IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 14, 2018, accepted June 4, 2018, date of publication June 22, 2018, date of current version August 7, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2847609

Unsupervised Learning Algorithm for Intelligent Coverage Planning and Performance Optimization of Multitier Heterogeneous Network

JURAJ GAZDA[1], EUGEN ŠLAPAK[1], GABRIEL BUGÁR[1], DENIS HORVÁTH[2], TARAS MAKSYMYUK[3], (Member, IEEE), AND MINHO JO[4], (Senior Member, IEEE)

[1]Department of Computers and Informatics, Technical University of Košice, 040 01 Kosice, Slovakia
[2]Center of Interdisciplinary Biosciences, Technology and Innovation Park, P. J. Šafárik University, 040 01 Košice, Slovakia
[3]Department of Telecommunications, Lviv Polytechnic National University, 79000 Lviv, Ukraine
[4]Department of Computer Convergence Software, Korea University, Seoul 02841, South Korea

Corresponding author: Minho Jo (minhojo@korea.ac.kr)

This work was supported in part by the Slovak Research and Development Agency under Project APVV-15-0055 and in part by the European Intergovernmental Framework COST Action CA15140: Improving Applicability of Nature-Inspired Optimisation by Joining Theory and Practice. The work of T. Maksymyuk was supported by the Ukrainian Government Project (Designing the methods of adaptive radio resource management in LTE-U mobile networks for 4G/5G development in Ukraine) under Grant 0117U007177. The work of M. Jo was supported by the National Research Foundation of Korea Grant through the Korean Government under Grant 2016R1D1A1B03932149.

ABSTRACT The densification of mobile network infrastructure has been widely used to increase the overall capacity and improve user experience. Additional tiers of small cells provide a tremendous increase in the spectrum reuse factor, which allows the allocation of more bandwidth per user equipment (UE). However, the effective utilization of this tremendous capacity is a challenging task due to numerous problems, including co-channel interference, nonuniform traffic demand within the coverage area, and energy efficiency. Existing solutions for these problems, such as stochastic geometry, cause excessive sensitivity to the pattern of the UE traffic demand. In this paper, we propose an intelligent solution for both coverage planning and performance optimization using unsupervised self-organizing map (SOM) learning. We use a combination of two different mobility patterns based on Bézier curves and Lévy flights for more natural UE mobility patterns compared with a conventional random point process. The proposed approach provides the advantage of adjusting the positions of the small cells based on an SOM, which maximizes the key performance indicators, such as average throughput, fairness, and coverage probability, in an unsupervised manner. Simulation results confirm that the proposed unsupervised SOM algorithm outperforms the conventional binomial point process for all simulated scenarios by up to 30% in average throughput and fairness and has an up to 6-dB greater signal-to-interference-plus-noise ratio perceived by the UEs.

INDEX TERMS Coverage planning, heterogeneous network, self-organizing map, unsupervised learning.

I. INTRODUCTION

Mobile networks are currently facing a rapid increase in traffic demand, forcing operators to enhance the existing radio access network (RAN) infrastructure. Conventional solutions such as cell splitting and sectoring have already reached their limits in terms of the effective trade-off between network performance and total cost of ownership [1]. Therefore, a heterogeneous network (HetNet) architecture has been proposed by the 3rd Generation Partnership Project (3GPP) as an acceptable alternative for improvement of the overall system throughput by introducing multitier coverage. These additional tiers are created by small cells in areas with particularly high traffic demand [2]. The important benefit of the small cells is their convenient deployment in any indoor or outdoor area (e.g., lampposts, metro stations, restaurants, and shopping malls). Moreover, indoor small cells are frequently purchased by the location owners, who also cover the electricity cost to power these base stations (BSs). This, in turn, results in a considerable reduction of capital expenditures (CAPEX) and operating

expenses (OPEX) for the mobile operators, while providing a significant improvement in overall network capacity and coverage quality [3].

The overall capacity of multitier HetNets is greater than that of conventional single-tier RANs, owing to the many additional cells, which overlay the coverage of the macrocell. Because these cells are considerably smaller than macrocells, the operator can reuse spectrum among them more effectively, providing more bandwidth per user and, theoretically, unlimited area spectral efficiency. Moreover, the small cell itself has higher spectral efficiency owing to the improved quality of shorter wireless channels. This advantage is especially noticeable in the macrocell edge areas, where macrocell user equipment (UE) suffers from a low signal-to-interference-plus-noise ratio (SINR) owing to the high distance from the serving BS. Offloading UEs from macrocell edges to small cells can substantially improve user experience and overall network performance [4]. Small cells can be installed in any arbitrary location that provides an advantage in flexibility of coverage deployment. However, the locations of the small cells cannot be completely random because their uncoordinated deployment could result in severe interference among cells, especially in dense scenarios. It is important to consider the area traffic demand during coverage planning to ensure the highest gain from any additional small cells. Therefore, optimized planning of HetNet coverage must be applied by the operator to avoid semi-blind placement of small cells and ensure the best network performance [5].

Typically, the deployment of small cells follows local design guidelines, which means that there are no strict requirements for electromagnetic compatibility or exposure limits to electromagnetic radiation because of low transmission power. Thus, installation of a small cell is similar to the installation of an additional Wi-Fi access point and should not be considered as a substantial CAPEX, such as that of a macrocell. However, technical and regulatory requirements remain, which must be considered before roll-out of a HetNet [6]. First, the network capacity must be planned considering the existence of the shared and neutral host infrastructure and the availability of a reliable backhaul connection, covered by novel flexible deployment standards and regulations. Nevertheless, solving the optimization problem of determining the best network configuration in terms of throughput and coverage can be computationally intensive and not feasible for dynamic systems such as HetNets [7]. The performance of a HetNet is extremely sensitive to parameters such as the UE location, mobility, and traffic demands. For example, a small cell does not provide any significant advantage in terms of capacity if current traffic demands are less than the capacity of the macrocell. Therefore, it is important to consider the geographic distribution of the traffic demands around the target coverage area to determine where to install the small cells. The complexity of optimal HetNet deployment is in the random movement of the UEs, which makes it difficult to predict the areas with the highest average traffic demand and to identify the most profitable allocation for the small cells. In this paper, we address this complex problem and propose a self-learning algorithm for optimization of HetNet deployment using self-organizing maps (SOMs). The proposed algorithm employs two mobility patterns that reflect the features of pedestrian and vehicular UE. Based on the simulated behavior of the UEs, the proposed algorithm attempts to deploy small cells in areas with higher loads, while considering the presence of co-channel interference among the small cells. In general, an SOM represents an unsupervised learning technique based on neural networks, which was introduced by Kohonen [8]. Further extensions of SOMs are presented in [9] and [10]. SOM networks have demonstrated excellent performance in various applications, such as document collection [11], image processing [12], and security-threat detection [13]. Recently, SOMs have been applied as the basic algorithm for the event-sensor deployment problem with random distribution of events [14]. In [15], Salem *et al.* apply the SOM algorithm to effectively minimize the probability of sensing errors in dynamic spectrum access networks. A detailed and comprehensive survey of different SOM applications in wireless communications is provided by Zhang *et al.* [16].

In this paper, we follow the conclusions regarding the efficient deployment of emerging HetNets based on artificial intelligence as presented in [17]. We utilize the unsupervised SOM algorithm for solving both HetNet coverage planning and performance optimization. Unlike previous studies, where random point processes (e.g., Poisson and binomial) are used to simulate the mobility of the UEs, we propose a combination of two different mobility patterns based on Bézier curves and Lévy flights. The former represents a periodic semi-dynamic pattern, which reflects the movement in public transport, and is interchangeable with the periods of fixed user positions, such as at home or at the office. The latter model employs a Lévy flight pattern with truncated power-law flight lengths [18]. The Lévy flight is a series of short movements with random longer movements. Such patterns have been observed in human behavior in Song *et al.* [19]. Thus, the proposed mobility model is noisier and more natural than a conventional random point process. The proposed approach targets explicitly topologically correct mappings of the spatial probability density distributions of high demand areas.

The major contributions of this paper can be summarized as:

- The new intelligent algorithm for HetNet coverage planning and performance optimization is proposed based on unsupervised SOM algorithm.
- The new realistic UE mobility pattern has been developed based on the combination of Bézier curves and Lévy flights random movement.
- Comprehensive simulations have been conducted for different number of small cells within the target coverage area.

- Practical recommendations for HetNet coverage planning have been derived by finding the best trade-off between average throughput per UE, coverage probability and fairness of resource allocation.

The remainder of this paper is organized as follows. Section II provides an overview of the recent related work on multitier HetNet coverage design. Section III presents a detailed explanation of the investigated system model and mobility patterns. Section IV describes the proposed SOM-based algorithm for optimization of the deployment of small cells. Simulation results and performance analyses are presented in Section V. Section VI concludes the paper.

II. RELATED WORK

The deployment of a two-tier HetNet consisting of macrocells and small cells has attracted significant attention recently for many reasons, including improved user throughput, reduced latency, and overall energy savings. In this paper, we focus mainly on the optimization of small cell deployment within the existing coverage of macrocells. Therefore, this section aims to advance our research in the context of existing works in this area. Although there are many research papers regarding the advantageous characteristics of small cell deployment, specific research targeting small cell optimization for outdoor scenarios remains limited.

We must acknowledge that the simulation of two-tier HetNets, where the investigation of different positions and transmission power of a small cell within the cell coverage of the macrocell, was conducted in [20]. The results rely mainly on the brute-force-based investigation of the various parameters, with a general recommendation regarding the small cell settings (large macro-to-small cell distances, relatively high transmit power, and high deployment density). Shin and Zain [21] investigate the coverage of the geo-clustering algorithm in HetNets, where small cells are added to complement the coverage holes of the macrocell coverage. This approach performs static clustering of network coverage without consideration of UE mobility. A sampling-based optimization method for small cell deployments with the objective of maximizing UE throughput is introduced in [22]. More general objectives such as maximization of the traffic offloading to small cells or minimization of the cost of service delivery are targeted in [23]. However, within the scope of the above-mentioned research, no specific UE dynamics have been considered that limit their application in general.

Applications that exploit the self-organizing principle in combination with the dynamic behavior of UEs are investigated in [24] and [25]. In [24], Park *et al.* propose an algorithm for outdoor small cell deployment that considers mutual interference and load variation among different coverage zones. The algorithm is based on the weighted sum of objectives, which belongs to the class of multi-objective optimization methods. Another study is provided in [25], where the authors analyze the impact of the imperfect small cell positioning concerning dynamic hotspots (HSs). Grid search based on metric calculations for several candidate locations of small cells is proposed and evaluated in [26]. Nomadic spatially quasi-static UE distribution is considered throughout the simulation scope. Zhou *et al.* [27] investigate the optimum location of small cells given a specific UE traffic pattern. They assume predefined probable locations for the small cells, and thus they are not obliged to search through the entire parametric search space, which reduces the complexity of the algorithm.

However, the existing (primarily static) model solutions remain significantly distant from realistic dynamic situations involving continuous changes in the physical locations of the UEs. Therefore, the dynamics we study herein consider not only the specific deterministic motion modulated by periodic cycles, but also the noisy, unpredictable influence of UEs following Lévy flight motion. We assume that the time-varying and spatially dependent requirements of the UEs can be achieved, for example, using appropriate heuristics. Considering the results obtained, the expansion of the portfolio of the existing heuristic approaches to the problem of optimizing small cells in HetNets is also expected.

III. PRELIMINARIES AND SYSTEM MODEL

In this section, we outline the system model considered in this paper, the investigated metrics characterizing the impact of HetNet deployment optimization on the UE experience, the traffic mobility patterns considered throughout the simulation scope and finally, the model simplifications.

A. SYSTEM MODEL

We consider the downlink transmission of a two-tier HetNet. The tiers operate at different frequencies, i.e no cross-tier interference is observed. Nevertheless, interference among BSs of the same type is present (i.e., co-channel interference). Whereas the traditional macrocell layer broadcasts traffic on a dedicated 2100 MHz carrier, small cells utilize the 3350–3650 MHz band. The reason for this derives from the recent FCC release document *Notice of Proposed Rulemaking*, where the standard development bodies issued 100 MHz of available spectrum to be used by small cells in the 3350–3650 MHz band [28]. This bandwidth is characterized in general by high penetration losses, and thus is not feasible for long-distance transmission. However, small cells can exploit the available spectrum for providing services and to offload traffic from the macrocells.

In our simulation setup, the BSs of both layers (i.e., macrocells and small cells) use the round-robin resource scheduling algorithm with full allocation, which shares the available resources among all UEs. Each UE is associated with a BS based on the calculated SINR, which provides improved results compared with conventional measurements of reference signal receiving power (RSRP).

We consider a model where the two-tier HetNet consists of N_{M} macrocells with dedicated positions and N_{S} small cells randomly distributed in the macrocell coverage area. Each macrocell has access to N_{RB} resource blocks (RB)

($N_{\mathrm{RB}} = 100$); for the small cells, $N_{\mathrm{RB}} = 20$ as proposed by 3GPP in [29]. Denoting the number of all BSs (macrocells and small cells) to be deployed as N_{BS}, the total number of BSs is given by $N_{\mathrm{BS}} = N_{\mathrm{M}} + N_{\mathrm{S}}$. Depending on where the UE is associated, we classify the UE as a macro-UE (the UE is associated with a macrocell) or small-UE (the UE is associated with a small cell).

Let us denote the positions of the small cells and macrocells by $\mathbf{y}_i \in \mathbb{R}^2$, where $i = 1, 2, \ldots, N_{\mathrm{M}} + N_{\mathrm{S}}$. The UEs, following a specific mobility model, are dispersed in the investigated region and their locations are denoted by $\mathbf{x}_j \in \mathbb{R}^2$, where $j = 1, 2, \ldots, N_{\mathrm{u}}$. Here N_{u} represents the number of UEs in the region. Assuming that the channel on each RB experiences independent and identical Rayleigh fading, the channel power gain of the link between the j-th UE and associated $i-$th BS in a RB is expressed as

$$g_{ij}(t) = A_{ij}\, g_{\mathrm{f},ij}(t)\, g_{\mathrm{pl},ij}, \tag{1}$$

where A_{ij} denotes antenna gain; $g_{\mathrm{f},ij}(t)$ is the exponentially distributed fading gain with unit mean representing the small-scale channel fading gain at time t; and $g_{\mathrm{pl},ij}$ is the path loss. Here, we consider the ECC-33 channel propagation model, as it demonstrates the closest agreement with the measurement results in the frequency range 2000–3500 MHz and is common for urban environments. The path-loss for the ECC-33 model can be defined as in [30]:

$$g_{\mathrm{pl},ij} = A_{\mathrm{fs}} + A_{\mathrm{bm}} - G_{\mathrm{b}} - G_{\mathrm{r}}[\mathrm{dB}], \tag{2}$$

where A_{fs}, A_{bm}, G_{b}, and G_{r} are the free space attenuation, the basic median path loss, the BS height gain factor, and UE height gain factor, respectively. These elements can be further characterized as

$$\begin{aligned} A_{\mathrm{fs}} &= 92.4 + 20\log_{10}(d) + 20\log_{10}(f), \\ A_{\mathrm{bm}} &= 20.41 + 9.83\log_{10}(d) + 7.894\log_{10}(f) + \\ &\quad + 9.56[\log_{10}(f)]^2, \\ G_{\mathrm{b}} &= \log_{10}(h_{\mathrm{b}}/200)\{13.958 + 5.8[\log_{10}(d)]^2\}, \\ G_{\mathrm{r}} &= [42.57 + 13.7\log_{10}(f)] \\ &\quad \times [\log_{10}(h_{\mathrm{r}}) - 0.585\,], \end{aligned} \tag{3}$$

where f denotes the frequency in GHz; h_{b} is the antenna height; h_{r} is the height of the UE; and d is the Euclidean distance between the associated cell-UE pair. The instantaneous received SINR of the dedicated RB of the j-th macro-UE located at $\mathbf{x}_j$, associated with the macrocell at location $\mathbf{y}_i$ at time t can be expressed as

$$\gamma_{ij}(t) = \frac{P_i g_{ij}(t)}{\sum\limits_{k=1,k\neq i}^{N_{\mathrm{M}}} P_k g_{kj}(t) + \sigma^2} = \frac{P_i g_{ij}(t)}{I_{ij} + \sigma^2}, \tag{4}$$

where P_i is the transmit power of the i-th macrocell; σ^2 is the additive white Gaussian noise variance; and I_{ij} is the interference power received by macro-UE j from macrocells other than the i-th macrocell.

For a small-UE, we have the similar description:

$$\gamma_{ij}(t) = \frac{P_i g_{ij}(t)}{\sum\limits_{k=1,k\neq i}^{N_{\mathrm{S}}} P_k g_{kj}(t) + \sigma^2} = \frac{P_i g_{ij}(t)}{I_{ij} + \sigma^2}, \tag{5}$$

where P_i denotes the transmition power of the i-th small cell; and I_{ij} is the interference power received by the small-UE j from neighboring small cells.

The throughput of the jth macro-UE when associated with the i-th macrocell at time t (the aggregated value accumulated through all assigned RBs) can be defined as

$$\Gamma_{ij}(t) = W \sum_{k=1}^{N_{\mathrm{RB}}} \log_2\left(1 + \frac{P_i g_{ij}^k(t) b_{ij}^k}{I_{ij}^k + \sigma^2}\right), \tag{6}$$

where $b_{ij}^k = 1$ if the k-th RB of the associated i-th macrocell is served to the j-th UE (otherwise $b_{ij}^k = 0$); W denotes the bandwidth per RB; $g_{ij}^k(t)$ and I_{ij}^k denote the channel power gain of the link between the j-th UE and the associated i-th BS in the k-th RB and interference received in the k-th RB, respectively. The variable N_{ij}^{RB} determines the number of RBs assigned to the j-th macro-UE in the i-th macrocell as follows:

$$N_{ij}^{\mathrm{RB}} = \sum_{k=1}^{N_{\mathrm{RB}}} b_{ij}^k = \left\lfloor \frac{N_{\mathrm{RB}}}{N_{\mathrm{c}}} \right\rfloor, \tag{7}$$

where $\lfloor \ldots \rfloor$ denotes the floor function, and N_{c} determines the number of macro-UE associated with the i-th macrocell ($N_{\mathrm{c}} \ll N_{\mathrm{u}}$). This concept ensures the proportional fairness distribution of the resources among the UEs of macro cells. associated with the i-th macrocell (round-robin scheduling). A similar assumption is derived for the UEs of small cells.

Throughout this paper, the network model is assumed to operate in a saturated mode where all BSs (both macrocells and small cells) transmit at their maximum power setting. This represents a worst-case scenario, and is a typical methodology used for network capacity dimensioning, as suggested in [31]. By applying the full-load condition in the network, the various deployment solutions are pushed to their ultimate limits in a systematic manner. Finally, we assume that UEs are capable of forwarding the traffic only over one cell at a time (either macrocell or small cell, depending on the conditions), which conforms with the LTE standard without the advanced carrier aggregation feature.

B. METRICS

With the objective of analyzing the impact of HetNet deployment optimization, we now introduce various metrics that allow us to numerically quantify our solutions. In addition to the most intuitive metric, the throughput Γ_{ij}, we have also chosen to consider fairness and coverage probabilities to achieve more general conclusions.

In wireless communication networks, all UEs expect to have uniform quality of experience (QoE) across the whole network coverage. If UEs in the network do not experience a similar QoE (expressed using, for example, the throughput

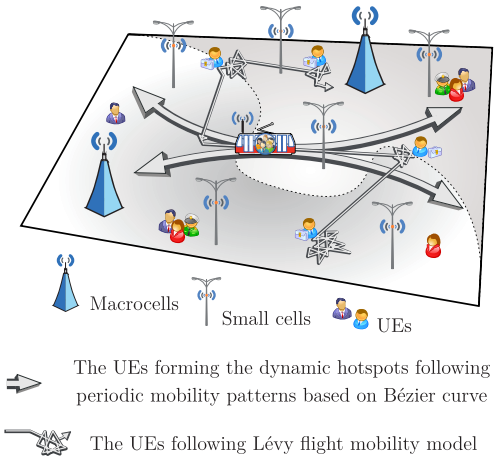

FIGURE 1. Schematic diagram of a HetNet model with a set of macrocell and small cell stations, and mobility patterns of HSs and UEs.

metric), the system is less fair and thus, we must introduce a function that quantifies the fairness. Here, we use Jain's index defined as

$$F(\{\Gamma\}) = \frac{\left(\sum_{j=1}^{N_u} \Gamma_j\right)^2}{N_u \sum_{j=1}^{N_u} \Gamma_j^2}. \tag{8}$$

For readability, we drop the index i from the definition of Γ_{ij} as we are interested in particular UE contributions to the fairness regardless of the associated BS. Jain's index has the following characteristics [32]:

- The index provides values in the interval (0, 1], where "1" means a completely fair allocation (all UEs have the same SINR value) and value close to "0" means totally unfair allocation (single UE has highest SINR value, while other UEs are out of coverage), respectively.
- The fairness is independent of the number of UEs active in the system.
- The fairness is a continuous function. Any change in the perceived $\{\Gamma\}$ is reflected in the fairness.

The coverage probability p_c determines the mean fraction of UEs that, at any given time, achieve an SINR greater than the given threshold β, i.e.,

$$p_c = \frac{1}{N_u} \sum_{j=1}^{N_u} \mathbf{1}_{\gamma_j \geq \beta}, \tag{9}$$

where $\mathbf{1}_{(...)}$ represents the indicator function.

C. TRAFFIC MOBILITY MODELS

In recent years, a number of techniques have been developed to model different transport mobility situations. The main idea behind the application of mobility models is that the outputs (i.e., trajectories) depict synthetic data that represents the realities of transport. One of the key points of our present work is the extension and application of mobility modeling principles [33]. This is mainly the formal decomposition of motion into its regular (trending or periodic) and random components. Furthermore, we focus particularly on the assemblies, including periodic and stochastic trajectories. It is clear that long-term stationarity is essential for the quality of SOM learning required to provide the (static) small cell layouts. In our applications, we consider the regular transport of UE clusters (semi-dynamic HSs) having a typical transport period of half an hour. Regular trajectories are parameterized by the function composition consisting of the Bézier curves, and logistic and harmonic functions. Independently generated Lévy flights are used to mimic the unpredictable and intermittent two-dimensional (2D) moves. An additional assumption is that the simulated entities do not change their manner of movement from random to regular, or vice versa.

1) REGULAR MOBILITY, PERIODIC TRAJECTORIES AS MODULATED BÉZIER CURVES

We describe the periodic regular movements of the semi-dynamic HSs. HSs can be used to calculate smooth pathways between obstacles using a Bézier curve characterized by control points. The application of Bézier trajectories is advantageous in describing the movement pattern with collision-free and obstacle constraints [34]. The basic idea is to create an assembly of smooth trajectories consisting of a system of regular visits to predetermined starting, target, and intermediate points. The intermediate points (only for the Bézier curves of order higher than two) are used to control trajectory curvature. Assume that M semi-dynamic HSs enumerated as $j = 1, 2, \ldots, M$ are dispersed in 2D. The simplified assumption is considered here, so that each HS consists of a constant number of UEs. We introduce Δt as the time step used to generate the mobility data. For a given simulation time interval $[\,0, T_{\text{sim}}\Delta t\,]$ that is greater than the period of motion T (i.e., $T_{\text{sim}}\Delta t \gg T$), a system of numerical trajectories can be generated that is parametrized by the modulated quadratic Bézier curves $\{\mathbf{x}_j(t) \in \mathbb{R}^2;\ t \in [0, T_{\text{sim}}\Delta t];\ j = 1, 2, \ldots, M\}$ with the particular trajectory

$$\begin{aligned}\mathbf{x}_j(t) = {} & (1 - \Phi_j(t))^2\, \mathbf{x}_{\text{init},j} \\ & + 2(1 - \Phi_j(t))\Phi_j(t)\, \mathbf{x}_{\text{interm},j} \\ & + \Phi_j^2(t)\, \mathbf{x}_{\text{tar},j}.\end{aligned} \tag{10}$$

Its structure depends on the temporal modulation mediated by the periodic activation functions $\{\Phi_j(t) \in [0, 1];\ t \in [0, T_{\text{sim}}];\ j = 1, 2, \ldots, M\}$, where

$$\Phi_j(t) = \frac{1}{1 + \exp\left[-a_1 - a_2 \sin\left(\frac{2\pi(t+t_j)}{T}\right)\right]}. \tag{11}$$

The persistent (time-invariant) properties of the HSs can be characterized by a system of four-tuples

$$\left\{\left\langle \mathbf{x}_{\text{init},j},\ \mathbf{x}_{\text{interm},j},\ \mathbf{x}_{\text{tar},j},\ t_j \right\rangle;\ j = 1, 2, \ldots, M\right\}, \tag{12}$$

where $\mathbf{x}_{\text{init},j} \in \mathbb{R}^2$ determines the starting point of the j-th HS trajectory and $\mathbf{x}_{\text{tar},j} \in \mathbb{R}^2$ is the vector directed to the target destination. An additional aspect is the control point,

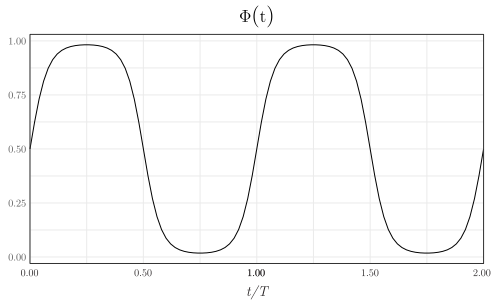

FIGURE 2. Time evolution of activation function $\Phi(t)$.

$\mathbf{x}_{\text{interm},j} \in \mathbb{R}^2$, which can be indirectly used to alter the curvature of $\mathbf{x}_j(t)$. Each $\Phi_j(t)$ is determined by the HS-specific phase shift t_j; all trajectories are characterized by the positive real-valued parameters T, a_1, and a_2. The parametric pair (a_1, a_2) determines the extent to which the logistic form tunes the periodic activity. Some important elementary facts must be considered before parametrization is suggested. As observed from (11), the choice $a_2 \gg a_1$ pushes Φ_j to the vicinity of asymptotic (boundary) values zero and one, which means that the limit points $\mathbf{x}_j|_{\Phi_j \to 0} = \mathbf{x}_{\text{init},j}$ and $\mathbf{x}_j|_{\Phi_j \to 1} = \mathbf{x}_{\text{tar},j}$ are never attainable. For some typical intermediate situation $\Phi_j = 1/2$, the weighted average $\mathbf{x}_j|_{\Phi \to 1/2} = (\mathbf{x}_{\text{init},j} + 2\mathbf{x}_{\text{interm},j} + \mathbf{x}_{\text{tar},j})/4$ is obtained, which underlines the known fact that $\mathbf{x}_{\text{interm},j}$ does not necessarily lie on the trajectory $\mathbf{x}_j(t)$. In Fig. 2, we depict an illustrative example of $\Phi_j(t)$, important for the parametric choices and further simulations.

2) LÉVY FLIGHT MOBILITY

Another variant of motion is provided by the Lévy model, where, in general, human mobility has been empirically observed to exhibit Lévy flight [19]. This combined motion is based on the following interpretation. If a UE travels to a certain relatively distant destination (in the city), the best direction of movement appears to be a straight line. Paradoxically, moving closer to a target point causes the information regarding the exact location of the UE to become vague; hence, another local search (though random) is required. In this precarious situation, a random strategy for the next route choice becomes the most efficient. The Lévy flight process arises as a repeated application of these two movement strategies. As UE moves are considerably subjective, and therefore unpredictable and interrupted by variable targets, we represent them through 2D Lévy flights.

In this segment, we present essential simulation details that illustrate this situation. Under the assumption that angular decisions are distributed uniformly and randomly, at each simulation step, the random angle of the j-th UE is generated by $\theta_j^{\text{LF}}(t) = 2\pi r_{\text{ang},j}(t)$, where $r_{\text{ang},j}(t)$ is drawn uniformly from [0, 1]. The instantaneous length of each iterative step is obtained as the product of the typical spatial step length ℓ_{u} and the random multiplier $r_{\text{stp},j}^{-1/\alpha_{\text{LF}}}$, where α_{LF} is the stability parameter and $r_{\text{stp},j}(t) \in [0, 1]$ is independently and uniformly generated. The position $\mathbf{x}_j^{\text{LF}}(t) \in \mathbb{R}^2$ (LF is the distinguishing label for a UE with this type of mobility) of the j-th UE is updated iteratively according to

$$\begin{aligned} \theta_j^{\text{LF}}(t) &\leftarrow 2\pi r_{\text{ang},j}(t), \\ \mathbf{x}_j^{\text{LF}}(t+\Delta t) &= \mathbf{x}_j^{\text{LF}}(t) + \\ &\quad + \ell_{\text{u}} \frac{\left(\cos(\theta_j^{\text{LF}}(t)),\ \sin(\theta_j^{\text{LF}}(t)\right)_{2D}}{[r_{\text{stp},j}(t)]^{1/\alpha_{\text{LF}}}}, \end{aligned} \tag{13}$$

where $(.,.)_{2D}$ is used to represent the vector in 2D.

An additional simulation detail is that a Lévy type of UE is not permitted to move a step that overcomes the rectangular boundaries of the system (i.e., no wrapping is assumed). Of course, the accumulation of limited movements could cause a marginal increase in the concentration of UEs in a border region. These subtle consequences are left for further study.

D. SYSTEM MODEL SIMPLIFICATIONS

Although the system model is designed to reflect the real HetNet as precisely as possible, there are still some simplified assumptions, which were introduced to decrease the computational complexity of the proposed algorithm.

- The spectrum bands of macrocells and small cells are completely separated to simplify the task of interference aware resource allocation for small cells. Such simplification is common for the real network deployment, especially with the advent of LTE-Unlicensed technology, which assumes utilization of unlicensed spectrum for the small cells.
- Round-robin resource scheduling is assumed to ensure that all simulation results will solely depend on the positions of small cells, rather than on other random parameters. Such simplification allows to ensure that SOM algorithm will find the HetNet topology with the most uniform QoE across coverage area.
- In real network scenario, UEs accept the QoE, which is not lower than some minimum threshold. In this paper, we assume that QoE requirements are the highest possible for each UE. This allows us to ensure that each UE will tend to maximize his throughput as much as possible. Such assumption is important for SOM algorithm to find the most effective network topology.
- In this paper, we limit our investigation to the two-tier HetNet architecture with fixed transmission powers of all macrocells and small cells. This allows us to reduce the algorithm complexity only to small cells positioning, which is the main goal of this paper. In our future research we will extend this problem to more tiers of coverage and flexible power adjustment of BSs, in order to reflect more realistic network scenarios.
- ECC-33 wireless channel propagation model is assumed as the best suitable for the chosen frequency range and urban environment. In real scenarios, the channel propagation models are different across coverage area due to

different type of landscape, height of buildings and link distances. However, this is very complex to simulate, so most of similar studies usually focus on the single channel propagation model.

- The UEs' mobility pattern is assumed to be the combination of Bézier curves and Lévy flights random movement to reflect more realistic behavior of UEs. Nevertheless, this pattern is still far from the real UEs' mobility, because it does not reflect the typical UEs' mobility in the big city. In our further research, we will implement more realistic patterns for UEs' mobility based on the machine learning algorithms.

IV. SELF-ORGANIZING MAP ALGORITHM: PROPOSED ALGORITHM FOR HETNET DEPLOYMENT OPTIMIZATION

In this section, we discuss the details of the application of a known SOM network to the problem of learning, where the parameters of the neurons define the coordinates of the cells of a small cell network.[1] If we know the trajectories of the UEs, then the SOM-mediated reduction projects these input data to the significantly reduced, yet representative, locations that are associated with the small cell positions in our application. Because the learning process is unsupervised (self-organized), it can reveal a hidden data structure, that is, it can perform categorization without a priori knowledge of the appropriate solutions. The ability of SOM networks to achieve self-organized states is equivalent to the ability to form clusters of input data. In our case, the data represent the moving coordinates of the UEs, and the persistent positions of the small cells belong to the locations of the neurons (or centers of clusters) forming the SOM content. To harmonize the concepts of the proposed model with the standard neural network dictionary, we can say that the neuron (which encodes the location of the small cell) whose position is closest to the data entry (for example, a UE position) is the best matching unit (BMU neuron). After each SOM iteration, the BMU vectors are modified; they are trained linearly to be closer to the appropriate data inputs. What is considered to be ''closest'' in the context of the problem being studied, of course, can be interpreted in different manners and is the primary idea and specificity of the proposed SOM application. In many cases, the proximity of the vector objects can simply be quantified by their distance defined for the metric space. Euclidean distance is frequently used as the standard in this area. However, the distance based on the purely geometric properties acquired by the conventional Euclidean prism is not suitable in cases where (i) heterogeneous properties (e.g., those associated with signal propagation) become important; (ii) other parties influence the process of indirect distance evaluation, e.g., based on the signal transmission between entities; (iii) if there is a manifestly functional asymmetry between a pair of entities (such as a transmitter-receiver) that violates symmetry, and thus makes the applicability of the distance concept difficult. Therefore, we propose a modified SOM application where the immediate SINR in (4) is employed instead of the traditional Euclidean distance. Based on the previous discussion, we assume that the SOM application in the HetNet network can be useful not only to improve HetNet coverage but also to explore other investigated metrics. In the next segment, we will briefly review the SOM preliminary studies and provide details of the SOM application in the HetNet deployment optimization process.

In the proposed SOM version and other implementations, the network consists of a set of neurons arranged on a 2D regular square mesh. To obtain an optimized (self-organized) 2D structure of the neurons (small cell equivalents), an iterative learning process is required. Synthetic data inputs presented to the SOM network are the trajectory samples of the UEs, which are partially random ($\mathbf{x}_j^{\mathrm{LF}}(t)$) or deterministic ($\mathbf{x}_j(t)$) in nature. In the following text, only $\mathbf{x}_j(t)$ symbols are used to represent the synthesized patterns for SOM learning. It is also worth mentioning that the standard theory of learning recognizes two types of training/learning practices, called sequential and batch algorithms. In this work, we use a sequential learning algorithm because it retains the time sequencing of the samples, which can be important for the analysis of periodic UE movements. Finally, we do not propose changing the macrocell topology structure, and thus the macro-UEs (the UEs that are associated at particular learning time steps with the macrocell) do not have any instantaneous impact on the SOM learning process.

The parallel dynamics of the moving UEs and the learning dynamics of the SOM are linked by specific time variables t' and t, where t' is a discrete-time variable describing the stages of SOM learning; the positions of the UEs acquired within the mobility models (mobility curves) are parametrized by the continuous (real) time variable t. The mutual relation between t and t' is given by $t' = \lfloor t/\Delta t \rfloor$, where Δt is the time step of the mobility model sampling.

In the following segment, we discuss the computational details of our specific SOM implementation. An SOM network is a recurrently evolving system of vectors–neurons $\mathcal{S}(t') \equiv \{\mathbf{m}_i(t') \in \mathbb{R}^2;\ i = 1, \ldots, N_{\mathrm{S}}\}$. The mobility models provide instantaneous data samples including $\{\mathbf{x}_1(t), \mathbf{x}_2(t), \ldots, \mathbf{x}_{N_{\mathrm{u}}}(t)\}$ consisting of N_{u} UE positions at a given time. The single-learning step for the occurrence and processing of data item $\mathbf{x}_j(t)$ includes:

(i) the *introduction* of particular input $\mathbf{x}_j(t)$ to the SOM. Here we emphasize the neuronal dependence not only on t' but also on the pair (j, t'), which stems from the fact that for serial update, the learning depends not only on t' but also on the order j of the respective data item $\mathbf{x}_j(t)$ (i.e., data items are visited in a typewriter manner).

(ii) the *identification* of the BMU vector $\mathbf{m}_{c(j,t')} \in \mathcal{S}(t')$ specified by its index $c(j, t') \in \{1, 2, \ldots, N_{\mathrm{S}}\}$;

[1] In this paper, we use the terms neuron and small cell interchangeably, as they represent the same element within the algorithm framework.

(iii) the *adjustment* of $\mathbf{m}_{c(j,t')}$ and neighbors $\mathbf{m}_i(t') \in \mathcal{S}(t')$ with the learning rate modified by the neighborhood function $h_{i,c(j,t')} \in \mathbb{R}^+$ (the form is specified below). The processes of BMU identification and the calculation of $h_{i,c(j,t')}$ depend on the distance (or SINR-based alternative) between the data samples and neurons. Given $\| \dots \|_{2D}$ is the Euclidean distance in 2D, then $c(j, t')$, the index representing the BMU, is identified using

$$c(j, t') \equiv \arg \min_{i=1,\dots,N_S} \| \mathbf{x}_j(t) - \mathbf{m}_i(t') \|_{2D}. \tag{14}$$

At this stage, neurons begin adapting their components in response to changes in their own network topology and changes in the data environment

$$\mathbf{m}_i(t'+1) = \mathbf{m}_i(t') + \alpha(t') h_{i,c(j,t')}(t') \left[\mathbf{x}_j(t) - \mathbf{m}_i(t') \right], \tag{15}$$

where $\alpha(t')$ is the learning rate and $i = 1, \dots, N_S$. However, not all neurons necessarily require modification; rather, local changes are typical for the SOM architecture. The time-varying limitations for the $\mathcal{S}(t')$ update are defined at the level of the neighborhood function

$$h_{i,c(j,t')}(t') \equiv \begin{cases} 1, & \text{if } \|\mathbf{m}_i(t') - \mathbf{m}_{c(j,t')}\|_{2D} \le \tilde{\sigma}(t') \\ 0, & \text{else.} \end{cases} \tag{16}$$

The convergence toward a static limit $\mathcal{S}^*$ is achieved by the decreasing learning rate (varying in a simulated annealing manner)

$$\alpha(t') = \alpha_{\text{start}} \, k_1^{t'}, \quad k_1 = \left(\frac{\alpha_{\text{tar}}}{\alpha_{\text{start}}} \right)^{1/T_{\text{sim}}}, \tag{17}$$

where $k_1 \in [0, 1]$ is a constant and $T_{\text{sim}} \gg T$ is a typical time scale, which at the same time characterizes not only the simulation of mobility but also the slowing of the learning (simulation length). The learning rate is delimited by the positive constants α_{start} and α_{tar}. The complementary heuristic $\tilde{\sigma}(t') = k_2 \alpha(t')$ is used for the process of the "narrowing" of the network focus. Because k_1 and k_2 have a significant impact on the convergence and results achieved, several meta-optimization tests have been applied to perform the selection of these constants.

If no algorithmic optimization (or pre-processing of input data) is performed for the respective SOM variant, the computational complexity of each iteration of the network can be given as $\mathcal{O}(N_S N_u^2)$ [10]. The above discussed literature (and references therein) reports that the main options to improve the SOM performance may be available in situations where the data environment is relatively sparse. We can mention that the computational load due to N_u^2 contribution can be simply reduced by representing the spatially dispersed UEs as the centroid using appropriate cluster mechanism. In lieu of detailed valuation of all UEs, we assume that some pre-treatment procedure from the class of coarse-graining procedures can also be applied.

We refer to the procedure presented above as the *conventional* application of SOM to the HetNet deployment optimization problem. However, as long as the Euclidean distance in (14) is the decisive factor for selection of the BMU, the Euclidean scenario finds somewhat limited application in the HetNet deployment problem. The problem of the optimal user-cell association can be extremely pronounced in the specific situations where the optimality UE association is achieved for a remote (in the Euclidean sense) BS. The reason for this seemingly paradoxical choice is the ability to gain a higher SINR of the receiver. In general, a remote BS can have a more favorable position in terms of ubiquitous interference, and can thus provide higher throughput to the associated UE.

Because this phenomenon cannot be captured using the conventional Euclidean metric, we choose to use the SINR "quasi-metric" formula. Our presented heuristic modification of the SOM is limited to one calculation step. To identify $c(j, t')$, we use the following *modified* variant of (14):

$$c(j, t') \equiv \arg \max_{i=1,2,\dots,N_S} f_{\text{SINR}}(\mathbf{x}_j(t), \mathbf{m}_i(t')), \tag{18}$$

where $f_{\text{SINR}}(.,.)$ represents the SINR function (see (5)) between the j-th UE and the i-th neuron (encoding the position of the small cell). Although this may not be optimal, the next steps of the SOM algorithm remain consistent with the conventional application. Regarding the modified form, the most serious doubts arise in connection with the relaxed metric properties that are irrevocably lost when the Euclidean description is changed to $f_{\text{SINR}}(.,.)$. Although the modified approach does not provide sufficiently precise arguments, this deficiency can be compensated by the empirical-computational, intuitive, and comparative approach (where e.g., a comparison of the conventional and modified SOMs is performed in the specific cases). A mathematical intuition can be based on a return to the metric concept. The inverse proportionality $1/\|\mathbf{x}_i - \mathbf{x}_j\|_{2D}^{\text{SINR}} \sim 1/f_{\text{SINR}}(\mathbf{x}_j, \mathbf{m}_i)$ is an elementary idea in this direction.

V. SIMULATION RESULTS

To evaluate the performance of the proposed approaches (based on the Euclidean metric and SINR metric), we distributed N_S small cells following the binomial point process (BPP) in an investigation region having an area of 1 km^2. A simple 2D BPP is typically used to model the positions of small cells of single tier or K-tier cellular networks. Unlike other stochastic geometry tools, the use of a BPP is becoming prevalent in coverage-limited scenarios to model a finite number of small cells distributed within the coverage area. As HetNets are commonly deployed in an increasingly irregular and random manner, modeling the locations of the small cells using tools from stochastic geometry would seem to be a fair solution to model the HetNet network performance. In general, the application of the BPP modeling approach in our discussion is twofold. First, the BPP modeling approach serves as a reference point for a performance comparison with the proposed

SOM application. Furthermore, a BPP also provides the initial HetNet topology for the SOM deployment optimization process.

The numerical tests were conducted using the 3GPP LTE-A specification and small cell parameters as described in [23] and [29] using Monte Carlo simulation. Throughout the simulation, we followed scenario #2a presented in [35], where the outdoor small cell deployment was considered jointly with the overlaid macrocell network. Separate frequency deployment of the macrocells and small cells was assumed. The detailed simulation parameters are introduced in Table 1.

TABLE 1. Simulation parameters.

Parameter	Value
Simulation method	Monte Carlo with 1000 trials
Simulation length (T_{sim})	$T_{\mathrm{sim}} = 800$ hours
Period T of $\Phi(t)$ function	$T = 0.5$ hours
Coverage area	$1km^2$
TX power (P_i)	macrocell: 40.3 W small cell: 6.3 W
BS height (h_{b})	macrocell: 45 m small cell: 2 m
UE antenna height (h_{r})	1 m
Antenna pattern	0 dBi (isotropic)
Carrier frequency (f)	macrocell: 2.1 GHz small cell: 3.5 GHz
Macro-layout (N_{M})	$N_{\mathrm{M}} = 2$ Inter-site distance $= 500m$
Small cell layout (N_{S})	$N_{\mathrm{S}} \in [2, 30]$
No. of RBs (N_{RB})	per macrocell: $N_{\mathrm{RB}} = 100$ RBs per small cell: $N_{\mathrm{RB}} = 20$ RBs
RB bandwidth (W)	180 kHz
System bandwidth	macrocell: 18 MHz small cell: 3.6 MHz
Path loss model	ECC-33
Thermodynamic temperature	290 K
Link scheduler	Round-robin
Number of Lévy Flight UE	100
Lévy flight mobility parameter (α_{LF})	$\alpha_{\mathrm{LF}} = 1.5$
Number of HSs	10, each consisting of 10 UEs
Total number of UE (N_u)	$N_u = 200$
Phase shift of HSs (t_j)	$t_j \in [-2, 2]$ hours (random)
SOM learning constant (k_1)	$k_1 = 0.9974$
SOM learning constant (k_2)	$k_2 = 16.67$
SOM learning constant (α_{start})	$\alpha_{\mathrm{start}} = 0.8$
SOM learning constant (α_{tar})	$\alpha_{\mathrm{tar}} = 0.002$

We first discuss in detail the specific manner in which the positions of the neurons were changed via sequential moves. Here, we illustrate the use-case with a Euclidean-based SOM; however, similar results were achieved for the SINR case. Even a visual inspection can indicate that the dynamics governing the radius of learning can have an important role in the efficiency of the convergence process. We initiated the SOM learning procedure with a neuron distribution following the BPP (Fig. 3a). Owing to the fast learning process and initial large radius, the neurons were grouped together and jointly traversed the region (see supplementary material)). As we proceeded with the review of the simulation snapshots, we observed that there were clearly identifiable early stages of the fast learning process where the instantaneous arrangement fluctuated irregularly with relatively low global throughput efficiency. Under this regime, the arrangements of the neurons reflected the combined effect of instantaneously large radius of learning (low competition, high plasticity), high learning rate, and a local cumulative effect of the mobile UEs, which can be localized in an extremely narrow region of the traffic funnel that acts as an intense temporary attractor for all neurons. Because the effect of stochasticity is ubiquitous, the analogy between our variant of SOM and the simulated annealing processes is clear. In simulated annealing, the temperature determines the probability of the transition to higher *energy* from some instantaneous configuration. In the specific search for the optimal configurations of neurons, the complex density deviations cause unpredictable *quasi-thermal* dynamics. The Lévy flight randomness, the regular perpetual motion of the traffic stream, and the funnel together with the gradual non-equilibrium reduction of the learning rate are the dynamical components that form the independent time scales shaping the environment of the neurons. From an intuitive viewpoint, the timing structure mixing the global and local trends resembles the standard annealing schedules. As can be observed, the *overheated* neurons wandered initially within a broad region of the search space (see video, supplementary material and Fig. 3b). When the learning rate became sufficiently slow (the movements of the neurons were less pronounced), a relatively sharp threshold transition occurred that was accompanied by a qualitative shift toward high learning (optimization) efficiency and considerable structural changes in the arrangement of the neural *degrees of freedom* (Fig. 3c). Structural tendencies supporting axial symmetry are readily observable. It can be seen that smaller, yet significant, changes due to slow learning have led to the disappearance of locally unstable neuro-spatial deviations. Gradually, almost axially symmetrical arrangements are formed that reflect the overall symmetry of the environment occupied by the moving UEs. Finally, when the learning rate approached zero, neural movements became limited (virtually static), and thus was considered adapted, and a fixed arrangement was obtained that integrated all the environmental movements. As a by-product, axial symmetry of the neural arrangements was developed, which was externally dictated by the axis of the transport funnel and the macrocell axis (Fig. 3d). The principal positive simulation finding is that final neuronal alignment is reflected by the enhanced global throughput value despite the fact that the global fitness function is not explicitly defined for the SOM network. By observing small, residual, deviations from the regularity of the agreements, we conclude that barriers to optimization appear to be an obstacle to the relaxation and further improvement of the learning (optimization) process.

Figure 4 explains how the SOM algorithm learns over time to improve the performance of deploying small cells.

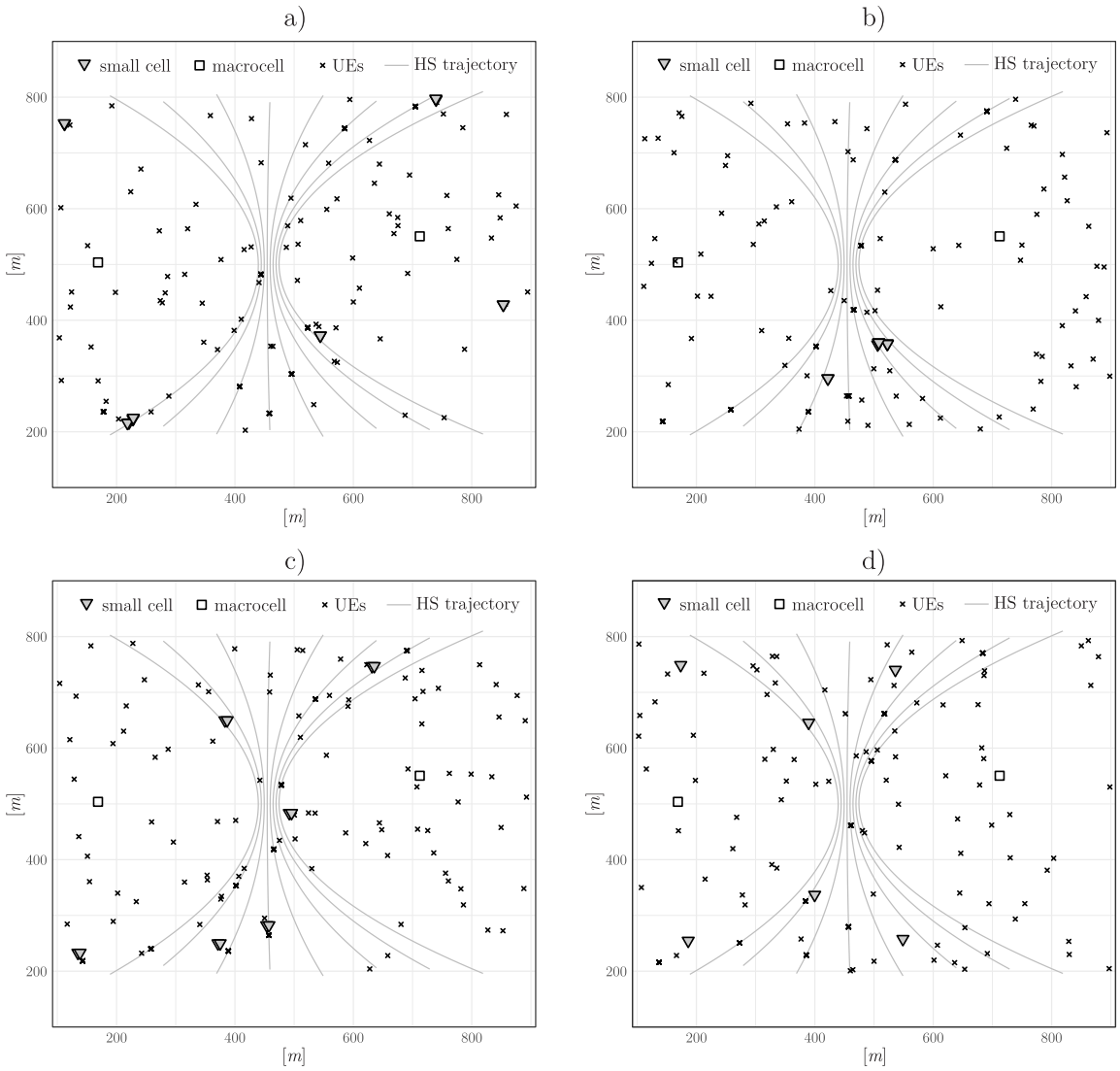

FIGURE 3. Visual illustration of the SOM learning framework. Four plots indicate the small cell topology design at different stages of the SOM learning process. The subplot notation a), b), c), and d) correspond to the different SOM learning stages depicted in Fig. 4.

Starting from a random initial BPP placing (point **a**), the algorithm attempts to determine improved small cell placement. It is noticed that initially the algorithmic process is "confused" regarding how to place the small cells effectively. Such behavior is normal considering the unsupervised nature of an SOM. Figuratively, its cognitive abilities are at the level of a baby taking its first steps. Thus, the average throughput per user is significantly less compared with BPP during this learning period. After several iterations, the SOM algorithm determines where small cells should be placed to improve the experience of the UEs. Then, in a short time, the average throughput per user increases more than twice (from point **b** to point **c**), reaching values unattainable by BPP or other existing solutions. Point **c** corresponds to the local optimal HetNet coverage. From point **c** to point **d**, we do not observe any significant changes in the network performance because the SOM algorithm performs only minor shifting of the small cells to determine the best configuration. The difference in throughput between points **c** and **d** can be neglected due to the additional noise in the system caused by the UE mobility.

To compare the performance of the proposed SOM algorithm with existing stochastic geometry algorithms, we conducted simulations for three key parameters of HetNet: average throughput per user, fairness of resource allocation, and coverage probability. Figure 5 displays the results for the average throughput per user. It is clear that both SOM implementations outperformed the conventional BPP deployment, especially when the number of small cells increased. However, the difference between the two SOM metrics (SINR and Euclidean distance) was minimal and diminished for a large number of small cells. This phenomenon can be explained by the features of the simulation scenario. Increasing the number of small cells per area naturally results in a smaller cell size. If the cell size is sufficiently

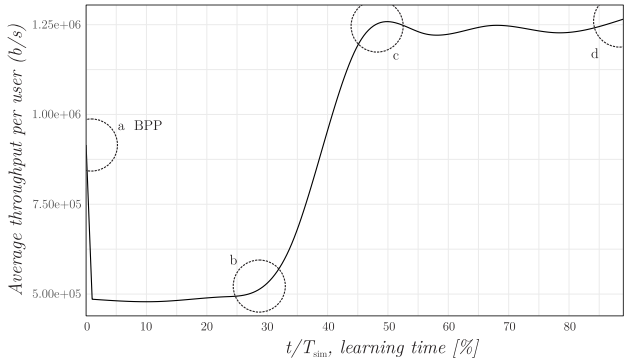

FIGURE 4. Average throughput per user vs. SOM training time. Initially, the throughput dramatically decreases owing to the large radius and fast learning rate. However, the throughput gradually improves over the learning time duration. Notations a), b), c), and d) indicate different stages of the SOM learning process.

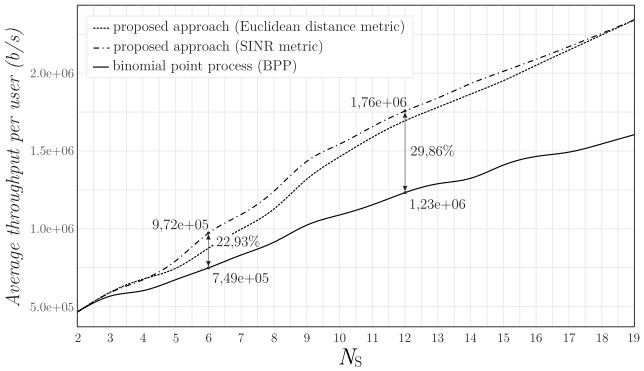

FIGURE 5. Average throughput per UE vs. number of small cells. The studied algorithms provide enhanced results in terms of the throughput compared to BPP. However, the differences between the proposed metrics diminish for $N_S = 18$.

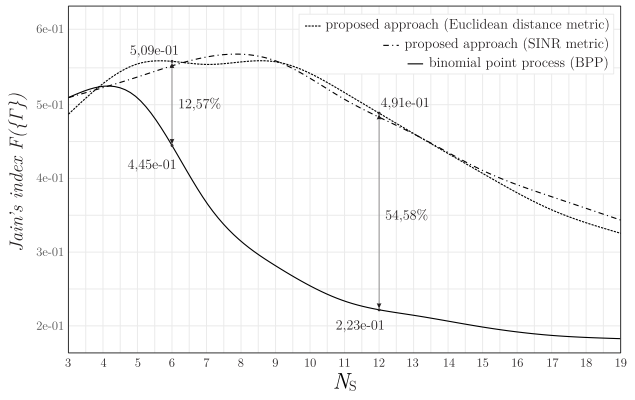

FIGURE 6. Fairness vs. number of small cells for the studied algorithms. Clearly, the fairness decreases with an increased number of small cells. The higher the number of small cells, the greater the degree of freedom with user-cell association present, resulting in a decrease of fairness. The proposed algorithms provide superior results in terms of fairness compared with BPP.

small, we can assume that the SINR variation over the cell is negligible; SINR depends only on the received interference power. Received interference power increases proportionally to the Euclidean distance from the serving BS. Thus, Euclidean and SINR metrics are closely related and indicate similar performance for all simulated parameters.

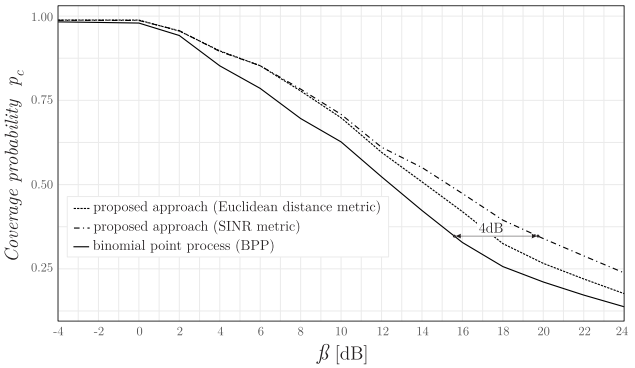

FIGURE 7. Coverage probability for $N_S = 6$. SOM with proposed SINR metric performs the best. SOM with Euclidean metric performs better than BPP, yet significantly worse than using the SINR metric.

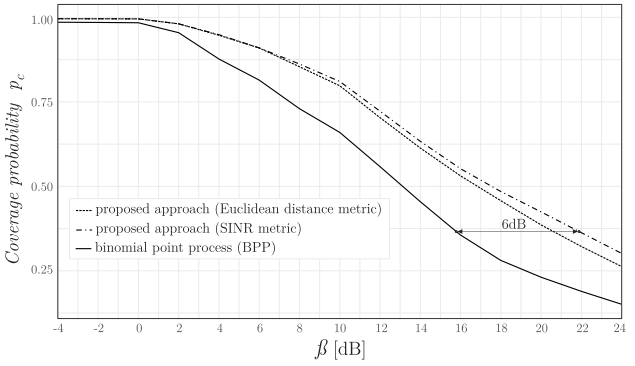

FIGURE 8. Coverage probability for $N_S = 12$. SOM with both proposed metrics significantly outperforms BPP topology design. However, the differences between the metrics do diminish to a virtually negligible level.

Although Fig. 5 displays the advantage of SOMs, this does not necessarily mean that the most effective network configuration is achieved. It is common in HetNet to observe a situation where some UEs experience significantly higher throughput than average, whereas the throughput for other UEs is significantly below average. Thus, it is important to determine what configuration of small cells provides a more uniform throughput distribution among the UEs. Figure 6 depicts fairness versus the number of small cells for the studied algorithms. As can be observed from the results, the fairness decreases with an increasing number of small cells. This result was expected because UEs tend to stay less time in cells with a smaller radius, resulting in a higher probability of small cell overload or underutilization. Nevertheless, the proposed SOM algorithm demonstrates superior fairness compared to the BPP, regardless of the metric used. This advantage can be explained by the self-optimization capability of SOM. Unlike the BPP, which tends to fill the area without coverage holes, the SOM algorithm determines the positions of the small cells according to the probability of traffic demand in each area. Thus, SOM avoids small cell placement in areas where traffic demand does not exceed the capabilities of the macrocell. Therefore, considerably less inequality is observed in HetNet if the small cell layout follows the SOM topology.

Summarizing the results in Figs. 5 and 6, we can conclude that the effective number of small cells, which satisfy both criteria, is from 6 to 12. If the number of small cells is less than 6, the average throughput is overly low, even though the fairness distribution is sufficiently fair. Conversely, increasing the number of small cells to more than 12 can improve average throughput per user. However, the fairness in this case degrades significantly and the majority of UEs will experience unacceptable throughput.

For the two chosen cases, we compare the coverage probability versus target value of SINR in Figs. 7 and 8. The difference between the two simulated scenarios indicates that each additional small cell decreases the SINR; this is because number of interfering transmitters is increased. These results correlate well with the results presented in Figs. 5 and 6 and help to highlight the advantages of the proposed SOM over the conventional BPP deployment of small cells. Coverage probability is an important parameter; it provides insight regarding the percentage of UEs who experience acceptable SINR values. Given the direct relation between SINR and throughput, it is clear that SOM deployment provides an improvement of overall HetNet performance compared with conventional BPP deployment regardless of the simulated scenario. Considering the fact that quasi-optimal network configuration depends solely on the statistics of traffic demand in a particular coverage area, the proposed SOM algorithm is a promising tool for the task of HetNet topology design.

VI. CONCLUSION

A new approach to HetNet topology design based on unsupervised learning was presented in this paper. The proposed approach aims to address the challenge of optimal small cell placement within a target coverage area considering the statistics of traffic load HSs and co-channel interference among the UEs of different cells. The HetNet topology is designed based on the unsupervised SOM algorithm. The proposed SOM-based algorithm can adjust the position of small cells using one of two criteria: average UE distance from the serving BS or average SINR perceived by each UE. The important advantage of this approach is that it can be trained for any type of UE mobility pattern and provides unlimited opportunities for HetNet design for any type of urban environment. Simulations were conducted to compare the performance of the proposed SOM algorithm with the stochastic BPP algorithm, which has been widely considered as the most feasible solution for two-tier HetNet topology design. We compared both algorithms based on the average throughput per UE, fairness of throughput distribution among UEs, and coverage probability. The results confirm that the proposed approach outperforms the BPP for all simulated scenarios. In particular, the SOM gain in average throughput was noticeable with an increase in the number of small cells. However, for a large number of small cells, the probability of unfair throughput distribution increased, due to the reduced coverage area of the small cells. The results of coverage probability indicate that in the case of 6 small cells, SOM provides SINR gains of up to 2 dB for the Euclidean distance metric and up to 4 dB for the SINR metric. In the case of 12 small cells, SOM provides an SINR gain of up to 6 dB for both studied metrics, meaning that for small cells, the difference between SINR and the Euclidean metric is negligible. In further research, we will provide additional insight into the HetNet coverage planning for different scenarios using unsupervised learning algorithms with different performance criteria.

REFERENCES

[1] P. Bhat *et al.*, "LTE-advanced: An operator perspective," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 104–114, Feb. 2012.
[2] B. Soret, H. Wang, K. I. Pedersen, and C. Rosa, "Multicell cooperation for LTE-advanced heterogeneous network scenarios," *IEEE Wireless Commun.*, vol. 20, no. 1, pp. 27–34, Feb. 2013.
[3] Y.-W. Chen and W.-H. Lu, "Small cells placement scheme in metropolitan area environment," in *Proc. 18th Asia–Pacific Netw. Oper. Manage. Symp. (APNOMS)*, Oct. 2016, pp. 1–6.
[4] J. Hoadley and P. Maveddat, "Enabling small cell deployment with HetNet," *IEEE Wireless Commun.*, vol. 19, no. 2, pp. 4–5, Apr. 2012.
[5] W. Guo and S. Wang, "Interference-aware self-deploying femto-cell," *IEEE Wireless Commun. Lett.*, vol. 1, no. 6, pp. 609–612, Dec. 2012.
[6] "Small cell siting challenges," Small Cell Forum, Dursley, U.K., Tech. Rep. 192.09.01, 2017.
[7] C. Kim, R. Ford, and S. Rangan, "Joint interference and user association optimization in cellular wireless networks," in *Proc. 48th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2014, pp. 511–515.
[8] T. Kohonen, "The self-organizing map," *Neurocomputing*, vol. 21, nos. 1–3, pp. 1–6, 1998.
[9] S. Kaski, T. Honkela, K. Lagus, and T. Kohonen, "WEBSOM—Self-organizing maps of document collections," *Neurocomputing*, vol. 21, nos. 1–3, pp. 101–117, 1998.
[10] T. Kohonen, "Essentials of the self-organizing map," *Neural Netw.*, vol. 37, pp. 52–65, Jan. 2013.
[11] T. Kohonen *et al.*, "Self organization of a massive document collection," *IEEE Trans. Neural Netw.*, vol. 11, no. 3, pp. 574–585, May 2000.
[12] C. Amerijckx, M. Verleysen, P. Thissen, and J. Legat, "Image compression by self-organized Kohonen map," *IEEE Trans. Neural Netw.*, vol. 9, no. 3, pp. 503–507, May 1998.
[13] O. Depren, M. Topallar, E. Anarim, and M. K. Ciliz, "An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks," *Expert Syst. Appl.*, vol. 29, no. 4, pp. 713–722, 2005.
[14] C. Koutsougeras, Y. Liu, and R. Zheng, "Event-driven sensor deployment using self-organizing maps," *Int. J. Sensor Netw.*, vol. 3, no. 3, pp. 142–151, 2008.
[15] T. M. Salem, S. Abdel-Mageid, S. M. Abdel-Kader, and M. Zaki, "ICSSSS: An intelligent channel selection scheme for cognitive radio ad hoc networks using a self organized map followed by simple segregation," *Pervasive Mobile Comput.*, vol. 39, pp. 195–213, Aug. 2017.
[16] Z. Zhang, K. Long, and J. Wang, "Self-organization paradigms and optimization approaches for cognitive radio technologies: A survey," *IEEE Wireless Commun.*, vol. 20, no. 2, pp. 36–42, Apr. 2013.
[17] X. X. W. Li and V. C. M. Leung, "Artificial intelligence-based techniques for emerging heterogeneous network: State of the arts, opportunities, and challenges," *IEEE Access*, vol. 3, pp. 1379–1391, 2015.
[18] I. Rhee, M. Shin, S. Hong, K. Lee, S. J. Kim, and S. Chong, "On the Levy-walk nature of human mobility," *IEEE/ACM Trans. Netw.*, vol. 19, no. 3, pp. 630–643, Jun. 2011.
[19] C. Song, T. Koren, P. Wang, and A.-L. Barabási, "Modelling the scaling properties of human mobility," *Nature Phys.*, vol. 6, no. 10, pp. 818–823, 2010.
[20] P. Tian, H. Tian, L. Gao, J. Wang, X. She, and L. Chen, "Deployment analysis and optimization of macro-pico heterogeneous networks in LTE-A system," in *Proc. 15th Int. Symp. Wireless Pers. Multimedia Commun.*, Sep. 2012, pp. 246–250.

[21] S. Y. Shin and I. F. M. Zain, "Cellular network planning for heterogeneous network using geo-clustering algorithm," in *Proc. Int. Conf. ICT Converg. (ICTC)*, Oct. 2013, pp. 448–449.
[22] X. Li, X. Tang, C.-C. Wang, and X. Lin, "Gibbs-sampling-based optimization for the deployment of small cells in 3G heterogeneous networks," in *Proc. 11th Int. Symp. Workshops Modeling Optim. Mobile, Ad Hoc Wireless Netw. (WiOpt)*, May 2013, pp. 444–451.
[23] M. H. Qutqut, H. Abou-zeid, H. S. Hassanein, A. M. Rashwan, and F. M. Al-Turjman, "Dynamic small cell placement strategies for LTE heterogeneous networks," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Jun. 2014, pp. 1–6.
[24] Y. Park *et al.*, "Effective small cell deployment with interference and traffic consideration," in *Proc. IEEE 80th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2014, pp. 1–5.
[25] A. Jaziri, R. Nasri, and T. Chahed, "System-level analysis of heterogeneous networks under imperfect traffic hotspot localization," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9862–9872, Dec. 2016.
[26] I. Bahçeci, "On femto-cell deployment strategies for randomly distributed hotspots in cellular networks," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2014, pp. 2330–2335.
[27] L. Zhou *et al.*, "Green cell planning and deployment for small cell networks in smart cities," *Ad Hoc Netw.*, vol. 43, pp. 30–42, Jun. 2016.
[28] T. Nakamura *et al.*, "Trends in small cell enhancements in LTE advanced," *IEEE Commun. Mag.*, vol. 51, no. 2, pp. 98–105, Feb. 2013.
[29] H. Holma and A. Toskala, *LTE for UMTS-OFDMA and SC-FDMA Based Radio Access*. Hoboken, NJ, USA: Wiley, 2009.
[30] V. S. Abhayawardhana, I. J. Wassell, D. Crosby, M. P. Sellars, and M. G. Brown, "Comparison of empirical propagation path loss models for fixed wireless access systems," in *Proc. IEEE 61st Veh. Technol. Conf. (VTC-Spring)*, vol. 1. May/Jun. 2005, pp. 73–77.
[31] S. F. Yunas, W. H. Ansari, and M. Valkama, "Technoeconomical analysis of macrocell and femtocell based HetNet under different deployment constraints," *Mobile Inf. Syst.*, vol. 2016, May 2016, Art. no. 6927678.
[32] D.-M. Chiu and R. Jain, "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks," *Comput. Netw. ISDN Syst.*, vol. 17, no. 1, pp. 1–14, 1989.
[33] C. Guang G. Jian, and D. Wei, "A time-series decomposed model of network traffic," in *Advances in Natural Computation (ICNC)* (Lecture Notes in Computer Science), vol. 3611, L. Wang, K. Chen, and Y. S. Ong, Eds. Berlin, Germany: Springer, 2005, pp. 338–345.
[34] W. Wang, J. Wang, M. Wang, B. Wang, and W. Zhang, "A realistic mobility model with irregular obstacle constraints for mobile ad hoc networks," *Wireless Netw.*, to be published.
[35] *Small Cell Enhancements for E-UTRA and E-UTRAN—Physical Layer Aspects*, document TR 36.872, 3GPP, 2013.

JURAJ GAZDA was a Guest Researcher with Ramon Llull University, Barcelona, Spain, and also with the Technical University of Hamburg-Harburg, Germany, under the supervision of Prof. H. Rohling. He was also involved in the development activities of Nokia Siemens Networks. In 2017, he was recognized as the Best Young Scientist of the Technical University of Košice. He is currently an Associate Professor with the Faculty of Electrical Engineering, Technical University of Košice, Slovakia. His research interests include spectrum pricing, techno-economic aspects of cognitive networks, co-existence of heterogeneous networks, and application of machine learning principles in 5G networks.

EUGEN ŠLAPAK received the B.A. and M.S. degrees in informatics from the Technical University of Košice, Košice, Slovakia, in 2016 and 2018, respectively. His master's thesis focused on the use of self-organizing maps for HetNet topology design, under the supervision of Prof. J. Gazda. He works part-time with the research team at the Intelligent Information Systems Laboratory, Technical University of Košice.

GABRIEL BUGÁR is currently a Research Assistant with the Faculty of Electrical Engineering and Informatics, Technical University of Košice, Slovakia. His research interests include cognitive networks, image processing, and e-learning.

DENIS HORVÁTH was a Junior Scientist with the Department of Theoretical Physics, Slovak Academy of Sciences, and then with the Department of Theoretical Physics and Astrophysics, P. J. Šafárik University, Košice, Slovakia. He is currently a member of the Research Group, Center of Interdisciplinary Biosciences, P. J. Šafárik University. He is also an experienced researcher, having worked in the interdisciplinary team for almost 25 years. His present mission is to develop and implement theoretical models of cognitive networks, artificial markets, and cancer progression. He has extensive experience in data analysis and numerical modeling, theory and simulation of complex and biological systems, and solutions of optimization problems.

TARAS MAKSYMYUK received the B.A. degree in telecommunications, the M.S. degree in information communication networks, and the Ph.D. degree in telecommunication systems and networks from Lviv Polytechnic National University, Lviv, Ukraine, in 2010, 2011, and 2015, respectively. He received the Post-Doctoral Fellowship from the Internet of Things and Cognitive Networks Lab, Korea University, under supervision of Prof. M. Jo. He is currently an Assistant Professor with the Telecommunications Department, Lviv Polytechnic National University. His research interests include cognitive radio networks, LTE in unlicensed spectrum, mobile cloud computing, massive MIMO, 5G heterogeneous networks, software-defined radio access networks, Internet of Things, big data, and artificial intelligence. He was recognized as the Best Young Scientist of Lviv Polytechnic National University in 2015, received the Lviv State Administration Prize for outstanding scientific achievements and contribution in 2016, and Lviv Metropolitan Prize for Best Scientists in 2017. He is currently an Associate Editor of the IEEE *Communications Magazine*, and an Editor of the *International Journal of Internet of Things and Big Data*.

MINHO JO received the B.A. degree from the Department of Industrial Engineering, Chosun University, South Korea, in 1984, and the Ph.D. degree from the Department of Industrial and Systems Engineering, Lehigh University, USA, in 1994. He is currently a Professor with the Department of Computer Convergence Software, Korea University, Sejong, South Korea. He is one of the founders of the Samsung Electronics LCD Division. He has published over 100 publications in reputed journals/magazines and international conferences. His current research interests include IoT, LITE-Unlicensed, artificial intelligence and big data in the IoT, blockchain and security, HetNets in 5G, cloud computing, wireless energy harvesting, 5/6G wireless communications, autonomous cars, and optimization and probability in networks. He received the Headong Outstanding Scholar Prize in 2011. He was the Vice President of the Institute of Electronics of the Korea Information Processing Society. He is currently the Vice President of the Korean Society for Internet Information. He is currently an Editor of the IEEE Wireless Communications, an Associate Editor of the IEEE Access, the IEEE Internet of Things Journal, and *Security and Communication Networks* (Wiley), and *Wireless Communications and Mobile Computing* (Hindawi). He is the Founder and Editor-in-Chief of *KSII Transactions on Internet and Information Systems* (SCI and Scopus indexed).