Applying deep reinforcement learning for self-organizing network architecture

*Note: Sub-titles are not captured in Xplore and should not be used

Yi-Hao Tu
Dept. of Electrical Engineering,
National Taiwan University of Science and Technology,
Taipei, Taiwan.
d11107005@mail.ntust.edu.tw

Yi-Wei Ma
Dept. of Electrical Engineering,
National Taiwan University of Science and Technology,
Taipei, Taiwan.
ywma@mail.ntust.edu.tw

Zhi-Xiang Li
Dept. of Electrical Engineering,
National Taiwan University of Science and Technology,
Taipei, Taiwan.
m11107511@mail.ntust.edu.tw

Jiann-Liang Chen
Dept. of Electrical Engineering,
National Taiwan University of Science and Technology,
Taipei, Taiwan.
Lchen@mail.ntust.edu.tw

Kazuya Tsukamoto
Dept. of Computer Science and Networks,
Kyushu Institute of Technology,
Fukuoka, Japan.
tsukamoto@csn.kyutech.ac.jp

***Abstract*—**Inefficient resource allocation and unstable connection quality for mobile devices are the primary challenges of Self-Organizing Networks (SON). Frequent handovers between base stations can result in a network burden imbalance. In contrast, unstable connection quality causes disconnection or signal interference between mobile devices and base stations, influencing network performance and reliability. In recent years, wireless communication technology has extensively utilized Reinforcement Learning (RL) to obtain the optimal strategy through continuous interaction between agents and their environments. Deep Reinforcement Learning (DRL) is based on Deep Neural Networks (DNN) architectures which allows it to handle increasingly complex network situations. This paper proposes a SON architecture based on DRL in response to the aforementioned challenges. We describe how the agent learns the optimal parameter settings through training based on various network scenarios to improve handover strategies and enhance overall network performance and resource utilization. Our proposed framework can also be applied to the present Fifth Generation (5G) network.

Keywords*—Deep reinforcement learning, handover optimization, mobility load balancing, mobility robustness optimization, self-organizing networks.*

I. Introduction

The Self-Organizing Networks (SON) conceptualism is proposed by Next-Generation Mobile Networks (NGMN). Advanced self-optimization strategy needs and accompanying SON application scenarios are discussed, and thus the 3rd Generation Partnership Project (3GPP) later embraced this model. To offer a solution for the automated setup and deployment of Long-Term Evolution (LTE), a collection of SON functions (SONF) was created. SON aims to assist mobile operators in improving network efficiency and performance by automating operations and operations to minimize network complexity and expense. SON offers self-configuration, -optimization, and -healing for decreasing human operation requirements [1-2].

The implementation of SON concentrates predominantly on the operation of the Radio Access Network (RAN). Despite the early interest, it has yet to become part of the end-to-end solution and fully meet mobile network operators' expectations. However, due to the exponential growth of mobile devices, current standards become increasingly inadequate to meet future demands. Fifth Generation (5G) seeks to fulfill the following functions: high transmission rates and low latency, control plane and data plane decoupling, heterogeneous network (HetNets) characteristics, network-densification and -virtualization, flexible spectrum allocation, and infrastructure sharing. Nowadays, the telecommunications infrastructure has to adapt to 5G. However, these factors will significantly increase network management requirements. To establish the authorization procedure for the future communication system, assuring its dependability, scalability, stability, and efficiency, the SONF remains an essential technology for mobile operators.

Mobility Load Balancing (MLB) and Mobility Robustness Optimization (MRO) are the critical characteristics of self-optimization [2-4]. MLB will divide the load across between imbalance two cells [2,5]. MRO is concerned with reducing the issue of handover. The MLB and MRO parameters have a close correlation and dependence on each other, despite operating independently. Conflicts may arise when adjusting handover parameters to optimize performance if they are adjusted in opposite directions. For example, MLB modifies the parameters of handover which balancing the burden among the neighboring cells. Meanwhile, MRO adjusts the handover settings to minimize the handover issue, and if this behavior loops, the network performance will suffer significantly [6].

This study focuses on how MLB and MRO operate, explains the reasons for the conflict and the corresponding challenges to be addressed, and outlines our main contributions, including a SON framework that supports Reinforcement Learning (RL) or Deep Reinforcement Learning (DRL). This proposed framework can be applied to 5G scenarios.

II. The conflict problem in SON

Before describing the operational concepts of the two functions, MRO and MLB [6]. This study identifies the causes of the conflicts of the handover triggering procedure, which introduces in this section.

Identify applicable funding agency here. If none, delete this text box.

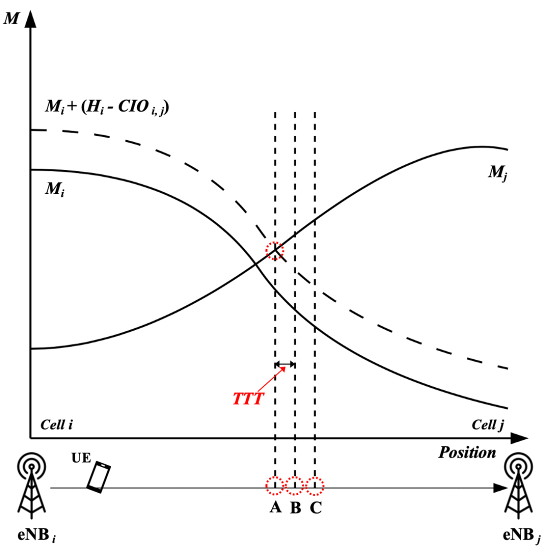

Fig. 1. In LTE, when the entering condition is met at point A, Event A3 can be triggered. At point B, UE sends a measurement report for the handover procedure trigger. The handover procedure is then completed at point C.

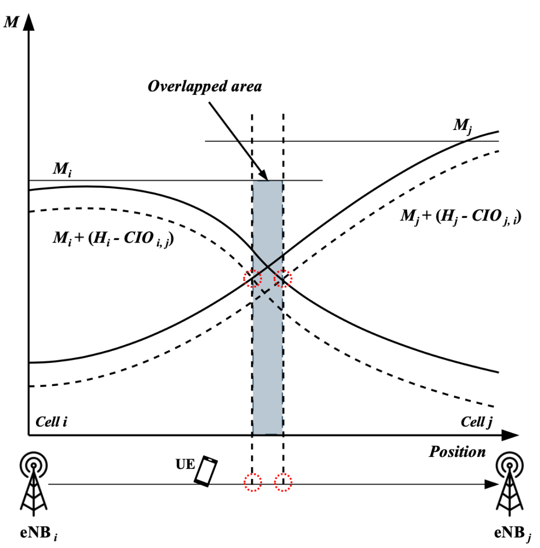

Fig. 2. If a user equipment (UE) moves through the overlapped area, it will trigger a ping-pong handover between *cell i* and *j*.

A. Handover trigger procedure by LTE

Figure 1 illustrates the LTE handover trigger procedure. The horizontal axis (x-axis) of the graph illustrates the spatial coordinates of the User Equipment (UE), while the vertical axis (y-axis) refers the magnitude of the strength of the signal received between the UE and the E-Urtan Node B (*eNB*). When UE moves from *cell i* to *cell j*, the intensity is taken into account two *cell* signals and see whether the signal strength matches the *cell i* A3 event [7] entrance condition, as shown in Eq. (1). When the entrance condition is satisfied and the Time To Trigger (TTT) has been triggered, the UE sends out a measurement report to *eNB i*. Afterward, the UE commences a handover procedure from *eNB i* to *eNB j*.

$$M_j > M_i + \left(H_i - CIO_{i,j}\right) \tag{1}$$

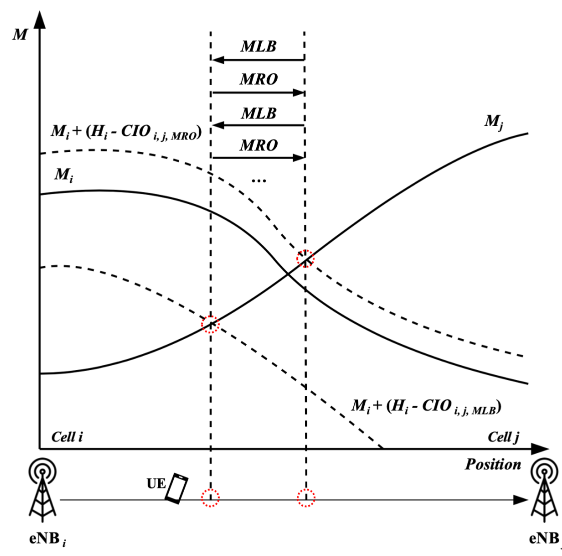

Fig. 3. MLB and MRO are being operated inappropriately, causing a conflict to arise.

where M_j indicates the *cell j* measurement results, M_i denotes the *cell i* measurement results, H_i represents the hysteresis parameter, and $CIO_{i,j}$ means *cell i* has set a specific offset for *cell j*.

B. MRO operational principle

The issues of the Radio Link Failure (RLF) and the ping-pong effect might arise from improper handover settings. Disconnection between the UE and the serving *eNB* due to the RLF event has a major detrimental effect on the user's session. The primary cause of RLF is attributed to the insufficient signal intensity of the eNB that is responsible for providing service to the UE, coupled with the presence of excessive interference [8]. A deficient handover trigger can lead to a RLF. Suppose the initiation of the handover trigger occurs prematurely, and the signal strength of the destination cell is insufficient. In such scenario, the RLF is probably going to take place not long after the handover process has been finished. Be aware that if the handover trigger is delayed, the signal strength of the source cell will already have decreased before the handoff takes place. Because of this, there is a greater possibility of encountering RLF either before or during the handover process [2]. Even while the Quality of Service (QoS) of the UE may not be directly impacted by the ping-pong effect, it does lead to an inefficient usage of network resources. As shown in Figure 2, H_j indicates the hysteresis parameter of *cell j*, $CIO_{j,i}$ which means *cell j* sets a specific cell offset for *cell i*. The incorrect setting of the handover parameter resulted in an overlapping area, and the A3 events of both cells met the entry conditions. When the UE passes through the area, a ping-pong effect occurs between the two cells [7].

Through the collection of messages from the UE, the MRO has the capacity to determine whether or not handover issues are happening inside the cell [2]. The RLF and ping-pong effects are to be mitigated to the greatest extent possible by optimizing the auto-handover parameters, namely H, CIO, and TTT. If there is a premature handover RLF, the MRO will postpone the process for the handover trigger. Conversely, if the MRO identifies a late handover RLF, they will initiate the handover trigger process earlier. Additionally, the MRO has a process for addressing ping-pong handover problems. To avoid the ping-pong handover, the overlapping areas in Figure 2 must be eliminated by adjusting the handover parameters using Eq. (2).

$$\left(H_i - CIO_{i,j}\right) + \left(H_j - CIO_{j,i}\right) > 0 \tag{2}$$

C. MLB operational principle

The primary objective of MLB is to address the issue of imbalanced traffic distribution among cellular network cells. Periodically monitor the load and make adjustments to the handover parameter CIO in instances where the load is found to be unbalanced [2]. Let's say that *cell i* is experiencing a high amount of usage, while neighbor *cell j* is relatively underutilized. In this scenario, if the MLB function of *cell i* is turned on, it will select *cell j* for load balancing. During the operation, *eNB i* increases $CIO_{i,j}$ to initiate handover sooner, Therefore, when the UE moves from *cell i* to *j*, the handover out of *cell i* will occur earlier, thereby reducing the load in *cell i*. Simultaneously, *eNB i* informs *eNB j* to turn $CIO_{j,i}$ down, This causes a delay in transferring from *cell j* to *i*, which also slows down the process of increasing the load in *cell i*.

D. MLB and MRO are in conflict

The initial method involved independent operations of MLB and MRO, but they are closely interconnected. Both functions optimize network performance by adjusting the handover parameter. As a result, there is a potential for a dispute to arise if they alter the same parameter in the opposite direction. As shown in Figure 3, to achieve load balancing in *cell i*, MLB has added $CIO_{i,j}$ to advance UE out of *cell i*. Unluckily, the entry condition value of the *cell i's* A3 event changed too low, and a premature handover of RLF occurred after the MLB operation. The MRO has detected an issue with RLF and has made adjustments to the $CIO_{i,j}$ to delay the handover trigger procedure. Eventually, however, MLB changed it again. These repeated behaviors cause loops, which are called "ping-pong". On the other hand, improper operation by MLB, which delays a handover trigger, can also result in a ping-pong handover or a late RLF handover, which leads to conflict.

The consecutive conflicts between MRO and MLB have resulted in performance degradation for both entities, presenting an urgent problem that necessitates resolution. On the one hand, the handover issue presents a conflict, which has implications for user experience and the inefficient utilization of network resources. However, as a result of the reciprocal interaction between the two functions, the efficiency of the MLB operation is compromised. This phenomenon may give rise to a prolonged state of overload, thereby exacerbating call-blocking and call-dropping occurrences.

III. Machine Learning Mechanism for SON

Machine Learning (ML) is a branch of Artificial Intelligence (AI) that focuses on teaching computers how to interpret information effectively. The four main ML types are Supervised, Unsupervised, Semi-Supervised, and Reinforcement Learning [8].

In Supervised Learning (SL), the algorithm undergoes the training process using input data that is labeled with a specific output. Consequently, its objective is to generate a function that establishes a mapping between the input and the desired outcome. In contrast, a Unsupervised Learning (UL) algorithm only has access to unlabeled data. The goal of UL is to identify patterns in input data in order to forecast future inputs.

The Semi-Supervised Learning (SSL) framework integrates both SL and UL techniques, using data with and without labels throughout the training phase to improve learning result accuracy. In the context of SSL, the training of a predictive model involves utilizing a lesser proportion of labeled data compared to unlabeled data.

In the field of Reinforcement Learning (RL), an autonomous agent acquires knowledge through iterative interactions with its environment, wherein it is provided with feedback in the form of positive or negative rewards. The primary objective is to select a course of action that maximizes the potential for future rewards.

IV. Intelligent Conflict Avoidance for SON

As mentioned earlier, ML has gained popularity because it can analyze dependencies and relationships such as MLB and MRO. It can be successfully applied to avoid/resolve conflicts between these two functions.

The issue of avoiding or resolving conflicts has been extensively discussed in various recent works. Rojas *et al.* [9] propose a hybrid coordination method based on ML and apply it to the coordination between MRO and MLB. The adaptability of the framework was proved by its ability to handle challenging networking situations such as Ultra-Reliable and Low-Latency Communications (URLLC). On the other hand, Rojas *et al.* [10], the authors explore an ML-driven framework for autonomous system model construction, one that considers the selected SONF's dynamics and streamlines the optimization process for resolving any MRO/MLB conflicts. They do this by automatically deriving the system model from the data collected by the framework.

Stamatelatos *et al.* [11] provide a more detailed analysis of the contributions made by previous studies in the field of SON coordination schemes that integrate ML capabilities in order to meet the requirements of 5G SON. Shubyn *et al.* [12] describe an innovative strategy for managing switching in heterogeneous 5G mobile networks. This strategy makes use of AI technology, in particular Recurrent Neural Networks (RNNs), which use both Gated Recurrent Units (GRU) and Long Short-Term Memory (LSTM) units. According to the results, using architectures based on GRUs is a realistic option because of their better performance in rapidly adjusting to environmental changes. Moysen *et al.* [13], a suggested scheme, solve possible conflicts between two critical functions within SON regarding the mobility of MLB and MRO. These potential conflicts may be caused by the fact that both functions need MLB and MRO to move about. The method described in [13] generates a prediction model by using previous UE measurements and regression analysis methods to make predictions on the network's performance. Based on the simulations, It has been proved that the suggested method can resolve conflicts by correctly forecasting network performance via the comprehensive analysis of UE data.

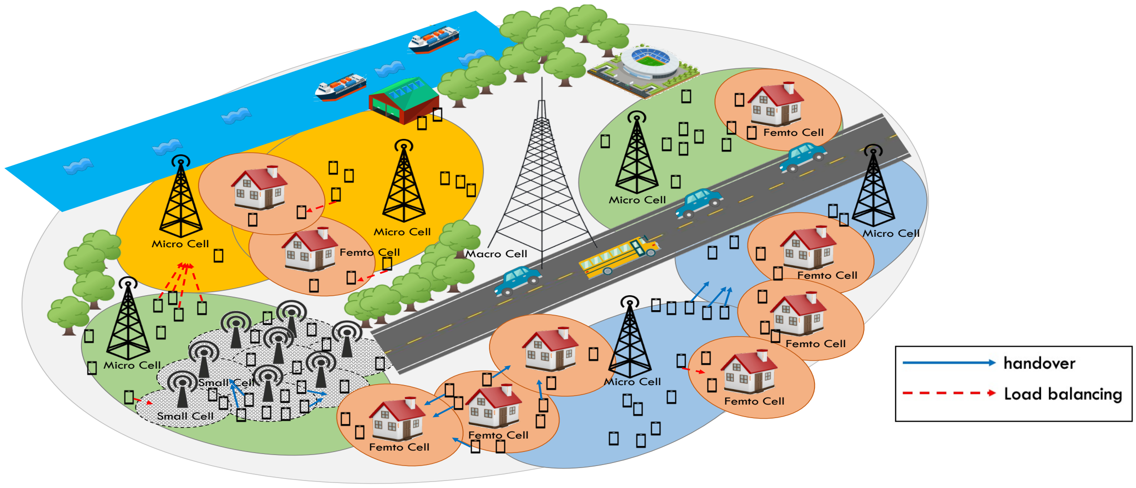

Fig. 4. The Basic HetNets scenarios for 5G Networks.

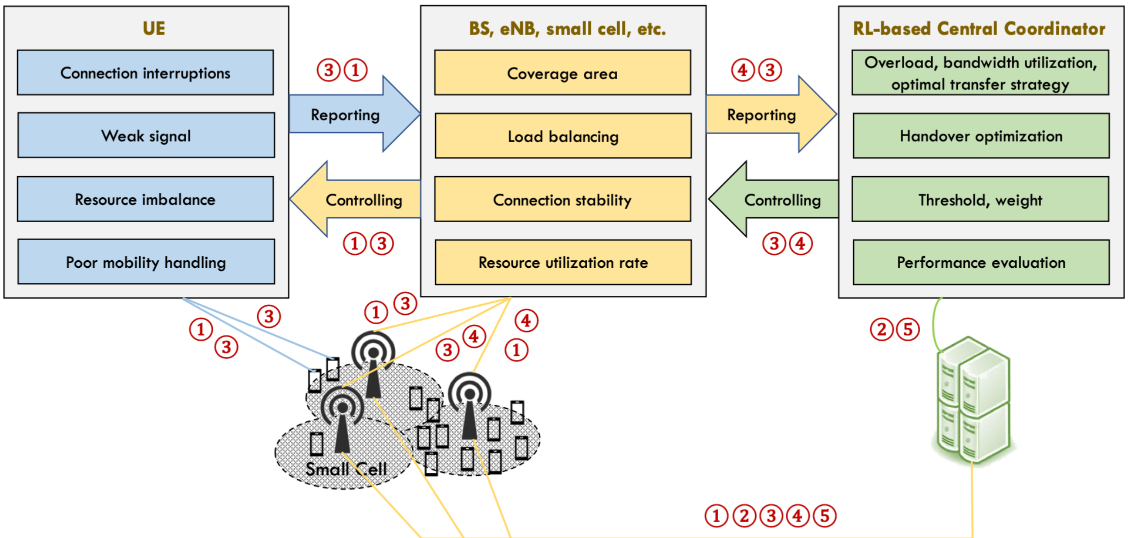

Fig. 5. The SON Framework for Heterogeneous Networks that utilizes DRL includes a sample of an imbalanced small cells.

Iacoboaiea *et al.* [14] propose a SON COordination (SONCO) design that utilizes RL (Reinforcement Learning) to enhance decision-making through learning from past experiences. In order to simplify things, the authors have utilized two types of function approximation and presented a case study. The findings demonstrate that the suggested SONCO design can balance fairness between SONF by allotting weights to SONF. A unified self-management method that is based on fuzzy logic and RL is proposed by Muñoz *et al.* [15]. In their paper, the authors provide a self-management system that is comprehensively based on fuzzy logic and RL [15]. This system mixes the two types of reasoning. The method modifies the switching parameters in a certain way and works to improve the significant performance indicators connected to load balancing (LB) and handover optimization (HO). The findings make it abundantly evident that the performance of their technique is superior to that of the simultaneous operation of individual units in the network. Paropkari *et al.* [16] proposed a novel foundational model for the purpose of acquiring and evaluating UE and network parameters. The objective of this model is to facilitate informed decision-making in relation to user mobility by identifying optimal strategies. Efforts are made to incorporate a multitude of variables in order to effectively forecast the handover, taking into account the newly introduced input parameters from the UE or network. The model put forth by the authors is characterized by its simplicity and adaptability, while effectively encompassing the intricacies of the parameters that govern the process of switching.

Huang *et al.* [17] present a highly effective approach that relies on the load level of neighboring cells. The initial step involves the formulation of the MLB objective, followed by the conversion of the MLB issue into a linear programming problem, which can be effectively addressed using established methodologies. Based on this premise, the determination of suitable parameters for MLB switching is contingent upon the analysis of load distribution within the neighboring cell. Ultimately, the structure of the MLB program is presented. Lateef *et al.* [18], the authors put forth a policy framework predicated on the Trigger-Condition-Action (TCA) model. This framework may be a foundation for developing decision tree logic or combined optimization algorithms for MRO and

MLB functions. Both of these types of optimizations are possible. The researchers found relevant self-coordination mechanisms for each category of conflict by doing research on several conflict kinds as well as hypothetical situations. The authors' proposed analysis holds promise for advancing future research aimed at generating practical solutions for SON that are free from conflicts.

Based on the above literature, this study found that RL can adjust the network configuration in real time according to the current environment state and demand to maximize performance. This dynamic adaptability enables SON systems to better cope with changes in network capacity, interference, and user needs. At the same time, DRL can optimize multiple goals, and by setting the corresponding reward function, the agent can find the best solution to balance the different goals. These advantages make RL and DRL powerful tools for improving the performance and efficiency of SON systems, automating and optimizing network configurations to provide better wireless communication services.

V. RL-Based Framework for SON

In 5G and Beyond 5G (as shown in Figure 4), HetNets characteristics will lead to the increase in network complexity, resulting in numerous potential conflicts that are extremely challenging. Based on the above, this paper proposes the SON framework based on DRL, as shown in Figure 5. ① In the process of coordinating MLB and MRO, the load imbalance in the network is first detected, and then the MLB mechanism is triggered to achieve load balancing. ② In this process, DRL is introduced as a method of coordination, the DRL model is built, and the agent is trained to choose the best action according to the state and reward of the environment. ③ Based on the model's decision, the agent switches the user device to a lighter small cell, monitors the small cell load and mobility toughness metrics, and provides feedback to the DRL model. ④ If a mobility resilience problem is detected, the MRO mechanism is triggered to select the appropriate policy based on the network state and objectives, such as adjusting handshake parameters or changing connection parameters. ⑤ The performance and parameters of the DRL model are continuously monitored and adjusted throughout the process to achieve the best results in coordinating MLB and MRO.

VI. Conclusion

In this article, we discuss the significance of the MLB and MRO concepts in the SONF and future networks and explore the main reasons for conflicts between these two functions. We will also present solutions to address these conflicts. Moreover, we will highlight the expected role and impact of ML in this process and explain the benefits of DRL for wireless networking scenarios. To address the challenges of SON networks in 5G, we propose a SON framework that supports reinforcement learning. Future work will involve identifying DRL-based conflict and coordination scenarios, implementing the proposed scheme, and evaluating the 5G Key Performance Indicator (KPI).

Acknowledgment

This work was supported in part by the Kyushu Institute of Technology – National Taiwan University of Science and Technology Joint Research Program, under the grant Kyutech-NTUST-2023-04.

References

[1] S. Feng, and E. Seidel, "Self-Organizing Networks (SON) in 3GPP Long Term Evolution", Nomor Research GmbH, May 2008.

[2] 3GPP standardization, "Self-configuring and self-optimizing network use cases and solutions (Release 9)", TR 36.902 v1.2.0, May 2009. http://www.3gpp.org/

[3] A. Schröder, H. Lundqvist, and G. Nunzi, "Distributed Self-Optimization of Handover for the Long Term Evolution", in *Proceedings of the 3rd International Workshop on Self-Organizing Systems*, 2008, pp.281-286.

[4] M. I. Tiwana, B. Sayrac, and Z. Altman, "Statistical Learning for Automated RRM: Application to eUTRAN Mobility", *in Proceeding of IEEE International Conference on Communications*, 2009, pp.1 - 5.

[5] R. Nasri, and Z. Altman, "Handover Adaptation for Dynamic Load Balancing in 3GPP Long Term Evolution Systems", *in Proceeding of the Advanced in Mobile Computing Multimedia*, 2007, pp.145-154.

[6] CMCC, "Dependencies among SON use cases and CCO priority", San Francisco, US, 3GPP TSG RAN WG3 Meeting #64, R3-091032, May 2009. http://www.3gpp.org/

[7] A. Lucent, "Exchange of handover parameters directly between eNBs", San Francisco, US, 3GPP TSG RAN WG3 Meeting #64, R3-091294, May 2009. http://www.3gpp.org/

[8] A. G. Barto, "Reinforcement Learning," in A Bradford Book, vol 15, Cambridge, 1998, pp. 665-685.

[9] D. F. P. Rojas, "Benefits and limits of machine learning for the implicit coordination on SON functions," Doctoral dissertation, Dissertation, 2022.

[10] D. F. P. Rojas and A. M. Thiel, "Machine Learning-based SON function conflict resolution," *in Proceeding of IEEE Symposium on Computers and Communications*, 2019, pp. 1-6.

[11] G. Stamatelatos, A. Sgora, and N. Alonistioti, "Intelligent SON Coordination in the 5G-and-beyond era," *in Proceeding of Global Information Infrastructure and Networking Symposium*, 2022, pp. 99-103.

[12] B. Shubyn, N. Lutsiv, O. Syrotynskyi, and R. Kolodii, "Deep learning based adaptive handover optimization for ultra-dense 5G mobile networks," *in Proceeding of IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering*, 2020, pp. 869-872.

[13] J. Moysen, M. G. Lozano, L. Giupponi, and S. Ruiz, "Conflict resolution in mobile networks: a self-coordination framework based on non-dominated solutions and machine learning for data analytics [application notes]," *IEEE Computational Intelligence Magazine*, 13(2), 2018, pp. 52-64.

[14] O. C. Iacoboaiea, B. Sayrac, S. B. Jemaa, and P. Bianchi, "SON coordination in heterogeneous networks: A reinforcement learning framework," *IEEE Transactions on Wireless Communications*, 15(9), 2016, pp. 5835-5847.

[15] P. Muñoz, R. Barco, and I. d. l. Bandera, "Load balancing and handover joint optimization in LTE networks using fuzzy logic and reinforcement learning," *Computer Networks*, 76, 2015, pp. 112-125.

[16] R. A. Paropkari, A. Thantharate, and C. Beard, "Deep-mobility: A deep learning approach for an efficient and reliable 5g handover," *in Proceeding of International Conference on Wireless Communications Signal Processing and Networking*, 2022, pp. 244-250.

[17] M. Huang, and J. Chen, "A conflict avoidance scheme between mobility load balancing and mobility robustness optimization in self-organizing networks," *Wireless Networks*, 24, 2018, 271-281.

[18] H. Y. Lateef, A. Imran, and A. A. Dayya, "A framework for classification of self-organising network conflicts and coordination algorithms," *in Proceeding of IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications*, 2013, pp. 2898-2903.