

RESEARCH

Open Access



Automatic clustering algorithms for indoor site selection in LTE

Rocío Acedo-Hernández^{1*}, Matías Toril¹, Salvador Luna-Ramírez¹, Carlos Úbeda² and María Josefa Vera³

Abstract

Small cell systems are a cost-effective solution to provide adequate coverage inside buildings. Nonetheless, the addition of any indoor site requires evaluating the trade-off between the coverage and capacity gain provided by the new site and its monetary cost. In this paper, a new automatic indoor site selection algorithm based on clustering techniques is presented. The algorithm calculates the number of antennas, radio heads, and baseband units needed for the area under study. Then, a clustering algorithm groups several radio heads of different buildings in a single pooled baseband unit, reducing deployment costs. The proposed clustering algorithm is based on a local refinement algorithm, whose starting point considers a new baseband unit for every new site, and then, possible reallocation to existing units is checked. To assess the method, the proposed indoor site selection algorithm is included in a network planning tool. The algorithm is tested in a real heterogeneous network scenario, taking into account vendor specifications and operator constraints. Results show that the use of the proposed clustering algorithm can reduce the total network cost by up to 58 % in a real scenario.

Keywords: Small cells, Clustering, LTE, In-Building Solutions

1 Introduction

Mobile data traffic is expected to increase considerably in the coming years. Specifically, a tenfold increase of mobile traffic from 2014 to 2019 is envisaged by several equipment vendors [1, 2]. In parallel, surveys predict that more than 70 % of this traffic will be generated indoors, but almost half of the houses and premises have poor indoor coverage currently [3, 4]. It is therefore necessary to develop a cost-effective solution that provides adequate coverage in buildings.

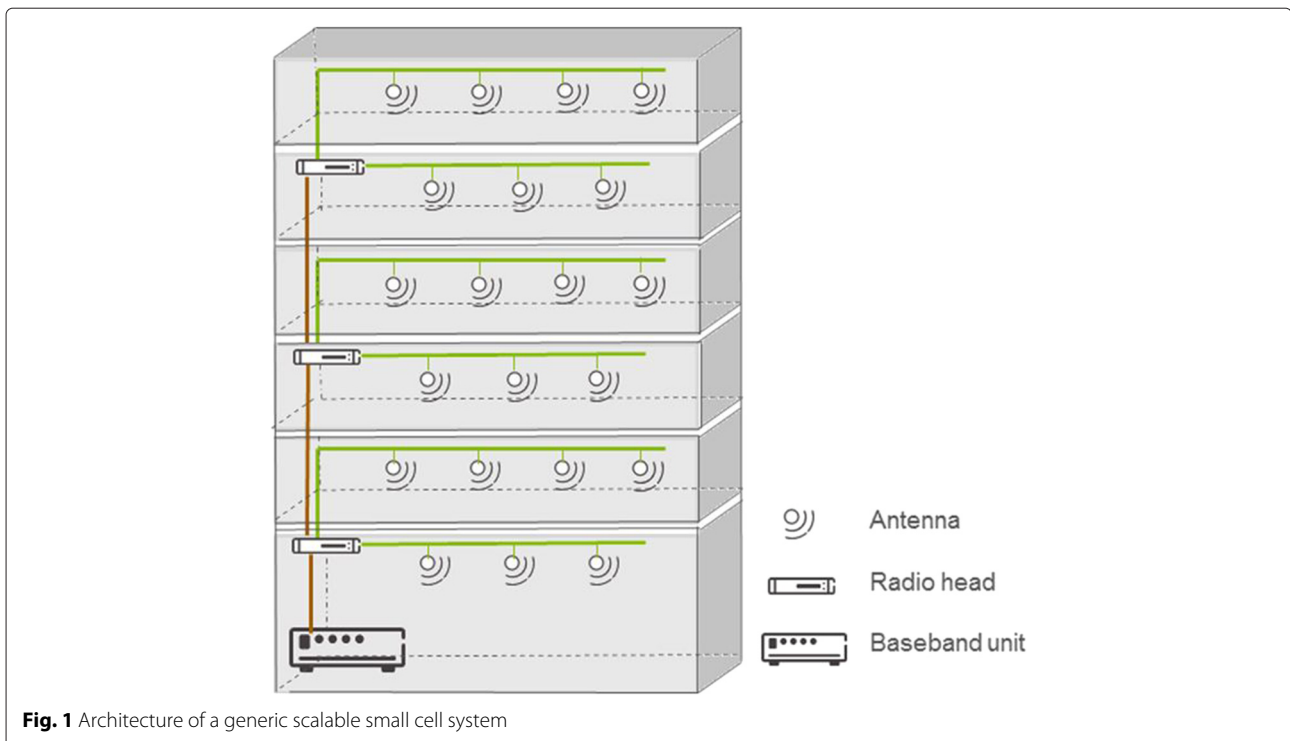
Providing indoor coverage and capacity has been a challenge since the start of mobile networking. Femtocell solutions [5–7] based on short-range low-cost low-power base stations are generally used to fulfill indoor capacity needs. Unfortunately, these systems only work correctly for small deployments, because cell planning and radio coordination become unmanageable when the number of cells increases [8]. For medium-to-large buildings, venues, and areas, macrocell features like coordination, seamless mobility, and interference management are needed. For

the last two decades, these needs have been overcome by distributed antenna systems (DAS) [9–11], where a cell's antenna is divided into several antennas covering the same area with a higher power efficiency, but still defining only one cell. This solution is still considered for multi-operator cases and neutral host applications. However, this approach becomes limited when new requirements for higher capacity and more advanced services appear, which is typically the case of medium-to-large buildings. In response to these new requirements, a promising solution is the deployment of small cells [12].

A small cell system is a hierarchical structure supplying coverage and capacity to some indoor area and behaving as one macrocell. The generic structure of a scalable small cell solution in a building is presented in Fig. 1. The considered system comprises a baseband unit controlling several radio heads, which are in charge of several antennas, usually in the same floor. The *baseband unit* is responsible for digital baseband signal processing. IP packets received from the core network are modulated into digital baseband signals and transmitted to the radio heads. The digital baseband signals received by the baseband unit from the radio heads are demodulated, and IP

*Correspondence: rah@ic.uma.es

¹Ingeniería de Comunicaciones, Universidad de Málaga, Málaga, Spain
Full list of author information is available at the end of the article



packets are transmitted to the core network. The baseband unit is also responsible for call and monitoring control processing. The *radio head* converts the digital baseband signal into radio frequency signals, which are amplified and transmitted to users by the antennas and vice versa.

Unlike femtocells, which have limited functionality, small cell solutions provide to the indoor segment the same functionality as the outdoor macrocell, i.e., the entire structure behaves as a macrocell enabling coordination with the rest of the mobile network, thus simplifying network operation and maintenance. At the same time, small cell solutions enhance system scalability and improve radio system performance compared to DAS systems.

Indoor site selection can be formulated as the discrete optimization problem known as *Capacitated Facility Location Problem (CFLP)* [13, 14]. Similarly to macrocellular network planning, the addition of a small cell system in an existing network is a trade-off between coverage/capacity improvement in the indoor area and the additional expenses due to the small cell infrastructure (also called *In-Building Solution, IBS*). Several works have studied the impact of site locations on cellular network performance in terms of coverage and/or signal to interference plus noise ratio (SINR). Thus, many different methods have been proposed to find the best location for macrocellular sites to improve network coverage, user connection quality, and/or network capacity [15–19]. Similar methods have been used for the optimal location

of wireless access points inside buildings [20–27]. All these approaches build a system model, over which a classical optimization algorithm is applied to find the location of access points maximizing some overall network performance indicator. In [28], an analysis of the properties of the optimal femtocell layouts for different radio performance criteria is presented. However, few studies have paid attention to deployment cost. Adding an IBS requires certain network resources (e.g., antennas, baseband units, and cabling), whose cost depends on decisions made during network planning. The current practice is to assign a baseband unit to every single building with an IBS. When medium/small buildings are considered, IBS resources are wasted because the majority of its ports are empty. For these cases, a clustering algorithm can reduce deployment costs by assigning the radio heads of different buildings into a single baseband unit. As a counterpart, there is an increase of cabling costs, which can be less than the initial solution with more baseband units, especially when linked buildings are close enough. To the authors' knowledge, no study has considered the use of clustering for sharing resources between buildings in the context of indoor site selection.

In this paper, a new clustering algorithm for the assignment of radio heads to baseband units in Long Term Evolution (LTE) is presented. The proposed algorithm is designed to be integrated in the site selection algorithm of a network planning tool. The algorithm is tested with a dataset constructed from a real heterogeneous network

scenario including macrocellular and microcellular sites (i.e., outdoor and indoor areas), vendor specifications, and operator constraints. The main contributions of this work are (a) a new clustering algorithm that minimizes indoor solution deployment costs, which can be integrated into a classical site location algorithm and (b) the results of the clustering algorithm in a real LTE heterogeneous environment. The rest of the paper is organized as follows. Section 2 formulates the indoor site selection problem from the operator perspective. Section 3 describes the proposed solution method. Then, Section 4 presents the performance analysis of the proposed method carried out with a system-level simulation tool. Finally, Section 5 presents the concluding remarks.

2 Problem formulation

Figure 2 shows an example of a small cell system, namely the Ericsson Radio Dot System (RDS) [29]. RDS has a centralized baseband architecture with three components: the Radio Dot (i.e., antennas); the indoor radio unit (IRU), acting as the radio head; and the digital unit (DU), working as the baseband unit. This work is focused on this structure, although results can easily be extrapolated to similar indoor solutions from other manufacturers.

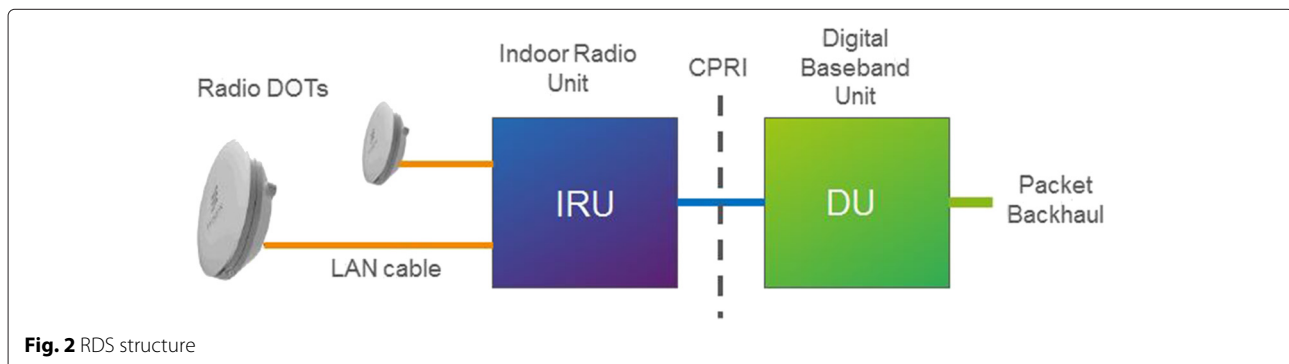
Figure 2 shows the RDS architecture and interfaces between elements. The DU provides pooled baseband processing for the system. It uses the Common Public Radio Interface (CPRI) standard to transfer synchronization, radio signals, and operation and management signals to the IRU. The IRU is a newly designed radio unit that incorporates existing macro software features (e.g., interference coordination, traffic management, and LTE combined cell), extending them with new indoor features (e.g., real-time traffic steering). The IRU connects to each Radio Dot using a proprietary IRU-Radio Dot interface over a conventional twisted pair local area network (LAN) cable. When co-located with the DU, an electrical CPRI interface is used, whereas a CPRI fiber interface is used for remote connection with the DU. Finally, the Radio Dot has two integrated antennas in a 10-cm form factor and weighs under 300 g. Each Radio Dot is connected to the IRU

through a dedicated LAN cable and remotely powered by Power over Ethernet (PoE). As in ethernet networks, the system employs a star topology, as opposed to the tree topology used in DAS. The ultra-compact design and use of LAN cabling simplify installation.

IBS expenses depend on the number of IBS (i.e., buildings) and the number and type of system elements per IBS to be deployed in the considered geographical area. The number of Radio Dots needed for each building is estimated from the Radio Dot coverage (provided by the manufacturer in the product technical specifications) and the surface of the building. A typical Radio Dot coverage is 500–800 m². The number of IRUs and DUs is determined by capacity restrictions. As shown in Fig. 3a, a star topology in RDS allows one DU to connect up to six IRUs and one IRU to connect up to eight Radio Dots. Thus, up to 48 Radio Dots per DU are supported. Alternatively, a cascade topology in RDS allows two IRUs to be cascaded to one DU port, supporting up to 96 Radio Dots per DU, as shown in Fig. 3b.

Different IBS structures can be used depending on the building size. Figure 4 shows some examples, depending on the area to be served. As shown in the figure, in medium-to-large buildings, the DU and IRU are co-located, while Radio Dots are set in a star topology. In the case of an existing nearby Radio Base Station (RBS), the DU can be co-located (e.g., a rooftop site on the same building). This structure enables advanced LTE coordination between outdoor and indoor coverage. In contrast, for large or very large buildings, IRUs are distributed over different floors or building segment levels. In the case of large venues, a central DU is shared by multiple facilities in the venue, i.e., not every facility has its own DU in this case.

The use of the latter structure can be generalized to small buildings that would not deserve their own DU in the proximity of large buildings that already have a DU. In these cases, a clustering algorithm is needed to evaluate when this shared structure is possible. When assigning the IRUs/Radio Dots of a building to a DU of another building, three factors must be taken into account: (a)



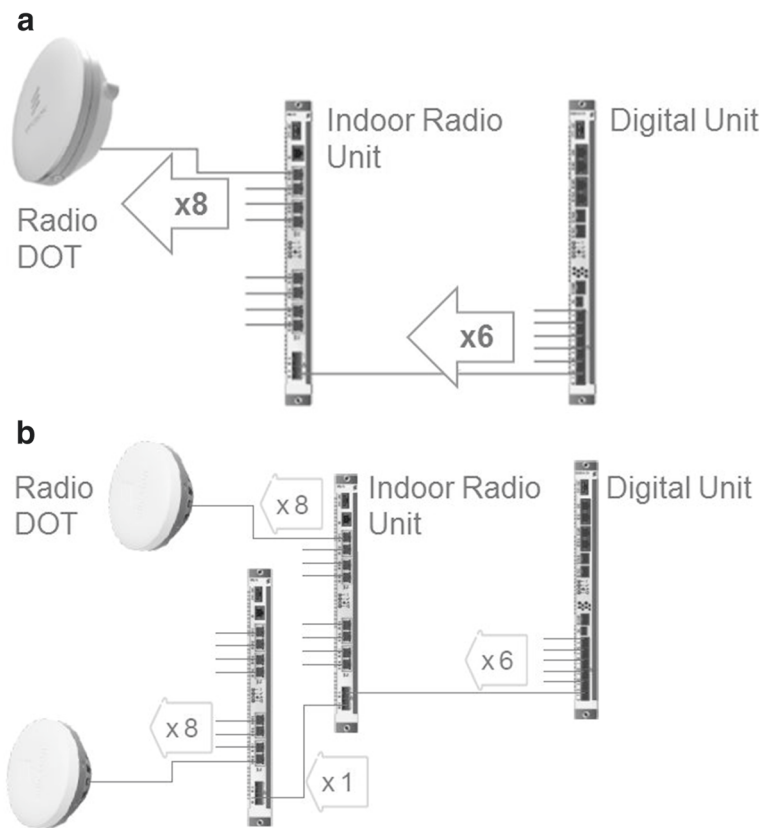


Fig. 3 RDS topologies: **a** star topology; **b** cascade topology

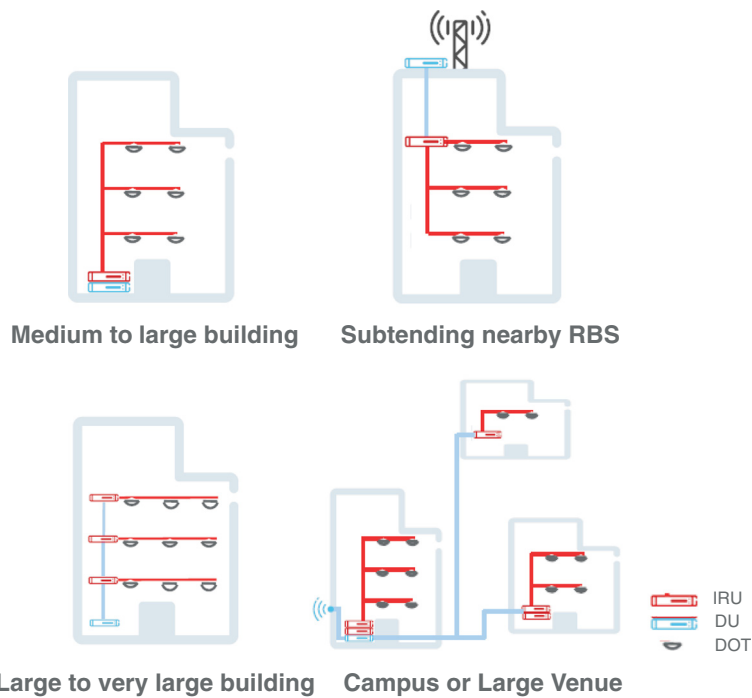


Fig. 4 Examples of IBS structures

the amount of required Radio Dots per building (coverage aspects), (b) the amount of Radio Dots per IRU and IRUs per DU (equipment capacity), and (c) the maximum distance between IRUs (cabling limit and expenses). The aim of the clustering algorithm is to group buildings into DUs as much as possible so that deployment costs are minimized, while satisfying coverage, capacity, and cabling constraints.

3 Solution method

Figure 5 shows the flow diagram of the indoor site selection algorithm, mainly consisting of four processes: dimensioning, clustering, site ranking, and recluster. Firstly, a dimensioning algorithm estimates the equipment needed to fulfill some coverage/capacity requirements in the indoor area of every candidate building in the scenario. Thus, an estimate of the number of Radio Dots and IRUs per building is obtained. Secondly, a clustering algorithm evaluates if candidate sites (i.e., buildings) can be

assigned to existing DUs. Then, new sites are selected iteratively. In each iteration, the most profitable candidate site is selected by ranking new sites in terms of radio network performance gain and deployment cost. When a new site requiring a DU is selected, a recluster algorithm checks if the assignment of building to DUs can be modified to reduce costs. The algorithm ends when all the buildings are included or when the deployment cost exceeds some budget constraint defined by the operator. The following paragraphs describe these processes in more detail.

3.1 Dimensioning

The dimensioning algorithm is applied independently to each building in the scenario. The aim of dimensioning is to estimate the number of antennas and IRUs needed per building. For the downlink of LTE, dimensioning is mainly driven by coverage issues [30, 31]. In this work, indoor coverage planning is based on simple geometrical

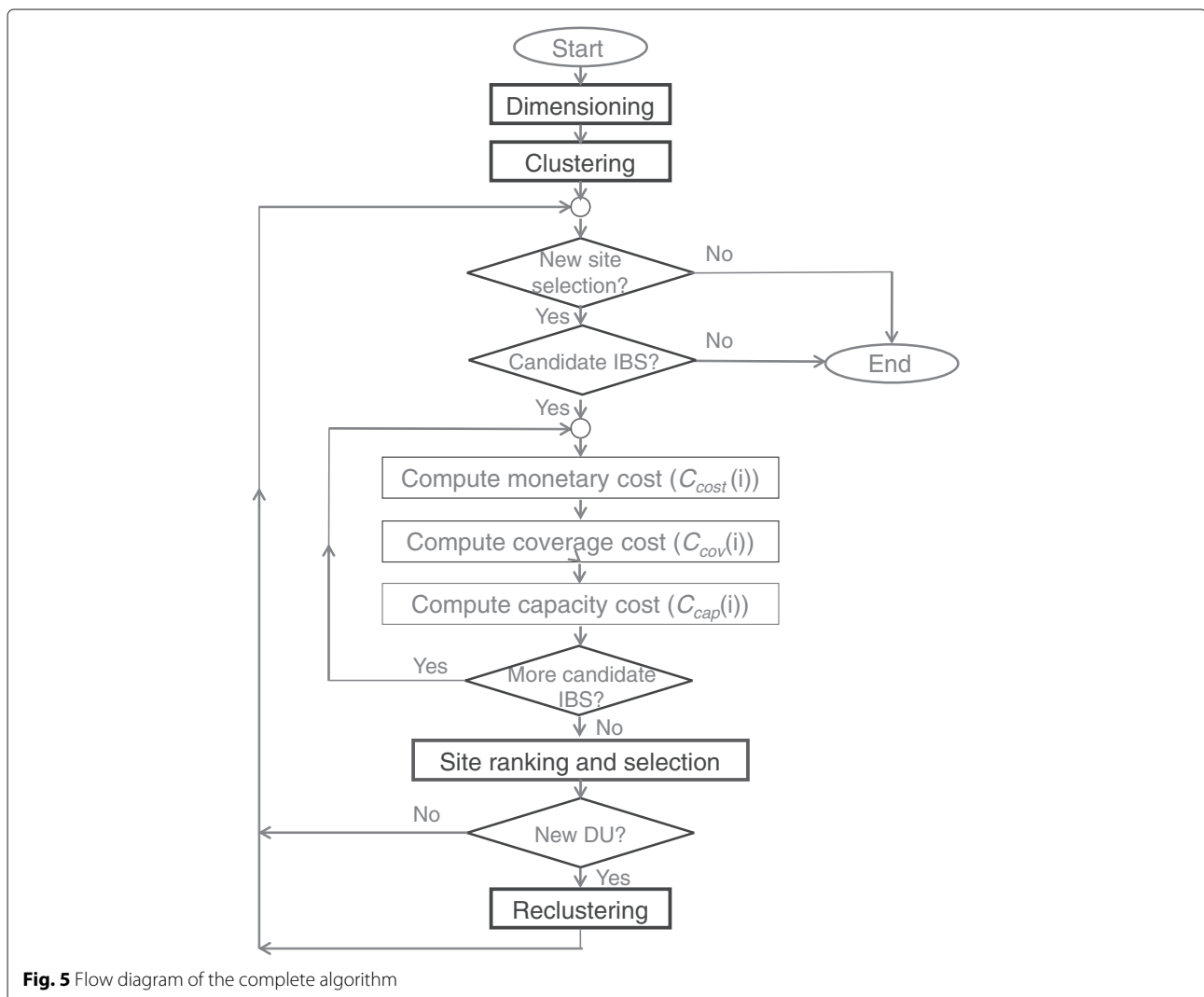


Fig. 5 Flow diagram of the complete algorithm

assumptions to reduce the computational load in the network planning tool.

To describe building geometry, geographical data can be extracted to define buildings as polygons, i.e., base coordinates and height. Figure 6 shows an example of a building of seven floors with a polygonal floor plan. A priori, the exact DU location within the building does not constrain the IBS design, since DUs are usually located in the facilities room (or next to the RBS, if it exists). It is assumed here that this room will be located in the lower floor and in the centroid of the building.

Once the floor surface and the number of floors are known for each building, the number of Radio Dots in building i is simply calculated by a rule of thumb as

$$N_{\text{Dot}}(i) = N_{\text{Dot/floor}}(i) * N_{\text{floor}}(i) , \quad (1)$$

$$N_{\text{Dot/floor}}(i) = S_{\text{floor}}(i) / \text{Dot}_{\text{cov}} , \quad (2)$$

where $N_{\text{Dot}}(i)$, $N_{\text{Dot/floor}}(i)$, and $N_{\text{floor}}(i)$ are the number of Radio Dots, the number of Radio Dots per floor, and the number of floors in building i , respectively; S_{floor} is the floor area (in square meters) of building i ; and Dot_{cov} is the Radio Dot coverage (in square meters), not depending on the shape of the building under study. Dot_{cov} is a parameter included in the technical specification of the product, and the floor area can be computed with the Gauss area formula [32] from the base coordinates of the building.

Once the number of Radio Dots has been estimated, the number of IRUs is calculated. Either of the two topologies shown in Fig. 3a, b might be used, leading to 6 or 12 IRUs per DU and eight Radio Dots per IRU. In practice, the star configuration is the typical configuration and is, therefore, considered hereafter.

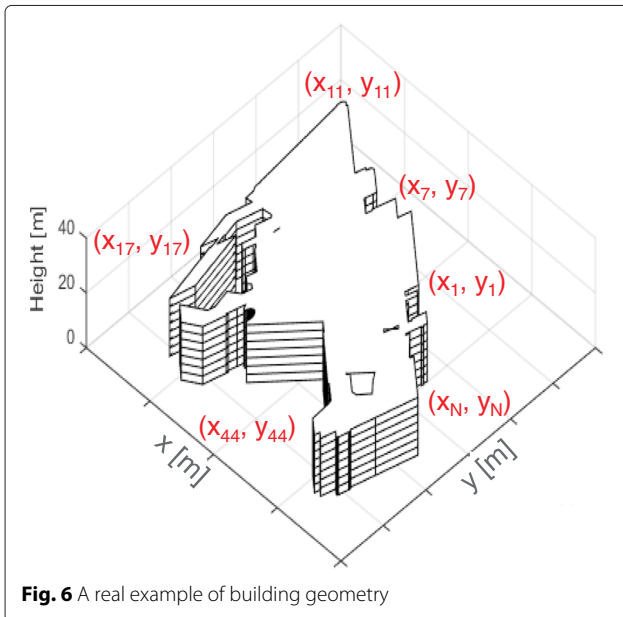


Fig. 6 A real example of building geometry

3.2 Clustering

Once the number of IRUs per building has been estimated in the dimensioning step, the clustering algorithm groups IRUs into DUs. The output of clustering is the assignment of every IRU to some DU.

The clustering of buildings into DUs can be formulated as an Integer Linear Programming (ILP) model as follows:

$$\text{Min} \sum_{i=1}^{N_b} \left(K_{\text{DU}} X_{ii} + \sum_{\forall j \neq i} K_{\text{fiber}} d_{ij} X_{ij} X_{jj} \right) \quad (3)$$

$$\text{s.t.} \sum_{j=1}^{N_b} X_{ij} = 1 \quad \forall i \quad (4)$$

$$\sum_{i=1}^{N_b} \omega_i X_{ij} \leq B_{aw} X_{jj} \quad \forall j \quad (5)$$

where N_b is the number of buildings in the planned area, K_{DU} and K_{fiber} are the DU cost and fiber cost per unit and meter, respectively, and d_{ij} is the distance in meters between buildings i and j . X_{ij} represents the binary decision variables defining the assignment of an IRU in building i to the DU in building j as

$$X_{ij} = \begin{cases} 1 & \text{if building } i \text{ is connected to building } j, \\ 0 & \text{otherwise,} \end{cases} \quad (6)$$

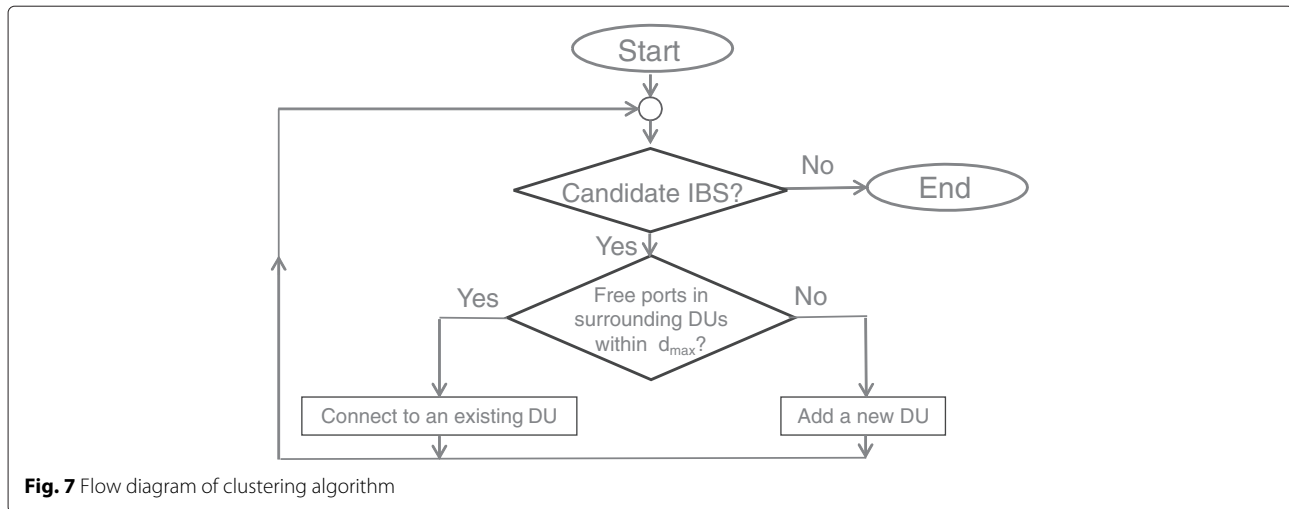
and

$$X_{ii} = \begin{cases} 1 & \text{if there exists a DU in building } i, \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

Note that some of these decision variables are fixed, since there might already exist some sites implemented in the network. Some of them may have their own DU (i.e., $X_{ii} = 1$), while others do not (i.e., $X_{ii} = 0$ and $X_{ij} = 1$ for some j). The variable ω_i denotes the number of IRUs needed in building i (estimated by the dimensioning algorithm) and B_{aw} is the maximum available connection for IRUs in a DU.

The objective function in (3) consists of a first term reflecting equipment cost and a second term reflecting cabling costs. The first constraint in (4) forces that each building must be assigned to only one DU (i.e., single homing). The second constraint (5) reflects hardware limitations so that no more than B_{aw} IRUs can be connected to the same DU.

For computational reasons, the combinatorial optimization problem in (3–5) is solved by a heuristic approach. As shown in Fig. 7, the clustering algorithm starts by assigning a new DU for every candidate site. Then, the algorithm checks if there is some DU with enough free ports located in adjacent buildings j within some distance, d_{max} , from the building under study. This d_{max} distance is the minimum of two distances, namely the maximum distance of the CPRI fiber interface due to physical limitations,



$d_{\max\text{phy}}$, and the distance above which the cost of cabling is larger than the cost of installing a new DU. Such a distance depends on the relation between equipment and cabling costs as

$$d_{\max} = \min(d_{\max\text{phy}}, K_{\text{DU}}/K_{\text{fiber}}) . \quad (8)$$

If $d_{ij} \leq d_{\max}$, the candidate site does not require a DU (as it is cheaper to assign its IRUs to the DU of a surrounding building with spare capacity) and the associated cost is due to fiber deployment (i.e., $C_{\text{cost}} = K_{\text{fiber}/m} * d_{ij}$). In contrast, if $d_{ij} > d_{\max}$, the candidate site needs a new DU (since cabling cost are more expensive than buying a new DU or there is no building with a DU with enough spare capacity in the surroundings) and the cost is that of a new DU ($C_{\text{cost}} = K_{\text{DU}}$).

3.3 Indoor site ranking and selection

The last step is the selection of the next candidate site to be implemented, for which potential sites must be ranked in terms of revenue. If only monetary issues are taken into account, a greedy site selection algorithm would favor candidate sites close to existing ones, since those are cheaper to deploy. However, coverage and capacity gains must also be taken into account. For this purpose, a general figure of merit (FoM) is defined to rank candidate sites as

$$\text{FoM}(i) = \omega_{\text{cov}} * C_{\text{cov}}(i) + \omega_{\text{cap}} * C_{\text{cap}}(i) - \omega_{\text{cost}} * C_{\text{cost}}(i) \quad (9)$$

where C_{cov} and C_{cap} denote the coverage and capacity gains obtained, respectively, if that candidate indoor site is added and C_{cost} is the deployment cost, including all radio and backhaul equipment. Parameters ω_{cov} , ω_{cap} , and

ω_{cost} are weights to prioritize objectives defined according to operator policies. Once all possible candidates are evaluated, the site with the largest FoM is selected.

3.4 Reclustering

The site selection algorithm works iteratively by selecting the next best candidate site. Normally, the operator runs several iterations of the site selection algorithm to define a set of new indoor sites that will be deployed in the future. The size of that set (and hence the number of iterations) depends on the budget.

As candidate sites are selected, new DUs may appear. After adding a site with a new DU, a reclustering algorithm checks if (a) that new DU can be avoided by relocating nearby DUs and (b) previously planned IRUs in the surroundings should be reallocated to the new DU to reduce cabling costs. If both options are feasible, the solution with the lowest cost is chosen.

Figure 8, show the two cases. The first case, shown in Fig. 8a, reflects the relocation of a previously planned DU. The left side of the figure represents the situation before reclustering where one building has its own DU (DU1) and one candidate building needing a new DU (DU2) has just been selected. A closer analysis shows that the addition of DU2 can be avoided if DU1 is relocated to a building closer to the new building selected. Thus, all the three buildings can be served by a single DU. The second case, shown in Fig. 8b, reflects the reassignment of a building to a new DU. The left side of the figure represents the situation before reclustering where one building has its own DU (DU1) and one candidate building is connected to that DU. When a new DU appears (DU2) as a result of selecting a new candidate site, surrounding sites are checked for new shorter (and cheaper) connections allowed by the new DU. The right side of the figure represents the



Fig. 8 Reclustering cases: **a** relocation of a DU; **b** reassignment of a building to a new DU

situation where a building has been reassigned to the new DU after reclustering.

4 Performance analysis

In this section, different tests are performed to assess the algorithms described previously. A real scenario obtained from a live heterogeneous network has been used. For clarity, the analysis setup is first introduced and results are then presented.

4.1 Analysis setup

The indoor site selection algorithm is included in a static system-level LTE simulator deployed in Matlab. The simulated scenario covers a geographical area of 30 km², including 176 macrocells. In this scenario, a smaller area of 12 km² with 449 buildings and 25 macrocells is used in the analysis. Figure 9 shows the area with the candidate buildings for inclusion of an IBS. Figure 9a shows the 2D map of the area, with buildings and nearby macrocells, and Fig. 9b shows the height of the buildings. The scenario is divided into a grid of points, representing potential user locations. Propagation losses to each macrocellular base station are calculated by a variant of the Okumura-Hata model [33] adjusted with the Ericsson 9999 model for 2600 MHz and urban scenario. Real building geographical data is used to identify indoor locations, for which

10 dB of penetration losses due to walls are added to the path loss. It is checked a posteriori that more than 59 % of potential user locations fall inside a building. A log-normal slow fading is also considered. Cell service areas are pre-calculated, and to avoid border effects, only points with a minimum Reference Signal Received Power (RSRP) are considered. Inside buildings, it is assumed that coverage and capacity requirements are satisfied with a specific number of Radio Dot units, computed by the dimensioning algorithm. Thus, no indoor propagation model is needed. Spatial traffic distribution is obtained from the average PRB utilization ratio and Timing Advance (TA) measurements collected per cell during the busy hour in the live network. Table 1 summarizes the main parameters in the simulator. With the simulation tool, the coverage and capacity gain of adding an IBS is estimated for each building. The coverage gain, C_{cov} , of a building is the share of pixels inside the building that was not covered properly before (i.e., coverage hole) and could be served by the new indoor site. For simplicity, it is assumed here that all grid points that fall inside a building experience an adequate service with an indoor solution and the indoor solution does not have any impact on the interference received by nearby macrocells due to wall isolation. Similarly, the capacity gain, C_{cap} , for each building is the share of traffic of nearby macrocells that would be captured by

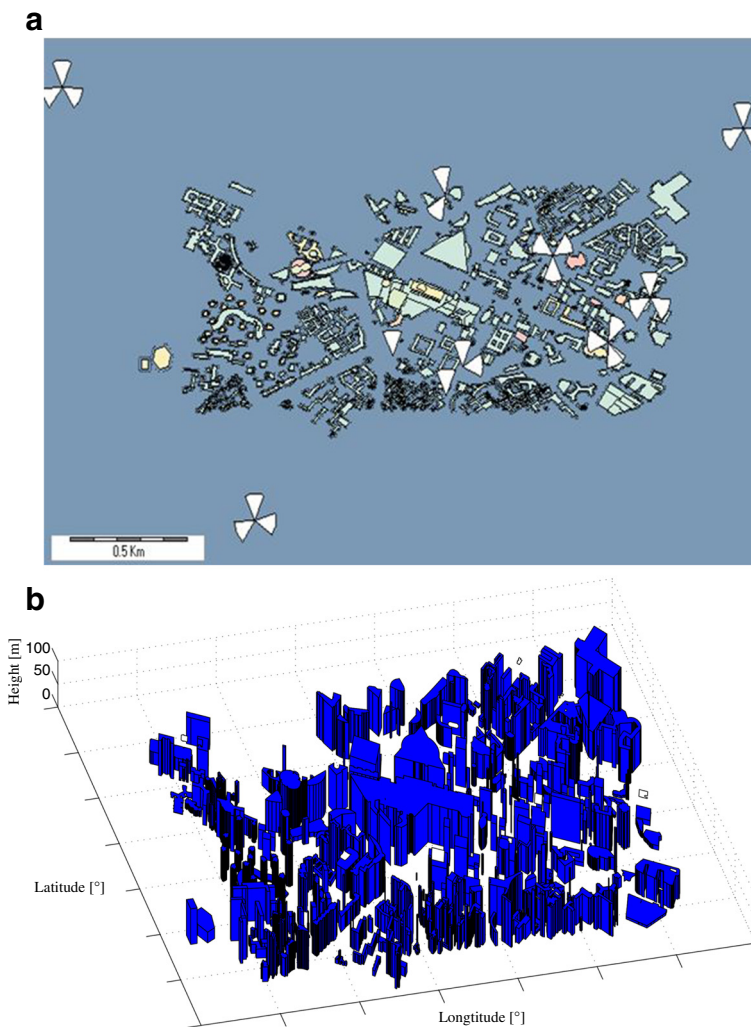


Fig. 9 Scenario used: **a** 2D scenario; **b** building heights

the new indoor site. Such an indicator aims to quantify the traffic relief in the macrocell achieved by the indoor solution. The inclusion of every new IBS affects coverage and capacity gains of other indoor sites, which must be considered for the next iteration in the site selection algorithm.

The indoor site selection algorithm is tested with four different clustering approaches of increasing complexity in the same scenario: (a) the *baseline* solution considering one DU per building (i.e., no clustering); (b) the proposed heuristic *clustering* algorithm with no reclustering (i.e., once a building is assigned to a DU, it cannot be reallocated); (c) the previous approach including the *reclustering* feature, where the assignment can be modified after the selection of every new site; and (d) the *exact* (i.e., optimal) solution, obtained by solving the ILP model in (3–5) with the Gurobi solver [34], provided that the set of sites to be added is known.

First, all methods are evaluated assuming that only monetary cost is used for site selection (i.e., $\omega_{cov} = 0$, $\omega_{cap} = 0$, and $\omega_{cost} = 1$, and thus, $FoM \equiv C_{cost}$). This can be done outside the simulator, since no coverage and capacity computations are needed. Later, all methods are included in the simulator to compute coverage and capacity estimations for selecting the best candidates. For ease of analysis, all weights are set the same (i.e., $\omega_{cov} = \omega_{cap} = \omega_{cost}$). Thus, it is evaluated how site selection based on other parameters impacts on monetary cost.

As the focus of this work is the clustering method, the site selection algorithm is run in all cases until all the candidate buildings in the scenario (449) have been selected. Thus, the main performance indicator to assess the clustering methods is the total monetary cost, C_{cost} . For ease of comparison, the cost of each solution is normalized by that of the baseline approach (i.e., without clustering).

Table 1 Simulation tool parameters

Simulation parameters	
Simulator type	System-level, static (grid-based)
Grid resolution (m)	20
Spatial traffic distribution	Irregular, based on PRB utilization ratio and TA measurements Overall network PRB utilization = 31 %
Antenna model	Antenna configuration MIMO (2 × 2)
Frequency (MHz)	2600
System bandwidth (MHz)	10
Number of PRBs	50
Propagation model (dB)	Outdoor: $PL = A - 13.82 * \log_{10}(h_{BS}[m]) + (B - 6.55 * \log_{10}(h_{BS}[m])) * \log_{10}(d[\text{km}])$ $A = 157.5, B = 44.9, h_{BS}[m] = 30$ Shadowing log-normal fading, 8 dB std. Correlation distance 20 m Outdoor to indoor: 10-dB penetration losses
Macrocell model	$P_{tx} = 46 \text{ dBm}, G = 13 \text{ dB}$
UE model	Antenna height 1.5 m Noise floor -114.45 dBm (per PRB)

Note that coverage and capacity criteria are only included to quantify the impact of these factors on the clustering algorithm. A thorough analysis of coverage and capacity performance in the simulated heterogeneous network is beyond the scope of the work.

4.2 Analysis results

In a first experiment, neither coverage nor capacity gains are considered in FoM. Figure 10 shows the normalized cost obtained by the different algorithms for different d_{max} values. Recall that, from (8), $d_{max} = K_{DU}/K_{fiber}$ if there is no physical limit for the IRU-DU interface. As shown in the figure, the proposed clustering and reclustering methods reduce the monetary cost considerably compared to the baseline solution (one DU per IBS), achieving near-optimal performance. It can also be seen that the total cost reduction strongly depends on the d_{max} value. Results show that the use of the proposed reclustering algorithms can reduce the total network cost by up to 58 % in the case of $d_{max} = 2500$.

When only C_{cost} is considered in FoM, selected sites are added one close to the other, since it is cheaper. However, when coverage and capacity factors are also considered, sites are selected from distant locations in the network, which has a strong impact on clustering algorithms. To prove this influence, in a second experiment, the clustering methods are tested when new candidates are randomly selected from different locations in the network. To simplify the analysis, d_{max} is fixed to 600. Figure 11 represents a box plot [35] with the quartiles of the normalized cost of the solution of each method for 100 different random orderings of the 449 buildings in the scenario. From the figure, it can be deduced that the order in which sites are selected affects clustering performance. However, cost differences are small when compared to the average cost reduction obtained by the clustering and reclustering algorithms. Only a few cases

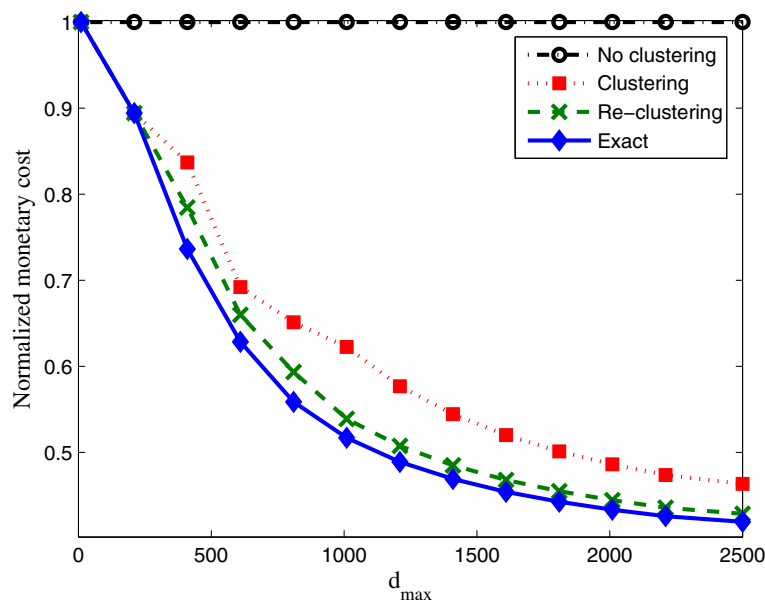


Fig. 10 Method comparison (only monetary criteria)

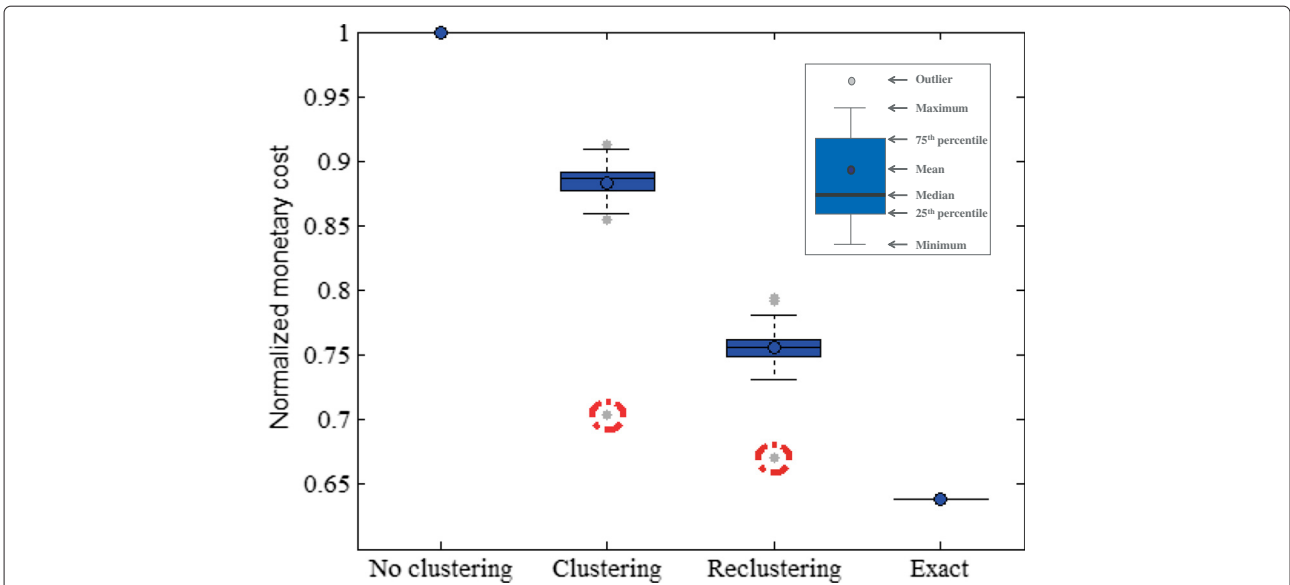


Fig. 11 Cost sensitivity in different methods to non-monetary criteria

(outliers, highlighted with a circle in the figure) show large deviations, with much lower normalized cost (which is good). As expected, these cases correspond to the case of selecting sites with the minimum monetary cost criterion.

In the last experiment, the complete site selection algorithm is tested, including coverage and capacity factors in (9). The aim is to quantify the impact of coverage and capacity criteria on the clustering performance in a real case. For ease of analysis, all weights in (9) are set the

same. Figure 12 plots the normalized monetary cost with every method for different values of d_{max} . As expected, it is observed that the baseline and exact methods perform the same as in the first experiment (shown in Fig. 10). Note that, for the baseline solution, one DU is set per building (i.e., monetary cost does not change), and in the exact method, coverage and capacity are not taken into account in (3). In contrast, the clustering and reclustering methods perform worse than in the first experiment, showing

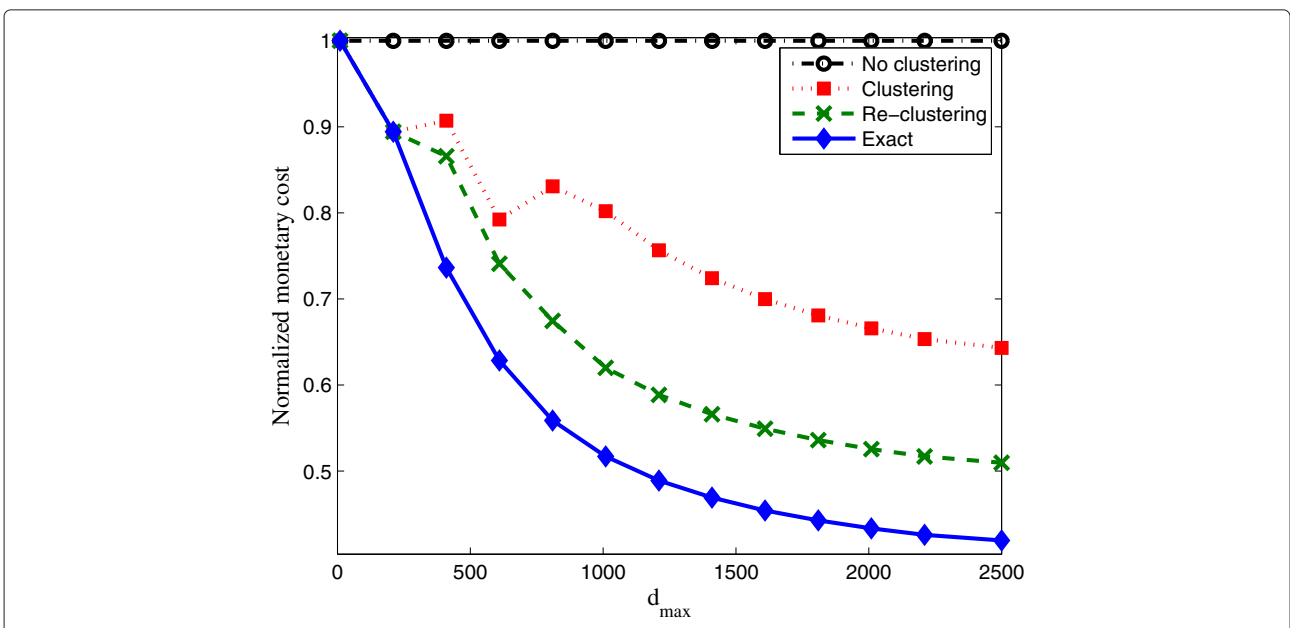


Fig. 12 Method comparison (complete site selection)

higher monetary costs. The difference is due to the different order of selecting new candidates when coverage and capacity issues are taken into account. Thus, network coverage and capacity improve at the expense of increasing deployment costs. More importantly, it is observed that the inclusion of coverage and capacity criteria has a larger impact on the basic clustering method than on the method with reclustering. Results show that, with equal weights for coverage, capacity, and monetary costs, the inclusion of the proposed reclustering algorithm can reduce the total network deployment cost by up to 49 % in a real scenario. With the basic clustering method, the cost can be reduced only by up to 35 %. Hence, it can be concluded that the need for reclustering is more critical in the real network case.

All the methods have been executed in a personal computer with a Intel(R) Core(TM) processor, 2.6-GHz clock frequency and 8 GB of RAM. When only monetary costs are considered, the proposed heuristic (re)clustering algorithm only takes less than 3 min to find a near-optimal solution with small deployment costs for the 449 buildings. In contrast, the exact solution built with Gurobi takes approximately 2 h.

5 Conclusions

In this paper, a novel automatic clustering algorithm for deploying IBS in a cost-effective manner has been proposed. In an initial dimensioning stage, the algorithm estimates the required number of elements in every building. Then, in a second stage, the clustering algorithms looks for the deployment solution with minimum cost by assigning different buildings into the same DU. The clustering solution is updated by a reclustering algorithm after adding every new site. The proposed clustering approach has been integrated into a site selection tool and tested with a dataset of a real heterogeneous scenario in a static system-level LTE simulator. Simulation results show that, when coverage, capacity, and monetary costs are considered, the inclusion of the proposed clustering algorithm can reduce the total network cost by up to 49 % in a real scenario.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work has been funded by the Spanish Ministry of Economy and Competitiveness (TIN2012-36455) and Optimi-Ericsson and Agencia IDEA (Consejería de Ciencia, Innovación y Empresa, Junta de Andalucía, ref. 59288), co-funded by FEDER.

Author details

¹Ingeniería de Comunicaciones, Universidad de Málaga, Málaga, Spain.

²Ericsson, Madrid, Spain. ³Ericsson, Málaga, Spain.

Received: 1 August 2015 Accepted: 13 March 2016

Published online: 22 March 2016

References

1. Cisco Systems Inc., Cisco visual networking index: forecast and methodology, 2014–2019. Technical report, Available: <http://www.cisco.com> (Accessed on Jul 2015) (2015)
2. Ericsson, Ericsson mobility report. Technical report, Available: <http://www.ericsson.com/res/docs/2015/ericsson-mobility-report-june-2015.pdf> (Accessed on Jul 2015) (2015)
3. S Saunders, in *2nd International Conference on Home Access Point and Femtocells*. The role of cooperation in establishing an efficient femto economy, (2007), pp. 1–5
4. V Chandrasekhar, J Andrews, A Gatherer, Femtocell networks: a survey. *IEEE Commun. Mag.* **46**(9), 59–67 (2008)
5. 3GPP TR 25.467, 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; UTRAN architecture for 3G Home NodeB; Stage 2 (Release 8). (v.8.1.0, 2009)
6. R Raheem, A Lasebae, M Aiash, J Loo, in *2013 Second International Conference on Future Generation Communication Technology (FGCT)*. From fixed to mobile femtocells in LTE systems: issues and challenges, (2013), pp. 207–212. doi:10.1109/FGCT.2013.6767218
7. R Raheem, A Lasebae, J Loo, in *2014 28th International Conference on Advanced Information Networking and Applications Workshops (WAINA)*. Performance evaluation of LTE network via using fixed/mobile femtocells, (2014), pp. 255–260. doi:10.1109/WAINA.2014.51
8. G Jeney, in *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*. Practical limits of femtocells in a realistic environment, (2011), pp. 1–5. doi:10.1109/VETECS.2011.5956403
9. AAM Saleh, AJ Rustako, RS Roman, Distributed antennas for indoor radio communications. *IEEE Trans. Commun.* **35** (1987)
10. KJ Kerpez, A radio access system with distributed antennas. *IEEE Trans. Veh. Technol.* **45**(2), 265–275 (1996). doi:10.1109/25.492850
11. H Osman, H Zhu, J Wang, in *2010 IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*. Downlink distributed antenna systems in indoor high building femtocell environments, (2010), pp. 1016–1020. doi:10.1109/PIMRC.2010.5672090
12. T Nakamura, S Nagata, A Benjebbour, Y Kishiyama, T Hai, S Xiaodong, Y Ning, L Nan, Trends in small cell enhancements in LTE advanced. *Commun. Mag. IEEE.* **51**(2), 98–105 (2013). doi:10.1109/MCOM.2013.6461192
13. S Melkote, MS Daskin, Capacitated facility location/network design problem. *European J. Oper. Res.* **129**, 448–495 (2001)
14. K Holmberg, M Rannqvist, D Yuan, An exact algorithm for the capacitated facility location problems with single sourcing. *Eur. J. Oper. Res.* **113**(3), 544–559 (1999). doi:10.1016/S0377-2217(98)00008-3
15. CY Lee, HG Kang, Cell planning with capacity expansion in mobile communications: a tabu search approach. *IEEE Trans. Veh. Technol.* **49**, 1678–1691 (2000)
16. AJ Nebro, F Chicano, F Luna, in *6th International Conference Numerical Methods and Applications*. Optimal antenna placement using a new multi-objective CHC algorithm, (2007)
17. L Raisanen, R Whitaker, Comparison and evaluation of multiple objective genetic algorithms for the antenna placement problem. *Mobile Netw. Appl.* **10**(1–2), 79–88 (2005)
18. E Amaldi, A Capone, F Malucelli, Planning UMTS base station location: optimization models with power control and algorithms. *IEEE Trans. Wireless Commun.* **2**(5), 939–952 (2003)
19. P Avella, S Mattia, A Sassano, Metric inequalities and the network loading problem. *Discret. Optim.* **4**(1), 103–114 (2007)
20. L Nagy, L Farkas, in *11th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*. Indoor base station location optimization using genetic algorithms, vol. 2, (2000), pp. 843–846
21. Z Ji, TK Sarkar, B-H Li, Methods for optimizing the location of base stations for indoor wireless communications. *IEEE Trans. Antennas Propag.* **50**(10), 1481–1483 (2002)
22. JKL Wong, AJ Mason, MJ Neve, KW Sowerby, Base station placement in indoor wireless systems using binary integer programming. *IEE Proc. Commun.* **153**(5), 771–778 (2006)
23. Y Ngadiman, YH Chew, BS Yeo, in *IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications*. A new approach for

- finding optimal base stations configuration for CDMA systems jointly with uplink and downlink constraints, vol. 4, (2005), pp. 2751–2755
24. T Fruhwirth, P Brisset, Placing base stations in wireless indoor communication networks. *Intell. Syst. Appl. IEEE.* **15**(1), 49–53 (2000)
 25. SJ Fortune, DM Gay, BW Kernighan, O Landron, RA Valenzuela, MH Wright, Wise design of indoor wireless systems: practical computation and optimization. *Comput. Sci. Eng. IEEE.* **2**(1), 58–68 (1995)
 26. L Pujji, K Sowerby, M Neve, Development of a hybrid algorithm for efficient optimisation of base station placement for indoor wireless communication systems. *Wirel. Pers. Commun.* **69**(1), 471–486 (2013)
 27. MA Abd Rahman, M Dashti, J Zhang, in *International Conference on Localization and GNSS*. Localization of unknown indoor wireless transmitter, (2013), pp. 1–6
 28. JM Ruiz, M Toril, S Luna-Ramírez, A femtocell location strategy for improving adaptive traffic sharing in heterogeneous LTE networks. *EURASIP J. Wireless Commun. Netw.* **2015**, 38 (2015). doi:10.0086/s13638-015-0246-0
 29. C Lu, M Berg, E Trojer, P Eriksson, K Laraqui, OV Tridblad, H Almeida, Connecting the dots: small cells shape up for high-performance indoor radio. *Ericsson Rev.* (2014). Available: http://www.ericsson.com/res/thecompany/docs/publications/ericsson_review/2014/er-radiodot.pdf (Accessed on March 2016)
 30. S Sesia, I Toufik, M Baker, *LTE: the UMTS Long Term Evolution, from theory to practice*. (Wiley, USA, 2009)
 31. H Holma, A Toskala, *LTE for UMTS: evolution to LTE-Advanced*. (John Wiley & Sons, UK, 2011)
 32. B Braden, The surveyor's area formula. *Coll. Math. J.* **17**(4), 326–337 (1986)
 33. J Milanovic, S Rimac-Drlje, K Bejuk, in *14th IEEE International Conference on Electronics, Circuits and Systems, 2007. ICECS 2007*. Comparison of propagation models accuracy for WiMAX on 3.5 GHz, (2007), pp. 111–114
 34. Gurobi Optimization. Available: <http://www.gurobi.com/> (Accessed on March 2016)
 35. W Navidi, *Statistics for engineers and scientists*. (McGraw-Hill Higher Education, USA, 2008)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
