to be in kilometers but the same results hold for any kind of arrangement that maintains the same initial-potential cell radius ratio.

Every BS supports up to C_{UE} UEs and provides C_{Bw} bandwidth, representing a generalization of wireless resources as in [14], assuming no inter-cell interference. Load is established by launching individual UEs associated with each BS with UEs being active based on probability p_{act}. The bandwidth of each session is uniformly distributed in the interval [L, H], where L is 6.54 Kbps being the smallest rate of the Adaptive Multi-Rate (AMR) speech codec and H representing web to video or ftp applications.

5.1 Evaluation of energy partition algorithms

For evaluating energy partition algorithms, we initiate simulations from an off-peak instance, where the radio access network supports sporadic spots of average load, lower than normal conditions. The simulation does not consider a temporal development, but is trying to find an optimal solution based on the initial configuration. Under-loaded $\mathrm{BS}_{\mathrm{low}}$ are introduced with specified probability $p(\mathrm{BS}_{\mathrm{low}})$ having an upper bound load limit $Load_{\mathrm{lim}}$. Table 1 summarizes the simulation

Fig. 6 ES algorithms performance

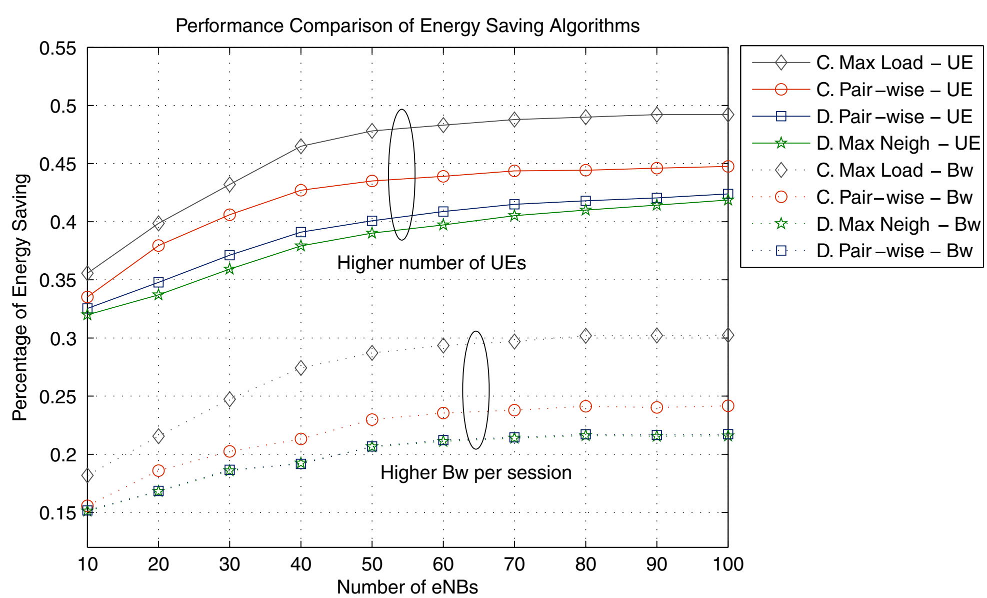
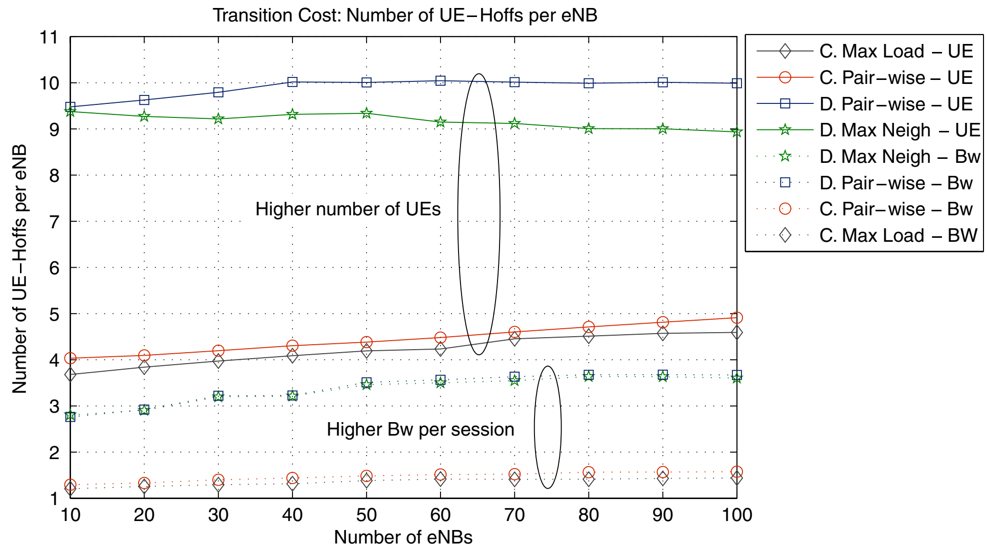


Fig. 7 ES Algorithms hoff overhead



parameters for the two scenarios. A centralized version of the pair-wise algorithm is also considered with the option of repeating the process returning the highest energy saving configuration. A prior simulation study indicated that 20 repetitions provide good results without increasing significantly the delay. We assess the initial energy partition algorithms using the following criteria:

- **Energy efficiency** represents the percentage of powered off BSs to the total number of BSs.
- **Load update cost** indicating the number of load update messages among neighbor BSs.
- **Handover cost** indicating the number of handovers that occur when BSs are powered-off.
- **Cost of collecting load information** referring to the load messages from particular BSs towards a centralized management entity.

Figure 6 presents the energy efficiency of the algorithms for high UE volume and high bandwidth per request, respectively. From the figure, centralized algorithms outperform distributed ones regardless of the network arrangement. Specifically, the centralized max load algorithm surpasses all the others, followed by the centralized pair-wise match and then by the distributed algorithms, which exhibit similar performance. The reason behind this is the scope of load information used by centralized algorithms from the complete network producing better energy partitions compared to the distributed ones that use information obtained from the neighbor BSs. The energy efficiency of all algorithms increases with the network size, primarily because the neighbor list of each BS grows. A generic observation is that higher overall resource consumption in combination with the high bandwidth per user limits the flexibility of transferring active sessions among coverage compensating BSs. Therefore, a load and mobility policy that restricts bandwidth demands on specific locations could prove useful for meeting the target conservation limits.

The handover cost related with configuring the energy state is illustrated in Fig. 7. The cost of centralized algorithms is similar, much lower than the equivalent cost of distributed. The main reason is based on the way different algorithms determine the energy partitions. Specifically, centralized algorithms determine the complete energy partitions before applying any related

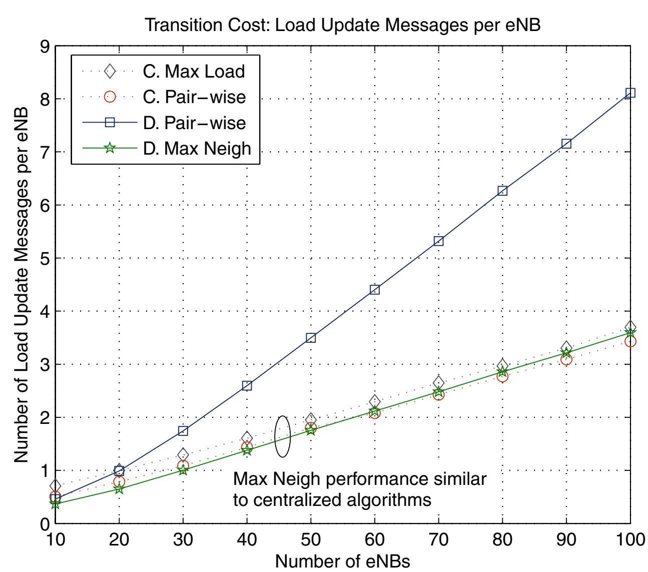


Fig. 8 High UE: load update overhead

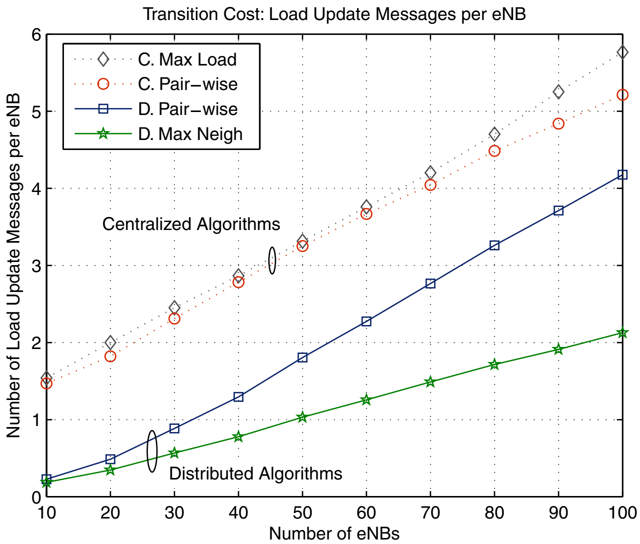


Fig. 9 High bw/request load update overhead

changes, whereas the distributed ones perform a step-wise operation. A higher UE volume with lower per session demands generates higher overhead, due to a larger number of transferred UEs.

The load update cost is shown in Figs. 8 and 9 for the high UE volume and bandwidth per user, respectively. Such update cost is essential for distributed algorithms but optional for centralized. Whereas the variants of the centralized algorithms exhibit similar performance, the load update cost of distributed algorithms depends significantly on the specific variant—with the pair-wise match generating higher cost compared to the max neighbor list. The reason is that each pair-wise step, is followed by a load update, while in the max neighbor list, an update is performed upon the new BS selection.

Considering the high UE volume scenario, the overhead of the distributed pair-wise algorithm is higher than the others. However, as the bandwidth per session demands increases, the overhead of centralized algorithms surpass the one of distributed. Higher bandwidth per session demand cause a nearly double increase in the overhead for centralized algorithms but a double overhead reduction for distributed ones. This overhead decline of the distributed algorithms is caused by the limited solution space, while the higher amount of compensating BS is the reason for the increased overhead for the centralized algorithms.

The overhead for collecting load information for centralized algorithms is of V magnitude, where V denotes the number of BSs included in such process. The exact amount of overhead also depends on the average packet size, i.e. packet header and on the number of load indicators. Finally, considering the worst-case complexity, centralized algorithms are V/N times more complex, where N is the average neighbor list size. In summary, for establishing energy partitions, centralized algorithms are more energy efficient, produce less handover overhead but cause more overhead for gaining load information and are more complex.

5.2 Evaluation of re-configuration algorithms

For evaluating re-configuration algorithms, we initiate simulations again from an off-peak instance, but this time the simulation process encounters traffic development, which is modeled as requests with specific source, destination and bandwidth requirements. Incoming users are assumed to arrive independently following Poisson with mean λ, while the session holding time is assumed to be an exponential with mean $\mu = 5$ min. The cell residence time is also assumed to be exponentially distributed with mean $n = 3$ min. Energy partitions are established based on the initial load conditions and then additional load is introduced at four random partitions by inserting new users with $\lambda = \{4, 5, 6, 7, 8\}$ to create congestion at selected regions.

A summary of the simulation parameters is illustrated in Table 2. The evaluation is performed measuring the contribution of each method to blocking and dropping rates with respect to the energy consumption. The blocking-dropping rate is defined as the percentage of blocked incoming and dropped handover sessions to the total amount of sessions while energy conservation is measured as defined priory. The blocking and dropping rates for each re-configuration algorithm including the simple centralized energy partition without re-arranging capabilities are shown in Fig. 10. The performance of the energy partition approach without re-configuration degrades rapidly. The distributed re-configuration then follows, while the centralized and BS-awaking selection algorithms introduce a considerably lower dropping and blocking rates especially as λ grows, with the BS-awaking scheme slightly outper-

Table 2 Simulation parameters for evaluating the re-configuration algorithms

BSs (V)	$C_{UE}\, C_{Bw}$	$[L_R, H_R]$	$p(\text{BS}_{\text{low}})$	$Load_{\text{lim}}$	p_{act}	[L,H]Bw
40	50, 12.6 Mb	[1, 5]	0.7	10%	0.3	6.54 k, 1 Mb

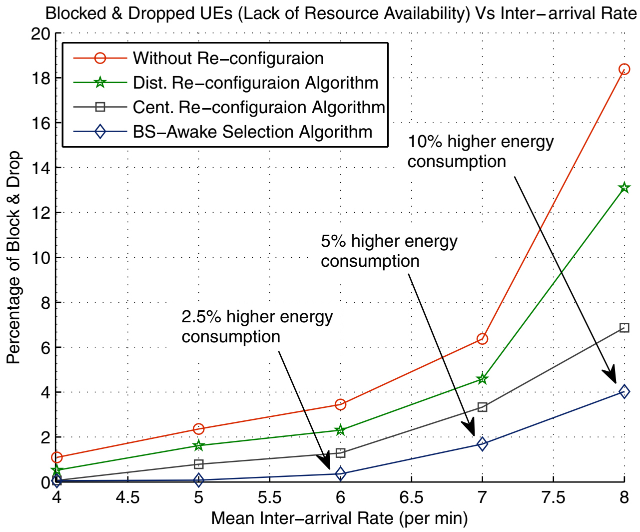


Fig. 10 Comparison of the re-configuration methods

forming the centralized. Distributed and centralized re-configuration algorithms cannot wake up additional BSs, maintaining constant energy consumption. This simplifies the comparative study, which focus on the performance of each individual algorithm.

The performance of all schemes is similar for $\lambda < 5$, beyond which their difference increases since further incoming load creates congestion on certain partitions. The centralized algorithm outperforms the distributed, because it uses information from the complete network producing more energy efficient re-configurations. In addition, the centralized scheme is faster in the sense that it is able to assess in a single step multiple congestion problems involving more than a single partition. Such rapid reaction is impossible for a distributed scheme because once a BS starts the re-configuration process neighbor BSs wait to avoid conflicts. This difference is investigated in this study by increasing λ on four regions introducing overloaded partitions at the same time, evident in Fig. 10 for $\lambda > 7$.

The algorithm that selects and wakes-up additional BSs inside overloaded partitions outperforms both centralized and distributed schemes at the cost of increased energy. Specifically, it selects additional BSs to re-power-on accounting for an energy increase of 2.5 to 10% respectively. For $\lambda < 6$ the energy consumption growth is 2.5%, while for $\lambda = \{7, 8\}$, it rises to 5 and 10% respectively. Centralized and distributed schemes are not increasing the energy consumption, since they only attempt to re-arrange the energy formation keeping the same number of operating BSs. Considering complexity, the centralized approach is the most demanding and introduces the highest overhead for retrieving load information. The distributed and BS-awaking one are less complex since a smaller number of BSs are involved producing also lower information exchange overhead.

6 Conclusion

This paper introduces the concept of energy partitions, associations of powered on and off BSs, formed collectively in a SON manner with the objective to match the operator offered capacity to the traffic demand. The unique benefit of energy partitions is flexibility, which permits energy re-configuration upon varying load conditions conserving maximum energy or upon a failure.

Centralized and distributed variants were developed for establishing energy partitions and for producing energy re-configuration. Our simulation results suggest that centralized algorithms perform better especially as bandwidth demand per user increases. In addition, in the re-configuration process centralized algorithms may react faster if congestion is spread among multiple partitions simultaneously. However, for regionally limited parts of the network, the advantage of having a global view is becoming less important, and a distributed algorithm could be competitive. Regarding the adoption of the BS-awaking algorithm, it could prove critical for cases where the re-configuration effort is higher than the energy consumption of one more awake BS, especially under circumstances of network instability.

Leveraging the insights described in this paper, one of the next steps for future work would be to analyze the limitations of the re-configuration process in terms of cost and stability. Moreover, we envision analyzing the performance of energy saving taking into account radio communication aspects, i.e. propagation and interference. Finally, we see some potential in further increasing energy saving by applying operator policies to influence user behavior, by rewarding conservation and penalizing users responsible for high consumption.

References

1. 3GPP TR 32.826 (2009) Telecommunication management; Study on energy savings management (ESM), (rel 9)
2. Fettweis G, Zimmermann E (2008) ICT energy consumption—trends and challenges. In: 11th WPMC
3. Cisco visual networking index: global mobile data traffic forecast update, 2009–2014. White paper (2010)

4. 3GPP TR 36.902 (2009) E-UTRAN; Self-configuration & self-optimizing network use cases & solutions, (Rel 9)
5. Marsan MA et al (2009) Optimal energy savings in cellular access networks. In: GreenComm workshop
6. Fehske AJ, Richter F, Fettweis GP (2009) Energy efficiency improvements through micro sites in cellular mobile radio networks. In: GreenComm2 workshop
7. Kim W-J et al (2009) Power efficient uplink resource allocation schemes in IEEE 802.16 OFDMA systems. IEICE Trans Commun E92-B(9):2891
8. Teyeb O, Van Phan V, Raaf B, Redana S (2009) Dynamic relaying in 3GPP LTE-advanced networks. In: EURASIP JWCN, vol 2009, article no 6
9. Izumikawa H et al (2010) Energy-efficient adaptive interface activation for delay/disruption tolerant networks. IEEE ICACT
10. Niu Z, Gong J, Yang Z (2010) Cell zooming for cost-efficient green cellular networks. IEEE Commun Mag 48(11):74–79
11. Zhou S et al (2009) Green mobile access network with dynamic base station energy saving. In: 15th ACM MobiCom
12. Schmelz LC et al (2009) Self-organisation in wireless networks use cases and their interrelation. In: 22nd WWRF
13. Dufkova K et al (2010) Energy savings for cellular network with evaluation of impact on data traffic performance. In: European wireless conference
14. Marsan MA, Meo, M (2009) Energy efficient management of two cellular access networks. In: ACM SIGCOMM GreenMetrics workshop
15. 3GPP TS. 25.111 (2009) Technical Specification Group RAN; LMU performance specification; UE positioning in UTRAN, (Rel 9)
16. Necker MC (2008) Interference coordination in cellular OFDMA networks. IEEE Netw 22(6):12–19
17. Karakayali MK et al (2006) Network coordination for spectrally efficient communications in cellular systems. IEEE Wirel Commun 13(4):56–61
18. Amirijoo M, et al (2008) Neighbor cell relation list and physical cell identity self-organization in LTE. Comms. workshop
19. Bollobas B (2001) Random graphs, 2nd ed. Cambridge University Press, Cambridge, UK