Spectral- and Energy-Efficient Antenna Tilting in a HetNet using Reinforcement Learning

Weisi Guo†, Siyi Wang§, Yue Wu§, Jonathan Rigelsford§, Xiaoli Chu§, Tim O’Farrell§
University of Warwick†, University of Sheffield§, United Kingdom
Email: weisi.guo@warwick.ac.uk†, {siyi.wang, yue.wu, j.m.rigelsford, x.chu, t.ofarrell}@sheffield.ac.uk§

Abstract—In cellular networks, balancing the throughput among users is important to achieve a uniform Quality-of-Service (QoS). This can be accomplished using a variety of cross-layer techniques. In this paper, the authors investigate how the down-tilt of basestation (BS) antennas can be adjusted to maximize the user throughput fairness in a heterogeneous network, considering the impact of both a dynamic user distribution and capacity saturation of different transmission techniques.

Finding the optimal down-tilt in a multi-cell interference-limited network is a complex problem, where stochastic channel effects and irregular antenna patterns has yielded no explicit solutions and is computationally expensive.

The investigation first demonstrates that a fixed tilt strategy yields good performances for homogeneous networks, but the introduction of HetNet elements adds a high level of sensitivity to the tilt dependent performance. This means that a HetNet must have network-wide knowledge of where BSs, access-points and users are. The paper also demonstrates that transmission techniques that can achieve a higher level of capacity saturation increases the optimal down-tilt angle.

A distributed reinforcement learning algorithm is proposed, where BSs do not need knowledge of location data. The algorithm can achieve convergence to a near-optimal solution rapidly (6-15 iterations) and improve the throughput fairness by 45-56% and the energy efficiency by 21-47%, as compared to fixed strategies. Furthermore, the paper shows that a tradeoff between the optimal solution convergence rate and asymptotic performance exists for the self-learning algorithm.

I. INTRODUCTION

Over the past decade, the complexity of the cellular radio-access-network (RAN) architecture has increased. Traditionally, the RAN has largely been an irregular homogeneous network. In recent years, multi-tier heterogeneous networks (HetNets) have emerged, and the RAN can include a variety of overlapping basestations (BSs) and access-points (APs). The motivation behind HetNets is that they can deliver an improved system capacity [1] at a lower energy consumption and cost expenditure level [2].

This paper considers how to balance the throughput for users across a RAN using antenna tilt adjustment [3]. Not only is balancing throughput important from a Quality-of-Service (QoS) perspective, but it has also been shown that by providing a more even throughput profile, the transmission energy efficiency can be improved [4].

A. Challenges

In order to deliver a uniform throughput to users inside the RAN, a variety of cross-layer techniques can be employed, such as: scheduling, soft-frequency-reuse, relaying, and antenna down-tilt adjustment. The paper considers how the down-tilt of BS antennas can be adjusted to maximize the throughput fairness in a multi-cell HetNet with a dynamic user-equipment (UE) distribution.

The challenges of obtaining the optimal down-tilt for multiple mutually interfering BSs are as follows:

- the stochastic nature of multi-path and shadow fading, as well as UE mobility, makes UE location difficult to estimate with sufficient accuracy for any tilt algorithms that rely on this information;
- the optimal down-tilt angle for mutually interfering BSs is difficult to compute explicitly even with UE location information;
- the actions of one BS affects not only its own performance, but the performance of all neighboring BSs through interference;
- the irregular radiation pattern of array antennas makes the optimization task complex;
- rapid convergence in numerical algorithm is needed in a live network environment, where conditions are changing.

Given a down-tilt resolution of \( \theta \) degrees and that each BS is affected by the actions of \( N_i \) interfering cells, the number of computation cycles for throughput is in the order of \( \theta^{2N} \) (a typical value of \( 20^{18} = 2.6 \times 10^{23} \)).

A further challenge is to consider HetNets, where cross-tier coordination is required for joint-optimal performance. However, the increased network complexity and cross-vendor coordination issues make coordination challenging. An alternative strategy employs distributed machine-learning algorithms to let nodes self-optimize their performance based on local observations. This is a feature of the wider Self-Organizing-Networks (SON) proposal [5].

B. Related Work

Existing research on antenna down-tilt optimization has focused on homogeneous networks and has employed either of the following strategies:

- **Fixed Strategy** utilizes prior or updated knowledge of traffic models to employ a constant antenna down-tilt for each BS [3] [4] [6] [7].
- **Adaptive Strategy** utilizes reinforcement learning to adjust the antenna down-tilts depending on the quasi-static traffic conditions [8].

What has been lacking in existing work is the analysis of how HetNet elements, such as femto-cells can impact the fixed and
self-learning strategies. This paper will show that the fixed strategy can achieve a reasonably good performance without updated knowledge of traffic models, but performs poorly when HetNet nodes, such as femto-cells are introduced. The paper will then focus on how the adaptive learning algorithm can improve on this.

C. Contribution

The paper proposes a novel exploration-based reinforcement learning that adapts its down-tilt strategy based on the local network conditions and the capacity saturation level of the transmission scheme. In the reinforcement learning algorithm, a tradeoff between exploitation and exploration is considered, and its impact on the asymptotic throughput performance and variance is shown. This is compared to the fixed strategy to show the relative advantages and disadvantages of solutions.

The system model is presented in Section II, which considers a dynamic multi-cell-multi-user (MCMU) setup. This is followed by examining the fixed strategy approach in Section III, and reinforcement learning approach in Section IV. For both approaches, the paper presents the key throughput and energy efficiency results, in the context of both a homogeneous and a heterogeneous network.

II. SYSTEM MODEL

A. Network Architecture and Layout

The paper considers an orthogonal-frequency-division-multiple-access (OFDMA) based Long-Term-Evolution (LTE) system with a number of co-channel micro-BSs deployed in a traditional hexagonal grid layout with wrap-around implementation. This has been demonstrated to yield a sufficiently accurate interference and performance model [9]. There are a number of indoor Femto-cell APs and UEs distributed in a hotspot area, which will be explained in further detail later. Figure 1a gives an example illustration of 2 BSs with UEs and an underlay of co-channel APs.

Link-layer capacity is calculated using the appropriate adaptive modulation and coding schemes [10]. System-layer throughput \( R \) is calculated using a round-robin scheduler in a bench-marked system simulator, with further details given in Table I.

B. Definitions and Parameters

The system considers an overlapping UE distribution, whereby some UEs are distributed uniformly over the network and a number of UEs are concentrated in a region, known as a hotspot. Typically, the APs are deployed at the centre of the hotspot. The modelled RAN has the following network elements and their associated parameter values can be found in Table I:

- **Micro-BS**: the system considers the actions of a BS \( n \in N_{BS} \), where there is a set of \( N_{BS} \) co-channel BSs.
- **Receiving Nodes**: each BS serves \( m \in M_{UE} \) UEs. There are \( M_{hotspot} \) UEs that are deployed according to a poisson

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth</td>
<td></td>
<td>20 MHz</td>
</tr>
<tr>
<td>Frequency</td>
<td></td>
<td>2.1 GHz</td>
</tr>
<tr>
<td>No. Micro-BS</td>
<td>( N_{BS} )</td>
<td>19 Hexagonal</td>
</tr>
<tr>
<td>No. Interference BS</td>
<td>( N_{I} )</td>
<td>18 Wrap-around</td>
</tr>
<tr>
<td>No. Femto-AP per BS</td>
<td>( K_{AP} )</td>
<td>3 (average)</td>
</tr>
<tr>
<td>UEs per BS</td>
<td>( M_{UE} )</td>
<td>10 (average)</td>
</tr>
<tr>
<td>Micro-BS Coverage Radius</td>
<td>( r )</td>
<td>500m</td>
</tr>
<tr>
<td>Transmission Scheme</td>
<td>( \lambda )</td>
<td>SISO, MIMO</td>
</tr>
<tr>
<td>Propagation Model</td>
<td></td>
<td>WINNER [12]</td>
</tr>
<tr>
<td>AWGN Power</td>
<td>( W )</td>
<td>( 6 \times 10^{-17} ) W</td>
</tr>
<tr>
<td>Shadow Fading Variance</td>
<td>( \sigma_s^2 )</td>
<td>9 dB</td>
</tr>
<tr>
<td>BS Antenna Height</td>
<td></td>
<td>20m</td>
</tr>
<tr>
<td>BS Antenna Pattern</td>
<td>( A )</td>
<td>[13]</td>
</tr>
<tr>
<td>BS Antenna Down-tilt Range</td>
<td>( \Theta )</td>
<td>( 0 - 20^\circ )</td>
</tr>
<tr>
<td>BS Transmit Power</td>
<td>( P_{t,BS} )</td>
<td>10W</td>
</tr>
<tr>
<td>AP Transmit Power</td>
<td>( P_{t,AP} )</td>
<td>0.1W</td>
</tr>
<tr>
<td>AP Antenna Height</td>
<td></td>
<td>1.5 m</td>
</tr>
</tbody>
</table>
distribution. The probability of a UE being a distance $d_m = D$ away from the serving BS is:

$$p_{\text{hotspot}}(d_m = D) = \frac{\lambda D}{D!} e^{-\lambda},$$  

(1)

where $\lambda$ is the distance from the centre of the hotspot to the serving BS. The remaining UEs ($M_{\text{UE}} - M_{\text{hotspot}}$) are deployed randomly and uniformly.

- **Femto-APs**: the system also include an under-lay of co-channel APs $k \in K_{\text{AP}}$, which follow the same distribution as given by (1), as shown in Fig. 1b. The paper assumes that the APs only cause interference to the outdoor UEs and transmit at full-buffer.

- **Down-Tilt**: each BS can adjust its down-tilt ($\theta_{n,t} \in \Theta$) in a range of down-tilt values.

- **Throughput**: at any particular time $t$, a BS can achieve a maximum aggregate throughput $R_{\text{BS},n,t,\theta} = \sum_{m=1}^{M_{\text{UE}}} R_{\text{UE},m,n,t,\theta}$ as a result of the down-tilt $\theta_{n,t}$.

- **Traffic**: at any particular time $t$, a BS experiences a traffic intensity $R_{\text{traffic},n,t}$. Traffic intensity is the aggregate traffic volume of circuit- and packet-switched data.

- **Constraint (Load)**: each BS experiences a load, which is a ratio of the traffic intensity and the maximum throughput achievable:

$$L_{n,t,\theta} = \frac{R_{\text{traffic},n,t}}{R_{\text{BS},n,t,\theta}}.$$

(2)

In order to satisfy the traffic demand, the value of $L_{n,t,\theta} \in [0, 1]$.

- **Throughput Fairness**: the UEs in each BS has a fairness index [14], which is a ratio of the cell-edge UE throughput (bottom 5%) and the cell-mean UE throughput:

$$F_{n,t,\theta} = \frac{M_{\text{UE}} R_{\text{UE},5\%,m,n,t,\theta}}{\sum_{m=1}^{M_{\text{UE}}} R_{\text{UE},m,n,t,\theta}}.$$

(3)

In order to maximize the throughput fairness across the coverage area, the paper attempts to max($F_{n,t,\theta}$) $\forall n, t$.

- **Energy Efficiency (EE)**: the transmission EE of a BS at a particular time $t$ is defined by the ratio between the transmit power used and throughput achieved [4] [15]:

$$\text{EE}_{n,t,\theta} = \sum_{m \in M_{\text{UE}}} \left( \frac{P_{\text{BS},m,n,t}}{R_{\text{UE},m,n,t,\theta}} \right).$$

(4)

The paper now examines how both the fixed strategy and adaptive learning strategy can provide a solution to the optimal antenna-tilting problem.

### III. Fixed Strategy

**A. Fixed Deployment Strategy**

The paper now considers the fixed strategy approach, whereby each BS employs a constant down-tilt angle, which is determined from prior-knowledge, e.g., experience, simulation results. Figure 2(a) and Figure 2(b) show the throughput fairness and down-tilt variation respectively, with hotspot location ($\lambda$) for half the users distributed in a hotspot manner ($M_{\text{hotspot}} = 0.5$), according to (1).

For a **homogeneous network** of micro-BSs, the maximum throughput fairness index is approximately 0.5 and the associated antenna down-tilt is between 5 to 1 degrees, depending on the hotspot location. These observations broadly agree with the results of previous findings in [4] [7]. It also shows that an improvement in fairness also leads to an improvement in the transmit spectral- and energy-efficiency. The resulting EE improvement is up to 35% compared to a fixed baseline tilt angle of 2°. The trend is that a closer hot-spot to the serving BS requires a steeper down-tilt to serve the hotspot UEs.

For a **heterogeneous network** of micro-BSs and femto-APs, the maximum fairness index is severely degraded due to the excess interference caused by APs. The impact is very high for hotspot locations far from the serving BS (below 0.1). The maximum fairness antenna down-tilt is between 16 to 2 degrees, depending on the hotspot and associated APs location. The trend is that a hot-spot closer to the serving BS requires a steeper down-tilt. The resulting EE improvement is up to 14% compared to a fixed baseline tilt angle.

From the results shown in Fig. 2(a) and (b) that can maximize throughput fairness is:

- **Homogeneous micro-BS Network**: fixing the antenna down-tilt in a narrow range of 1-5 degrees can guarantee a sufficiently high level of fairness, irrespective of hotspot location.

- **Heterogeneous femto-AP and micro-BS Network**: feedback knowledge of the hotspots is required in order to pre-determine the antenna down-tilt, which can vary in a wide range between 2-16 degrees.
The results in Fig. 2(a) and (b) demonstrate that even with only 1-2 hotspot femto-APs per BS, the effect can be dramatic and the strategy requirement has changed from a fixed down-tilt to a dynamic down-tilt depending on the hotspot location. Note that the presented results are for a symmetrical hotspot distribution (same \( \lambda \) for all BSs) in the multi-BS network. If the hotspot location varies for each BS, results could be different.

The BS antenna down-tilt in a homogeneous network is relatively insensitive to this, but the heterogeneous network is sensitive. Therefore, the fixed strategy for a heterogeneous network such as the one considered, would need network-wide knowledge of where the hotspots and APs are. Clearly this is quite infeasible, and a fixed strategy is sub-optimal in HetNets.

B. Hotspot Dispersion

Whilst the APs can not move, UE hotspots can migrate or disperse (e.g., shopping centers and train stations). The system model can be adjusted to account for a variation in the number of hotspot UEs (\( M_{\text{hotspot}} \)). The results in Fig. 2(c) show how the network is deployed for a hotspot density of \( \frac{M_{\text{hotspot}}}{M_{\text{BS}}} = 0.5 \), but as the density varies from 0 to 1, the throughput fairness index can change either positively or negatively depending on the network deployment. For a fixed strategy designed for a certain UE hotspot intensity:

- **Homogeneous micro-BS Network**: an unexpected increase in hotspot intensity means that the throughput fairness is significantly increased, because all UEs experience the same throughput (common location).
- **Heterogeneous femto-AP and micro-BS Network**: an unexpected increase in hotspot intensity means that the throughput fairness is marginally decreased, because all UEs experience severe interference from APs, and the BS antenna tilt is incorrectly adjusted.

Given that a fixed strategy can not account for variations in user density distribution, especially when femto-APs are deployed, a more dynamic algorithm should allow the BSs to learn what down-tilt strategy to adopt at any particular stage. This will be considered in Section IV.

C. Capacity Saturation

For realistic transmission schemes, there is a capacity saturation level, which corresponds to a saturation in the mutual information [16]. For LTE with 64-QAM and 6/7 turbo-codes, this is approximately 4.3 bit/s/Hz for SISO and 8.6 bit/s/Hz for 2x2 MIMO with spatial multiplexing. It has been proposed that energy-efficient BSs can adjust its transmission scheme based on the traffic load.

The paper now considers the impact of a changing capacity saturation level on the optimal antenna down-tilt angle. The results presented in Fig. 3 show that as the saturation level increases, the optimal down-tilt angle that maximizes throughput fairness decreases. This is because the higher-order transmission schemes increase the throughput disparity between the high SINR and low SINR users. In order to address this disparity, as the capacity saturation level increases, the antenna down-tilt needs to be decreased in order to compensate for the low SINR users further from the cell-centre.

In the next section, the paper investigates how a reinforcement learning algorithm can take into account the dynamic changes in user distribution, AP location and capacity saturation levels.

IV. REINFORCEMENT LEARNING

A. Distributed Solution

In complex HetNets, centralized planning and optimization requires near-perfect knowledge of all relevant parameters across the network in order to make an informed decision. This centralized process can be inefficient due to: long delays in feedback process, high complexity of joint-optimal algorithms, lack of feedback from low-complexity nodes.

An alternative method of optimization employs unsupervised machine-learning, whereby the agents (cells) within the network can learn through a combination of choosing actions and observing consequences. A rule for what actions to take is called a strategy. Existing research has demonstrated that machine learning can assist the dynamic optimization of complex issues [17].

B. Definitions

For the reinforcement learning algorithm, we define the following:

- **Agents**: the BSs, where the observed BS is \( n \in N_{\text{BS}} \).
- **Environment**: the network containing all BSs, APs and UEs subject to a certain offered traffic rate that has a uniform temporal and spatial distribution.
- **Action**: each agent can take an action to change its antenna down-tilt angle, \( \theta_{n,t} \in \Theta \).
- **State**: the current traffic load \( (L_{n,t}, \theta_{n-1,t} \in [0,1]) \) experienced by the agent concerned, after taking an action in the previous time frame.
- **Feedback**: by taking an action that leads to a state, the BS has an interference impact on the network containing other BSs.
Generally, a large small where the parameter \( \Delta \) from antenna tilt adjustment in the next section. These notions will be shown in terms of the results obtained for a known hotspot of UEs located at \( \lambda \)

- **Observed Reward**: defined by \( R_{\pi} \), which is the fairness ratio between cell-edge and cell-mean capacity, as given by (3).
- **Strategy**: a strategy \( (\pi_n \in \Pi) \) is individual to the BS. It dictates that for a given state \( L \), what action \( \theta \) to take in order to maximize the reward.

C. Strategy Exploitation vs. Exploration

In the absence of a full state transition map (i.e., Markov-Decision-Process (MDP)), there is often a tradeoff between the rate of exploitation and the rate of exploration [18] [19]. As shown in Fig. 4, over-exploration of strategies can lead to greater understanding of the MDP, but not being able to sufficiently exploit the most beneficial strategy. Under exploration will lead to rapid convergence to a single strategy, which is likely a sub-optimal strategy, and the network is not able to discover new and better strategies, as shown in Fig. 4.

A variety of exploration methodologies exist, and a popular one is the Boltzmann Exploration [18], where the chance of randomly exploring a strategy is weighted by the expectation of reward of all strategies. Under the Boltzmann Exploration algorithm [18], the strategy selection is based on a weighted probability biased in favor of likelihood to yield high rewards:

\[
p_{\pi} = \frac{e^{\frac{R_{\pi}}{\Delta}}}{\sum_{\pi' \in \Pi} e^{\frac{R_{\pi'}}{\Delta}}},
\]

where the parameter \( \Delta \) adjusts the level of exploration: a small \( \Delta \) favors exploitation and a large \( \Delta \) favors exploration. Generally, a large \( \Delta \) can guarantee asymptotic optimality [19]. However, there is a hidden variance tradeoff, as shown in Fig. 4 higher rewards generally mean higher reward variance. These notions will be shown in terms of the results obtained from antenna tilt adjustment in the next section.

V. ADAPTIVE STRATEGY RESULTS

In this section, the investigation considers BSs configured for a known hotspot of UEs located at \( \lambda = 250m \) with a hotspot intensity of \( \frac{M_{\text{ hotspot}}}{M_{\text{UE}}} = 0.5 \). Over time, this intensity grows to 0.8, as more UEs gather there. The BS has no knowledge of this and has to learn through exploration of strategies. The results are presented in Fig. 5 for a homogeneous network and in Table II for both the homogeneous and heterogeneous networks.

Figure 5 shows the variation in throughput fairness (\( F \)) with different number of learning iterations. For a fixed strategy, which was pre-set for a hotspot intensity of 0.5, it can achieve approximately \( F = 0.56 \). The resulting EE improvement is up to 35\% compared to a fixed baseline tilt angle of 2\°.

Now that the hotspot intensity has changed to 0.8, an exploitative algorithm \( (\Delta = 0.1) \), can achieve rapid solution convergence (6 iterations) and a high asymptotic fairness of \( F = 0.8 \) with a small variance of \( \sigma^2 = 0.015 \). The resulting EE improvement is up to 61\%.

For a more explorative algorithm \( (\Delta = 0.5) \), the solution achieves a higher asymptotic fairness of \( F = 0.86 \), but with a greater variance \( (\sigma^2 = 0.026) \) and slower convergence of 15 iterations. In this case, the results show that an increase in throughput fairness of 56\% is achieved. The resulting EE improvement is up to 74\%. When there is over exploration \( (\Delta = 1.0) \), the performance decreases because the BS is always looking for alternative tilt strategies, even when the optimal has already been found. The over-explorative results show that the fairness achieved is \( F = 0.72 \), with a large variance \( (\sigma^2 = 0.061) \).

The results in Fig. 5 show that in order to achieve the highest asymptotic reward (fairness), an optimal exploration rate exists. Typically this requires a degree of exploration and results in a medium variance, as shown previously in Fig. 4. Under-exploration leads to rapid convergence and low reward variance. Over-exploration leads to poor convergence and high reward variance.

It is worth noting that the optimal solution for the learning and fixed strategies should be the same, but the learning algorithm allows self-optimization without network knowledge and with a relatively rapid convergence speed. However, it incurs a
TABLE II
PERFORMANCE METRICS OF DIFFERENT STRATEGIES: VARIANCE, ENERGY-EFFICIENCY AND ERROR.

<table>
<thead>
<tr>
<th>Method</th>
<th>Metric</th>
<th>Iterations</th>
<th>Var.</th>
<th>Eff.</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homogeneous Network</td>
<td>Fixed</td>
<td>0</td>
<td>0</td>
<td>≤35%</td>
<td>≤62%</td>
</tr>
<tr>
<td>Learning (Exploit)</td>
<td>6-7</td>
<td>0.015</td>
<td>≤61%</td>
<td>≤21%</td>
<td></td>
</tr>
<tr>
<td>Learning (Explore)</td>
<td>12-15</td>
<td>0.026</td>
<td>≤74%</td>
<td>≤14%</td>
<td></td>
</tr>
<tr>
<td>Heterogeneous Network</td>
<td>Fixed</td>
<td>0</td>
<td>0</td>
<td>≤14%</td>
<td>≤94%</td>
</tr>
<tr>
<td>Learning (Exploit)</td>
<td>12-14</td>
<td>0.034</td>
<td>≤19%</td>
<td>≤28%</td>
<td></td>
</tr>
<tr>
<td>Learning (Explore)</td>
<td>19-21</td>
<td>0.071</td>
<td>≤24%</td>
<td>≤17%</td>
<td></td>
</tr>
</tbody>
</table>

Variance in the asymptotic performance and requires a suitable learning rate ($\Delta$) to balance the tradeoff between asymptotic performance and convergence speed. Table II summarizes the merits and disadvantages of various tilt adjustment strategies in homogeneous and heterogeneous networks. The reinforcement learning process can be conducted on a simulator prior to deployment (offline) or on a real-time network (online).

VI. CONCLUSIONS

This paper has considered the challenge of how to improve the throughput fairness and associated energy efficiency of a multi-cell network by adjusting the antenna down-tilt.

It has been found that a fixed tilt strategy can yield a good performance for either a homogeneous network with a dynamic traffic intensity pattern or a heterogeneous network with a static traffic intensity pattern. Furthermore, the results also show that the optimal down-tilt angle decreases with increased capacity saturation level of the transmission scheme. In order to account for dynamic traffic patterns and transmission schemes in a heterogeneous network, the system requires network-wide knowledge of where the hotspots are and what transmission scheme is being utilized.

The proposed dynamic tilt strategy employing a reinforcement learning algorithm does not need BS or network-wide knowledge of hotspot locations. The Boltzmann exploration algorithm converges to a near-optimal solution rapidly (6-15 iterations), for an appropriate exploration factor. The results show that it can increase the throughput fairness by 45-56% and the transmit energy efficiency by 21-47% compared to the fixed strategy. The tradeoff between asymptotic performance, variance and convergence rate is also discussed. For a given performance variance and convergence rate tolerance, there is an associated exploration rate and asymptotic performance.

REFERENCES