Explainable Decentralized Federated Learning for Energy-Efficient Base Station Sleep Control

Narges Movahedkor*, Reza Shahbazian†, Seyed Ali Ghorashi‡
*Department of Computer Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran
†Department of Informatics, Modeling, Electronics and System Engineering (DIMES), University of Calabria, Rende, Italy
‡School of Architecture, Computing and Engineering, University of East London, London, UK.
S.A.Ghorashi@uel.ac.uk

***Abstract*—Given the capacity and performance boosts offered by the 5G cellular networks, energy consumption at the base stations (BSs) has increased tremendously. This paper proposes a decentralized federated learning (DFL)-based intelligent BS switching, integrated with explainable artificial intelligence (XAI) methods, to mitigate the concerns for energy consumption in dense 5G networks. This entails collaboration among distributed but interconnected networks to learn the best policies for BS switching without any central controller, so that knowledge sharing can be ensured while privacy and communication efficiency are maintained. Very importantly, we further researched the XAI techniques to provide better transparency on the decision-making of the switching control agent and create some trust in the learned policies. Such explainability allows us to derive the most important factors affecting BS switching decisions and how these contribute in enabling energy savings while maintaining quality of service (QoS). Extensive simulations conducted to validate our proposed framework in presenting valuable XAI analysis have elaborately provided the basis for understanding the learned strategies and key factors driving energy-efficient BS management.**

***Index Terms*—Base Station Switching, Explainable AI, Decentralized Federated Learning, Deep Reinforcement Learning, Energy Efficient, 5G**

I. Introduction

The growing 5G and beyond necessitate dense small base station (SBS) networks for coverage and capacity, but raises concerns about energy consumption [1], [2]. The large energy footprint of these networks hinders infrastructure sustainability. To mitigate, energy conservation through station switching is suggested in several past studies [3], [4], however, it may negatively impact service quality and user association with inefficient base stations.

This paper proposes a novel explainable base station switching scheme using decentralized federated learning and deep reinforcement learning, along with a new deep learning model for traffic prediction, expanding our previous work [5]. The scheme aims to maximize system performance by optimizing energy efficiency and ensuring acceptable quality of service (QoS) [6], thereby reducing energy consumption and improving total system cost compared to existing methods.

Crucially, we recognize the inherent complexity of deep reinforcement learning models, which necessitates explainability and interpretability in their critical applications, such as 5G networks [7]. Specifically, we address the 'black box' problem inherent in many DRL solutions, illuminating the reasoning behind the agent's actions. Explainability allows operators to understand the decision-making process behind enabling or disabling certain activities, fostering trust and transparency. This is particularly important in dynamic 5G environments where rapid adjustments are required. Thus, this study primarily aims to integrate explainability into the proposed BS switching architecture to build confidence and guide wise judgments in energy-efficient 5G network operations, thereby enhancing the understanding of the agent's decision-making process. We achieve this by implementing a novel method for visualizing and quantifying the influence of different input features on the agent's actions, allowing for a detailed audit trail of its decision-making.

II. Related Works

BS on/off switching is a promising technique to reduce energy consumption in 5G networks. However, optimal strategies are computationally challenging [8]. Traditional rule-based approaches [9], [10] have evolved to heuristic and greedy methods [11]–[13], and more recently, reinforcement learning (RL) [14]–[18] and deep reinforcement learning (DRL) [19], [20] are used to handle dynamic environments and high-dimensional spaces, incorporating traffic demand and forecasting [8], [21], [22]. Essentially, a key challenge for learning-based decision systems is ensuring trustworthy and dependable decisions. In this respect, interpretability improves confidence, accountability, and transparency, allowing users to understand neural networks, identify weaknesses, and mitigate biases. Thus, there has been a growing interest in interpretability methods across AI disciplines like computer vision [23], counterfactual reasoning [24], and decision support systems [7] underscores the need for rationale in complex systems like 5G network management. Existing centralized control systems for BS switching often overlook privacy, QoS impacts, and single points of failure. This work proposes a DRL-based distributed federated learning (DFL) framework with private experience sharing, energy efficiency optimization, and QoS degradation minimization. We also incorporate explainability to provide insight into the DRL agent's decisions, ensuring transparent and reliable BS switching.

III. DFL Traffic-Aware Base Station Switching Management

Cellular networks manage fluctuating mobile traffic requirements, but intermittent traffic can cause some BSs to function under capacity, leading to energy inefficiency and improper usage. The BSs must thereby possess dynamic operating modes to maximize energy consumption [8]. This paper proposes a dynamic switching management framework for BSs to address mobile traffic needs by altering BS operation modes based on live traffic loads. We also use a coarse-grained time discretization to reduce overhead of regular mode changes within thirty-minute intervals. We also recognize that the scheme must be transparent, providing insight into mode changes for performance improvement and debugging in real-world deployments.

A. Problem Formulation

Aiming to minimize long-term system costs, we modeled the problem as a Markov Decision Process (MDP) $J =< S, A, T, C, \gamma >$, with states S, actions A, transition probabilities T, cost function C, and a discount factor $\gamma \in [0, 1]$. We divide a 5G network into non-overlapping grid regions, each served by a BS $BS_i \in B$. At the time slot t, the BS state s_t includes the predicted traffic volume $(\widetilde{P}_t, a_{t-1})$ and the previous mode (active/sleep). The action space A comprises active or sleep modes, and the state transition probability is defined by $T(s_{t+1}|s_t, a_t)$. We define the overall system cost at time slot t is by the following key components:

a) Energy Consumption: BS energy usage consists of fixed (due to hardware, cooling) and load-dependent (power amplification) components. We also consider a fixed energy consumption $E_{i,t} = E_i^{fixed\_sleep}$ for when BS_i is in sleep mode. When active, the consumed energy is defined as $E_{i,t} = E_i^{fixed\_active} + \omega \times \rho_{i,t}$ where $E_{i,t}$ is the energy consumed at time t, E^{fixed_sleep} and E^{fixed_active} are fixed energy consumptions, $\rho_{i,t}$ is the traffic load, and ω is the power amplification factor.

b) QoS Degradation : Deactivating BSs negatively impacts user QoS, increases delays and the load on remaining BSs and wastes potential capacity, which could be used to cover traffic demands. We define this degradation as $Q_{i,t} = \lambda * (trans\_cost + \frac{1}{(C_{i,t} - \rho_{i,t})})$ where $trans\_cost$ is the transmission cost, and λ is a load-dependent penalty factor defined by $(w_i \times \rho_{i,t})$.

c) BS Switching Cost: Assuming that the transition cost from active to sleep mode is negligible, we define the switching cost of a BS from sleep to active mode as SW_i.

Therefore, we define the overall cost for BS_i at time t as follows, considering the energy efficiency of the individual BS, while ensuring the overall QoS and Switching cost: $C_{i,t} = \omega_{e_i} \times E_{i,t} + \sum_{i \in B} \omega_{Q_i} \times Q_{i,t} + \sum_{i \in B} \omega_{SW_i} \times SW_i$ where ω_E, ω_Q, and ω_{SW} are weights for energy, QoS, and switching costs, respectively.

IV. Proposed Framework

The architecture of our proposed framework is depicted in Fig. 1. Each BS in this configuration carries a traffic prediction model and a DRL decision maker for a decentralized BS switching management module, both of which share weights with other BSs in a DFL manner [5]. We discuss each component in greater detail in the following sections.

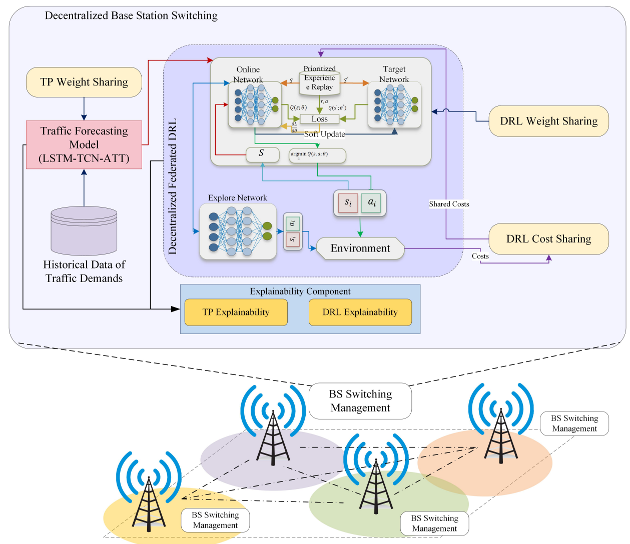


Fig. 1: Overall architecture of the proposed framework

A. Attention-based Traffic Forecasting

The proposed traffic forecasting model employs BiLSTM (Bidirectional Long Short-Term Memory), TCN (Temporal Convolutional Network), and self-attention mechanisms within a DFL framework. Two stacked BiLSTM layers, enhanced with batch normalization and ReLU activation, capture sequential traffic patterns while mitigating vanishing gradients [8], [25]. A TCN layer and self-attention mechanism extract long-term dependencies and relevant features, respectively. The model is trained collaboratively across BSs using DFL, ensuring both enhanced prediction accuracy and data privacy.

B. Dynamic BS Switching

We propose a novel DRL framework for optimal BS switching, combining traffic predictions with a PDDQN (Prioritized Double Deep Q-Network) architecture. To promote collaboration, a decentralized weight and cost sharing mechanism is implemented, which enables collaboration as well as adaptability to varied traffic patterns. Exploration is enhanced via an explore network, inspired by [8], sharing the network structure but with perturbed weights ($\widetilde{W} = W + \Delta W$, where $\Delta W = \alpha \cdot random(-1, 1) \cdot W$). Actions from both networks are evaluated, and if the explore network's result is superior, its weights $\widetilde{W}$ are partially incorporated into the actor network through $W' = W + \sigma \widetilde{W}$, with a decreasing factor σ to facilitate knowledge transfer.

C. Explainability

Here we talk about explainability for the proposed traffic forecasting and DRL-based BS-switching units, aiming for transparency and trustworthiness. We use attention mechanisms and action maps to scrutinize the model decision process, enhancing the model understanding and supporting the practical deployment.

1) Attention Mechanism: We use attention mechanism in our traffic prediction architecture to emphasize influential traffic patterns. Attention weights are derived from two average and max pooling layers applied after the TCN, quantifying the importance of different temporal contexts, namely the overall traffic trends, and peak values and sudden fluctuations, respectively. As the TCN layer's dilation of $[1, 2]$ yields two responses, we can assume capturing short-term and long-term traffic patterns dependencies. Therefore, for each time step, we have a 2×2 matrix of attention weights: $\begin{bmatrix} W_{avg,short} & W_{max,short} \\ W_{avg,long} & W_{avg,long} \end{bmatrix}$ where $W_{avg,short}$ and $W_{max,short}$ represent the average and max pooling output relating to the short range temporal data, while$W_{avg,long}$ and $W_{max,long}$ refer to the long range temporal data.

2) Action Mapping: We use action mapping to explain the deep Q-network's (DQN) decision-making, derived from the Q-value differentials. Given a network state s and predicted traffic patterns, the DQN outputs Q-values $Q(s, a)$ for each possible action. As our objective is to minimize total cost, the DQN selects the action that minimizes the expected cost, i.e., $a = \arg\min_a Q(s, a)$. To quantify the relative preference for each action, we compute the Q-value difference, ΔQ, between the activate and sleep actions $\Delta Q = |Q(s, activate)| - |Q(s, sleep)|$. The difference in Q values indicates the DQN's preference for an action, with negative ΔQ indicating sleep preference and positive Q meaning activation preference, also reflecting the confidence of the model.

D. Training

The model training process utilizes time steps and episodes, updating DDQN and traffic forecasting components periodically. This includes 12-24 hour updates for actor and target networks, and 48-hour time windows for traffic forecasting training and weight sharing among BSs.

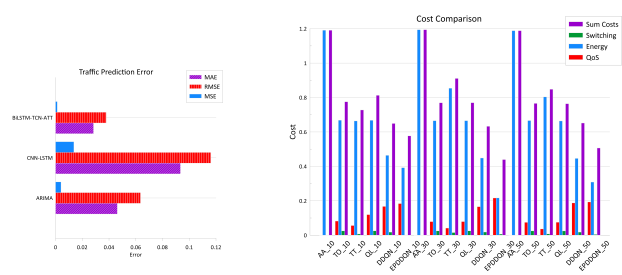
V. Experiments

In the following evaluations, we primarily focus on interpretability analysis rather than re-evaluating performance metrics, referring to previous research for detailed experiment settings, cost, and accuracy [5]. We use a real-world mobile traffic dataset, collected by Telecom Italia in Milan, Italy, in November 2013 [26][1], from 22:00, October 31, 2013, to 22:50, December 19, 2013, over 10 minutes. Experiments are conducted on a 64-bit computer with 24 processing cores and 16 GB RAM using Python and PyTorch for implementation, and splitting data by 0.6 for training and validation. The cost values in this table are obtained from [8] and [27]. We have made the software code for evaluations **publicly available** at GitHub.

[1]Publically available on: https://ieee-dataport.org/documents/telecom-italia-and-opnet-datasets-network-traffic-prediction

A. Performance Evaluation

This section briefly discusses the performance evaluation of the proposed approach, regarding the traffic prediction accuracy and its effectiveness in reducing the total cost. Figure 2 depicts the performance evaluation results regarding the traffic prediction and cost evaluations. Against ARIMA and CNN-LSTM benchmarks, we assessed our suggested BiLSTM-TCN-ATT traffic prediction model. Trained with Adam optimizer and MSE loss, the proposed model surpassed the baselines with about 50% reduction in MSE (Mean Squared Error), RMSE (Root Mean Squared Error), and MAE (Mean Absolute Error), therefore showing its excellent precision in traffic volume and extreme traffic conditions prediction (Figure 2a). We assess the performance of our BS switching control approach to various baseline approaches, including traffic-oblivious switching (TO) [3], dynamic two-threshold (TT) [28], [29], Q-learning, DDQN [27], [30], [31], and a scenario with always active BSs. The findings indicate that our approach surpasses the baseline methods, achieving noticeable more energy savings, while maintaining lower switching and total costs. This is particularly when there are numerous BSs, which indicates its scalability.



(a) (b)

Fig. 2: Performance evaluation comparisons.

B. Traffic Prediction Insights

Figure 3 compares attention scores across BSs for different traffic aspects, revealing distinct patterns for each BS. This indicates capturing localized traffic characteristics, emphasizing varying factors for accurate predictions. For instance, BS3 shifts focus from overall trends to peak loads around time step 1120 (Figures 3b and 3a), while BS1 and BS2 show stable and fluctuating attention patterns, respectively, suggesting different predictabilities. Analyzing visualizations also reveals how BSs respond to traffic patterns. BS5 is highly sensitive to recent traffic, shown by fluctuations in short-term attention, while maintaining stable long-term focus (3a and 3c). BS1 displays consistently low scores for average traffic trends (3a and 3b), while BS6 shows insensitivity to overall short-term trends (3a) but focuses on recent peak events, with a temporary shift between time steps 1100 and 1140. Figure 4 depicts attention patterns for individual BSs and the overall average trend, highlighting significant variability and dynamic focus adjustments. BS4 and BS9 rely on long-term trends, with BS9 temporarily shifting to peak values between time steps 1110 and 1150. BS8 shows high attention variability and sensitivity to peak values. By contrast, the overall analysis shows a balanced consideration of all attention types, with a network-wide preference for long-term peak values and recent trends around time step 1130. Visual representations of attention scores can reveal model behavior, highlighting dynamic patterns for each BS and network, system-wide shifts, justifying predictions, and building trust due to adaptation to local traffic.

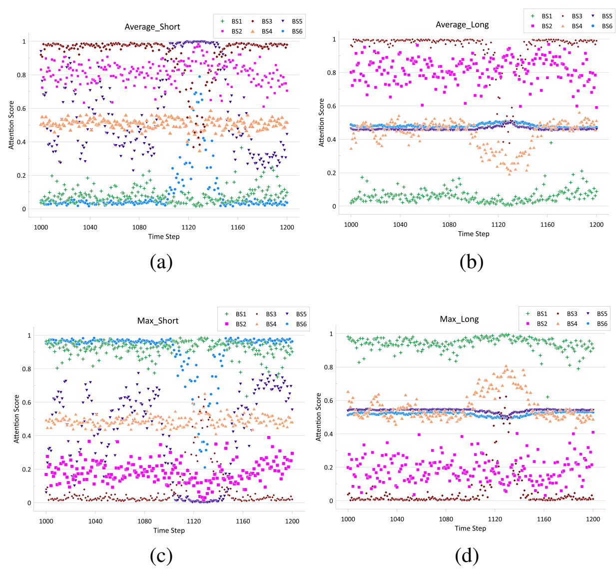


Fig. 3: Attention score comparison of different BSs, regarding traffic patterns and time-related factors

C. BS Switching Interpretation

We utilize action maps, showing Q-values, model confidence (ΔQ), and selected actions, to understand the DQN's decision-making in BS0 and BS3. These maps reveal how the DQN balances energy efficiency and performance based on local traffic conditions, selecting actions with lower Q-values to minimize total cost.

Figure 5 compares action maps for two sample BSs, showing a correlation between traffic and DQN decisions, with increased activity during higher loads. BS0 displays a dynamic map with generally high confidence periods. The varying confidence levels imply potential dependence on particular traffic types. By contrast, BS3 exhibits a single confidence spike followed by lower levels, indicating varying certainty. The actual network mode aligns with DQN actions, validating control. Also, BS3 displays fewer active periods and shows sleep modes covered by BS0 active modes, highlighting collaborative DFL.

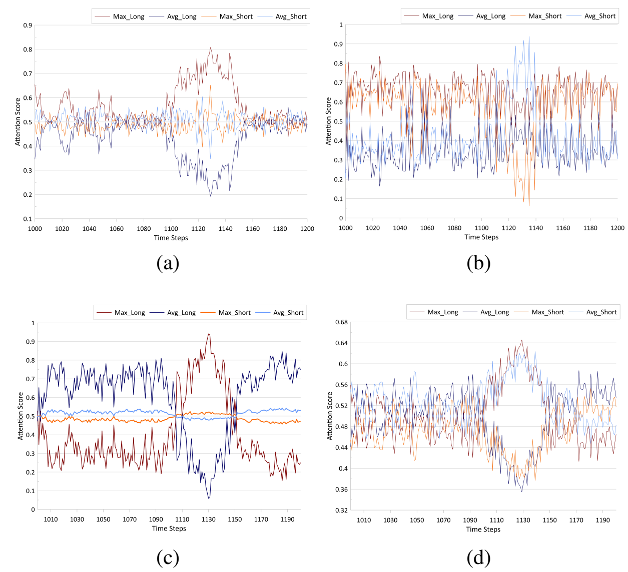


Fig. 4: Comparing the attention scores for individual BSs

Therefore, action maps aid in understanding localized behaviors and identifying anomalies in switching management. For example, both BSs are active from time steps 6460 to 6500 despite decreased traffic, suggesting potential energy savings through sleep mode. Confidence level fluctuations show varying certainty, and network activation reveals sensitivity to circumstances. Anomalies in confidence, actions, or traffic patterns can trigger real-time monitoring and adaptive control for network stability.

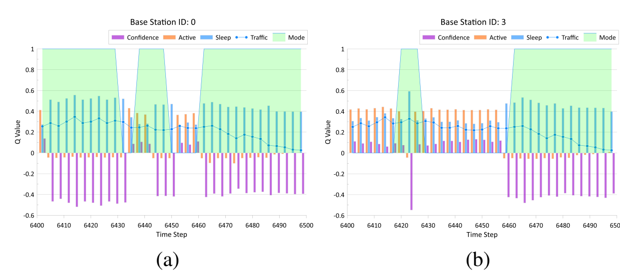


Fig. 5: Evaluating action maps of two sample BSs, alongside the traffic and model confidence

VI. Conclusion and Future Works

This paper expands on our previous framework's (EPDDQN) performance improvements in energy savings and cost reduction by over 20%. It explores the explanation of its decision-making through visualizations of attention scores and action maps. Through evaluations, we showed how the model dynamically changes network states and localizes traffic patterns, emphasizing temporal dependencies. This understanding improves openness and confidence, highlighting the model's interpretability, and making it a promising solution for balancing service quality and energy savings. Future studies aim to enhance the explainability of the EPDDQN approach by addressing complex network topologies and traffic patterns. We also intend to examine how these techniques affect the interpretability of learned representations and the effect of federated learning approaches on convergence speed and model performance.

Acknowledgment

We acknowledge financial support from PNRR MUR project FAIR (PE0000013).

Declaration of AI Assistance

We acknowledge the use of AI-based tools for language refinement, grammar correction, and formatting adjustments. The final version has been reviewed and approved by the authors.

References

[1] Hauke Holtkamp, Gunther Auer, Samer Bazzi, and Harald Haas. Minimizing base station power consumption. *IEEE Journal on Selected Areas in Communications*, 32(2):297–306, 2013.

[2] Adrian Ichimescu, Nirvana Popescu, Eduard C Popovici, and Antonela Toma. Energy efficiency for 5g and beyond 5g: Potential, limitations, and future directions. *Sensors (Basel, Switzerland)*, 24(22):7402, 2024.

[3] Chang Liu, Balasubramaniam Natarajan, and Hongxing Xia. Small cell base station sleep strategies for energy efficiency. *IEEE Transactions on Vehicular Technology*, 65(3):1652–1661, 2015.

[4] Mingjie Feng, Shiwen Mao, and Tao Jiang. Base station on-off switching in 5g wireless networks: Approaches and challenges. *IEEE Wireless Communications*, 24(4):46–54, 2017.

[5] Narges Movahedkor and Reza Shahbazian. Decentralized federated deep reinforcement learning framework for energy-efficient base station switching control. In *2024 11th International Symposium on Telecommunications (IST)*, pages 455–460. IEEE, 2024.

[6] Salma Kazemi Rashed, Reza Shahbazian, and Seyed Ali Ghorashi. Learning-based resource allocation in d2d communications with qos and fairness considerations. *Transactions on Emerging Telecommunications Technologies*, 29(1):e3249, 2018.

[7] Sören Schleibaum, Lu Feng, Sarit Kraus, and Jörg P. Müller. Adesse: Advice explanations in complex repeated decision-making environments. In Kate Larson, editor, *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24*, pages 7904–7912. International Joint Conferences on Artificial Intelligence Organization, 8 2024. Human-Centred AI.

[8] Qiong Wu, Xu Chen, Zhi Zhou, Liang Chen, and Junshan Zhang. Deep reinforcement learning with spatio-temporal traffic forecasting for data-driven base station sleep control. *IEEE/ACM transactions on networking*, 29(2):935–948, 2021.

[9] Eunsung Oh and Bhaskar Krishnamachari. Energy savings through dynamic base station switching in cellular wireless access networks. In *2010 IEEE Global Telecommunications Conference GLOBECOM 2010*, pages 1–5. IEEE, 2010.

[10] Eunsung Oh, Kyuho Son, and Bhaskar Krishnamachari. Dynamic base station switching-on/off strategies for green cellular networks. *IEEE transactions on wireless communications*, 12(5):2126–2136, 2013.

[11] Abhinav Kumar and Catherine Rosenberg. Energy and throughput tradeoffs in cellular networks using base station switching. *IEEE Transactions on Mobile Computing*, 15(2):364–376, 2015.

[12] Madhu Sudan Dahal. Energy saving in 5g mobile communication through traffic driven cell zooming strategy. *Energy Nexus*, 5:100040, 2022.

[13] Yi Yang, Zhixin Liu, Heng Zhu, Xinping Guan, and Kit Yan Chan. Energy minimization by dynamic base station switching in heterogeneous cellular network. *Wireless Networks*, 29(2):669–684, 2023.

[14] Ali El Amine, Jean-Paul Chaiban, Hussein Al Haj Hassan, Paolo Dini, Loutfi Nuaymi, and Roger Achkar. Energy optimization with multi-sleeping control in 5g heterogeneous networks using reinforcement learning. *IEEE Transactions on Network and Service Management*, 19(4):4310–4322, 2022.

[15] Ali El-Amine, Mauricio Iturralde, Hussein Al Haj Hassan, and Loutfi Nuaymi. A distributed q-learning approach for adaptive sleep modes in 5g networks. In *2019 IEEE wireless communications and networking conference (WCNC)*, pages 1–6. IEEE, 2019.

[16] Daniela Renga, Zunera Umar, and Michela Meo. Trading off delay and energy saving through advanced sleep modes in 5g rans. *IEEE Transactions on Wireless Communications*, 22(11):7172–7184, 2023.

[17] Ali El Amine, Jean-Paul Chaiban, Hussein Al Haj Hassan, Paolo Dini, Loutfi Nuaymi, and Roger Achkar. Energy optimization with multi-sleeping control in 5g heterogeneous networks using reinforcement learning. *IEEE Transactions on Network and Service Management*, 19(4):4310–4322, 2022.

[18] Meysam Masoudi, Mohammad Galal Khafagy, Ebrahim Soroush, Daniele Giacomelli, Simone Morosi, and Cicek Cavdar. Reinforcement learning for traffic-adaptive sleep mode management in 5g networks. In *2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1–6. IEEE, 2020.

[19] Junhong Ye and Ying-Jun Angela Zhang. Drag: Deep reinforcement learning based base station activation in heterogeneous networks. *IEEE Transactions on Mobile Computing*, 19(9):2076–2087, 2019.

[20] Hyungyu Ju, Seungnyun Kim, Youngjoon Kim, and Byonghyo Shim. Energy-efficient ultra-dense network with deep reinforcement learning. *IEEE Transactions on Wireless Communications*, 21(8):6539–6552, 2022.

[21] Yuchao Zhu and Shaowei Wang. Joint traffic prediction and base station sleeping for energy saving in cellular networks. In *ICC 2021-IEEE International Conference on Communications*, pages 1–6. IEEE, 2021.

[22] Gunhee Jang, Namkyu Kim, Taeyun Ha, Cheol Lee, and Sungrae Cho. Base station switching and sleep mode optimization with lstm-based user prediction. *IEEE Access*, 8:222711–222723, 2020.

[23] Junqi Jiang, Francesco Leofante, Antonio Rago, and Francesca Toni. Robust counterfactual explanations in machine learning: A survey. In Kate Larson, editor, *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24*, pages 8086–8094. International Joint Conferences on Artificial Intelligence Organization, 8 2024. Survey Track.

[24] Matteo Bianchi, Antonio De Santis, Andrea Tocchetti, and Marco Brambilla. Interpretable network visualizations: A human-in-the-loop approach for post-hoc explainability of cnn-based image classification. In Kate Larson, editor, *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24*, pages 3715–3723. International Joint Conferences on Artificial Intelligence Organization, 8 2024. Main Track.

[25] Sepp Hochreiter, Yoshua Bengio, Paolo Frasconi, Jürgen Schmidhuber, et al. Gradient flow in recurrent nets: the difficulty of learning long-term dependencies, 2001.

[26] Gianni Barlacchi, Marco De Nadai, Roberto Larcher, Antonio Casella, Cristiana Chitic, Giovanni Torrisi, Fabrizio Antonelli, Alessandro Vespignani, Alex Pentland, and Bruno Lepri. A multi-source dataset of urban life in the city of milan and the province of trentino. *Scientific data*, 2(1):1–15, 2015.

[27] Hyebin Park and Seung Hyun Yoon. Deep reinforcement learning for base station switching scheme with federated lstm-based traffic predictions. *ETRI Journal*, 2024.

[28] Bingjie Leng, Xueying Guo, Xi Zheng, Bhaskar Krishnamachari, and Zhisheng Niu. A wait-and-see two-threshold optimal sleeping policy for a single server with bursty traffic. *IEEE Transactions on Green Communications and Networking*, 1(4):528–540, 2017.

[29] Ioannis Kamitsos, Lachlan Andrew, Hongseok Kim, and Mung Chiang. Optimal sleep patterns for serving delay-tolerant jobs. In *Proceedings of the 1st International Conference on Energy-Efficient Computing and Networking*, pages 31–40, 2010.

[30] Zahra Rezaei and Behrouz Shahgholi Ghahfarokhi. Energy and spectrum efficient cell switch-off with channel and power allocation in ultra-dense networks: A deep reinforcement learning approach. *Computer Networks*, 234:109912, 2023.

[31] Jingchu Liu, Bhaskar Krishnamachari, Sheng Zhou, and Zhisheng Niu. Deepnap: Data-driven base station sleeping operations through deep reinforcement learning. *IEEE Internet of Things Journal*, 5(6):4273–4282, 2018.