Received April 19, 2022, accepted May 6, 2022, date of publication May 11, 2022, date of current version May 18, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3174089

Energy Saving Technologies and Best Practices for 5G Radio Access Network

RUMENG TAN[1], YING SHI[1], YINGYING FAN[2], WENTAO ZHU[3], AND TONG WU[4]

[1]China Telecom Research Institute, Guangzhou 510630, China
[2]Shanghai Research Institute, ZTE Corporation, Shanghai 201203, China
[3]China Mobile Group, Design Institute Company Ltd., Beijing 100080, China
[4]National Institute of Metrology, Beijing 100029, China

Corresponding author: Tong Wu (wut@nim.ac.cn)

ABSTRACT This article identifies energy-saving potential of the fifth generation (5G) Radio Access Network, and describes main energy-saving principles and technologies. It explores how to use network energy saving technologies, such as carrier shutdown, channel shutdown, and symbol shutdown in 5G network, that have been inherited from 4G. Some enhanced technologies for 5G like equipment deep sleep and symbol aggregation have also been introduced in this article. However, it is far from enough and an innovative energy-saving solution should be considered. To meet the requirements and development of intelligent and self-adaptive energy-saving solution, Artificial Intelligence (AI) and big data analysis are introduced to form a more precise energy-saving strategy based on site-specific traffic and site-related conditions, thus improving the efficiency and reducing the manpower. Finally, two commercial application practices of AI-based energy-saving solution are elaborated. One is the practice of AI-based service awareness energy saving for 4G/5G collaborative networks, the energy benefits can be improved up to 20%; The other practice is the adoption of a new architecture Active Antenna Unit (AAU) with beam pattern optimization, its energy benefits can be promoted by 30%. These two practices could help mobile network operators (MNOs) to achieve the most energy-efficient network with good network performance and lower Operating Expense (OPEX).

INDEX TERMS 5G, artificial intelligence, base station, energy efficiency, energy saving, radio access network.

I. INTRODUCTION

First of all, climate change is arguably the biggest challenge of our time. Even a Nobel Prize was awarded in 2021 to researchers in this field. And many of MNOs have set ambitious targets about carbon neutral and/or net-zero reported by GSMA [1], for instance, Orange has set 2040 for net-zero and 2050 for carbon neutral based on 1.5-degree centigrade climate changing. Meanwhile, in the guide for the cellular mobile network operators (MNOs) by GSMA about climate target, the goal is to reduce carbon emissions by 45% from 2020 to 2030 [2], [3].

In order to achieve these targets, we need new ways of thinking how we do green network, because the pressure would be huge. The total energy consumption by the mobile industry would be tripled in ten years, according to some forecast, and 5G network would be a big part of the increase [4]. And the biggest challenge for mobile network energy saving is about the base stations (BS), which account for about 73% of total energy usage of a typical mobile network according to GSMA [5]. So the critical thing to do is to reduce carbon footprint in every possible way, but not at the cost of user experience promised.

The good news is we don't have to re-invent all the wheels to 5G energy saving. Network energy saving technologies, such as carrier shutdown, channel shutdown, symbol shutdown etc., that have been emerged since the 4G era, can be leveraged to mitigate 5G energy consumption. Furthermore, enhanced technologies like deep sleep, symbol aggregation shutdown etc., have been developing in the 5G era. In this article it aims to detail these fundamentals.

However, it is far away from being enough, an innovative energy saving solution should be taken into consideration. Powered by AI and big data, a more precise energy saving

The associate editor coordinating the review of this manuscript and approving it for publication was Francisco Rafael Marques Lima.

strategy based on site-specific traffic and site-related conditions, improves the efficiency and reduces the manpower.

It is however not just about the network itself. Because the other industries can also benefit from a greener network either with direct reduction of carbon emission, or using 5G as an underlying technology to make their own businesses more efficient, more productive and therefore greener.

II. ENERGY SAVING DESIGN PRINCIPLES FOR BS

A. DYNAMIC SCALING

The dynamic capability is the capability of scaling power consumption with real-time changes of service loads. The dynamic capability is decomposed into the capability of scaling resource consumption with changes of service loads, and the ability of equipment power consumption to scale as resource consumption changes.

1) RESOURCE ON-DEMAND SCALING

Based on the network performance and user experience, time-, space-, and frequency-based resource occupation and air-interface power resource allocation are adjusted. During low- and medium-load period, reducing resource overhead through proper scheduling is a key path for energy-saving.

The evolution of wireless networks is a process of expanding network capacity and resources to cope with increasing service requirements. The resources are increased to meet the peak capacity of the network. However, traffic is tidal, and the proportion of peak scenarios in a day is not so high. It is therefore necessary to adjust the resources to the required demand.

2) POWER CONSUMPTION SCALES WITH RESOURCES

When resource consumption decreases, equipment power consumption should also decrease. The resource herein may be a bandwidth, a quantity of carriers, a quantity of channels, a transmit power, or the like. For example, the transmission power consumption varies with the number of RF (radio frequency) channels and PA (power amplifier) output power, while the compute power consumption typically varies with the bandwidth and the quantity of channels [6].

B. SHUT-DOWN CAPABILITIES

In order to get power consumption to scale with resources, the static power consumption may become an obstacle. In other words, when scaling down resources the power consumption reaches a level (the static power consumption) where it does not scale any more. In such cases, a shutdown capability of the hardware component(s) is desirable, i.e., a deactivation of the component(s) which decrease the static power consumption stepwise. A drawback with this is that the component is taken out of service, and it will require a delay to be re-activated.

The energy saving gain from shutdown capabilities depends on three dimensions as illustrated in Fig.1; S: Scenario (shutdown scenario) T: time (shutdown duration) D: Depth (shutdown depth). An improvement in a shutdown scenario, shutdown duration, and shutdown depth leads to an improvement in an overall energy-saving capability.

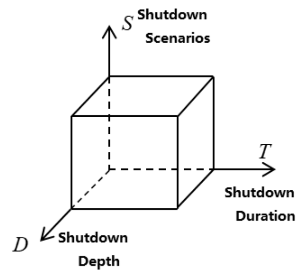


FIGURE 1. Three dimensions of shutdown capabilities.

Extended shutdown scenarios can be that parameters are adjusted such that the shutdown can be applied more often. An improvement of the shutdown duration can be achieved by reducing the common information overhead in no-load or light-load scenarios, or to perform aggregation scheduling for small-packet services. Finally, the shutdown depth is related to how many hardware components that can be de-activated.

Furthermore, the three dimensions are not independent, but are mutually coupled and mutually constrained. When one dimension is strongly expanded, the other two dimensions may be affected. For example, increasing the shutdown depth increases the energy saving gain at the time when the shutdown takes effect, but increasing the shutdown depth may increase the wakeup duration [6]. The increase of the wakeup duration affects the application scenarios and the possibility of entering the sleep state. The overall gain of the whole day may not be the optimal state. Therefore, these three dimensions need to be considered comprehensively in the design of the solution, and strive to find a good balance between them.

C. TRADE-OFFS

Energy-saving involves various trade-offs, as was seen in the section on shutdown capabilities. One trade-off is between **network performance and energy saving**. Based on the impact on performance, many energy-saving are lossy as they scale resources and the number of resources is closely related to performance. Therefore, it is difficult to achieve lossless performance. On the other hand, as mentioned in the preceding analysis, there is a large margin for network experience in off-peak hours, such as night. If the margin is properly squeezed out, there is energy-saving to harvest. However, how to squeeze this space will involve another issue, the qualitative and quantitative analysis of the impact of various energy-saving measures on performance. In many scenarios, this impact cannot be accurately modeled in a white box manner. Although certain parameter thresholds are designed for most features to control the impact on performance, it is difficult to provide accurate guidance on how to adjust parameters. Artificial intelligence (AI) might be a tool to handle this, as will be discussed in section VII.

Another trade-off is between **network energy saving and User Equipment (UE) energy saving**. In most cases, network energy saving is decoupled from UE energy saving. However, trade-off exists in some scenarios. For example, to save power for UEs, the network performs a lot of adaptation in the design of common signaling, and these adjustments usually increase the overheads on the network side. However, an Internet of Things terminal has a strict requirement on low power consumption, which exacerbates this contradiction. The terminal may selectively listen to a paging of a network for energy saving, but the network cannot arbitrarily reduce its paging density.

The number of UEs in a cell fluctuates. However, the network needs to implement flexible scaling of common signaling overheads to reduce the power consumption of the base stations in light-load scenarios. The UE side needs to be aware of and compatible with the dynamic adjustment of the common signaling on the base station side to reduce the impact on processes such as initial UE access and cell handover and re-selection.

III. BASIC ENERGY-SAVING FUNCTIONS FOR 5G RAN

5G will add many new functionalities to the RAN, and will also deploy more resources in the form of carriers, bandwidth, and transmit antennas. This not only means increased capacity and capabilities of the network, by following the design principles in the previous section there are also large opportunities for energy saving technologies.

However, the complexity of the 5G network and the difficulty of operation and management will also increase. This also holds for the management and control of energy saving technologies, as also other aspects such as coverage, interference, user QoS (quality of service) and so on need to be considered when using the different energy saving technologies. It is difficult to use fixed or simple algorithm to guarantee optimal performance. Therefore, AI needs to be introduced and how to use AI technology to realize energy saving applied to operation and management of wireless communication units becomes more and more important. Therefore, section V will be dedicated to AI-driven energy saving, while this section will focus on the energy saving technologies themselves.

A. TIME-DOMAIN ENERGY SAVING

The scheduler allocates a certain number of symbols for downlink data in accordance with the system load and service data forecasts and turns off the PAs to save energy when there is no information being transmitted. Symbol power saving sometimes refers to as cell discontinuous transmission (DTX) [7], as illustrated in Fig.2.

However, even an empty LTE (long term evolution) carrier transmits the cell-specific reference signals (CRS) approximately every 0.2 ms, which means that only very short (2-3 Orthogonal Frequency Division Multiplexing symbols) component de-activations are possible. 5G has addressed this issue by drastically reducing the signal load in idle mode,

FIGURE 2. Illustration of time-domain energy saving. During symbols with no information to transmit, the PA is de-activated.

to allow for more time-domain energy saving. The most necessary control signals for access are still broadcasted but many other signals are only transmitted on-demand from the accessing user devices. Furthermore, the periodicity of the necessary control signals can be configured between 5 ms and 160 ms. For a coverage-providing carrier the default periodicity is 20 ms [8]. This means that in idle mode a 5G carrier can have approximately 100 times longer silent periods than a LTE carrier (20 ms vs. 0.2 ms) [9], during which time-domain energy saving can be applied.

Furthermore, also when the carriers are loaded, time-domain energy saving can be applied. Due to the bursty nature of traffic, empty symbols are common also in loaded traffic scenarios. It is also possible to create empty symbols, time slots, and even sub-frames, by smart scheduling decisions. Hence, time domain energy saving is a very important energy saving technology in 5G.

B. SPATIAL-DOMAIN ENERGY SAVING

In spatial domain energy saving, a number of the RF channels are de-activated in low-load scenarios [10], [11], see Fig.3. This is sometimes referred to antenna or multiple input multiple output (MIMO) muting, or RF channel shutdown. The idea is that the remaining capacity should be enough to serve the traffic, but coverage and data rates may be affected since some of the PAs and thereby available output power is taken out of service. However, there are methods to compensate this, e.g. by common signal power boosting.

Since 5G in general, and massive MIMO in particular, uses more antennas and RF channels than previous generations, spatial domain energy saving is an important energy saving technology in 5G.

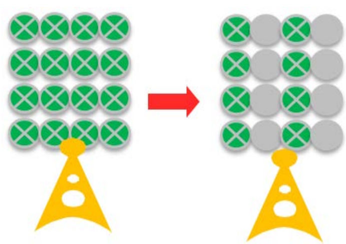


FIGURE 3. Illustration of spatial-domain energy saving. During low-load periods, a number of antenna elements and corresponding RF channels are de-activated.

C. FREQUENCY-DOMAIN ENERGY SAVING

Frequency domain energy saving relates to reducing the available frequency bandwidth and thereby save energy. There are two classes of frequency domain energy saving:

1. Large-scale, in this case an entire carrier is shutdown, and consequently it is sometimes referred to as carrier shutdown. Carriers corresponding to some frequency bands or frequencies are shut down, as illustrated in Fig.4. When these carriers are mapped to independent physical hardware components (e.g. a radio unit), corresponding components can be shut down and thereby save energy.

2. Small-scale, such as bandwidth shrinking and subcarrier shutdown. In this case only carrier specifications are adjusted. The entire carrier is still working, and component shutdown cannot be implemented. In this case, the specifications of certain processing units can only be adjusted to reduce power consumption, which means that the dynamic capability is crucial for the achievable energy saving.

FIGURE 4. Illustration of large-scale frequency-domain energy saving, carrier shutdown. One of the two carriers is shutdown, and the related radio unit is de-activated to save energy.

With the roll-out of 5G on new frequency bands, in particular large-scale frequency-domain energy saving, such as carrier shutdown, will be an important energy saving technology. With low load in the network and overlapping network coverage, some carriers can be shutdown to save energy. This can also be coordinated over different radio access technologies. Taking the overlapping network of NR (New Radio) and LTE as an example, we can think that NR is the capacity system and LTE is the basic coverage system. If the NR traffic is low, the NR can be intelligently shut down and the traffic transferred to LTE at the same time. Once LTE service exceeds a certain threshold, the NR carrier is re-activated. In this way, the power consumption of the whole network changes with the traffic volume.

IV. DEVELOPMENT OF ENERGY-SAVING TECHNOLOGIES FOR 5G

The beginning of network energy saving came with the fact that many sites had traffic peaks and troughs, which means certain parts of BSs could be shut down to save energy, and these included carrier frequency block, carrier frequency shutdown, channel shutdown and symbol shutdown. Details of these energy-saving technologies and their rationales can refer to [12].

With the evolution of 5G network, new technologies used in 5G provide faster and more services which make 5G system more complexity than 4G. However, providing services with cellular network, transferring signals with multiple channel, etc., 5G and 4G have something in common. Thus 5G energy saving does not have to start from scratch, all energy saving technologies of 4G can be leveraged.

However, it is far away from being enough, more new technologies and enhanced technologies need to be developed in 5G era such as equipment deep sleep, LTE/NR carrier cooperative shutdown, enhanced channel shutdown and symbol aggregation shutdown. The rationales of these new developed technologies in 5G are described in [12].

On the other hand, the standards organizations are accelerating the development of relevant energy-saving standards. The 5G NR standard allows more components to switch off or go to sleep when the base station is in idle mode and requires far fewer transmissions of always-on signaling transmissions. In 3GPP Release 18, a new study item - Network Energy Savings [13] - was established as a priority in Dec. 2021, aimed to further explore the energy-saving potentials.

This study item plans to study and develop a network energy consumption model especially for the base station, KPIs (key performance indicators), an evaluation methodology and to identify and study network energy savings techniques in targeted deployment scenarios. The study should investigate how to achieve more efficient operation dynamically and/or semi-statically and finer granularity adaptation of transmissions and/or receptions in one or more of network energy-saving techniques in time, frequency, spatial, and power domains, with potential support/feedback from UE, potential UE assistance information, and information exchange/coordination over network interfaces. The study not only evaluate the potential network energy consumption gains, but also assess and balance the impact on network and user performance, e.g. by looking at KPIs such as spectral efficiency, capacity, user perceived throughput, latency, UE power consumption, complexity, handover performance, call drop rate, initial access performance, service level agreement (SLA) assurance related KPIs, etc. The energy-saving techniques to be developed should avoid having a large impact to such KPIs.

V. AI-BASED ENERGY-SAVING SOLUTIONS

A network intelligent solution based on AI [14] and big data [15] is proposed after review of the research and application progress of AI in the cellular network field [16]–[19]. The solution believes that AI and big data can be introduced into a network on three levels: network element intelligence: operation and maintenance (O&M) intelligence; and service intelligence, with the principles of tiered, on-demand and phased. In this way, ubiquitous intelligence can be achieved.

A. WHAT IS AI

Artificial intelligence is the area of computer science focusing on creating machines that can engage on behaviors that humans consider intelligent. It combines computer science, physiology and philosophy that is a broad topic consisting of different fields, from machine vision to expert system.

The element that the fields of AI have in common is the creation of machines that can "think." The ability to create intelligent machines has intrigued humans since ancient times, and with the advent of the computer and over 50 years research of AI programming techniques, the dream of smart machines is becoming a reality [14].

According to the GSMA technical report [16], artificial intelligence is made up of 3 principal branches, big data, automation and artificial intelligence. Big data gathers large data sets on which analytics are applied to gain insights and enhanced decision making. Automation is where machines follow pre-programmed rules to run processes, generally used for repetitive tasks. The final area is most advanced – Artificial intelligence where machines perform cognitive functions similar to those attributed to humans.

AI is the ability of a computer or machine to emulate human tasks through learning and automation, generally understood to be the simulation of the higher order functions of intelligent beings in areas such as visual processing, speech processing and analytics.

AI algorithms take decisions as a consequence of the application of advanced analytical techniques and may be applied in combination with automated advanced feedback loops to solve problems.

In general, a more detailed AI should seek to emulate or simulate higher-order biological system functions including visual processing, speech/natural language processing, outcomes prediction, objects or data categorization, and problem-solving. However, it should be noted, AI must exclude software systems based on traditional rule-based and determined algorithms, for example, where a specific process or algorithm is designed/programmed by one or more people, which are not based on AI tools or techniques. This is because AI should include a significant element of learning from or adapting to data either as the whole process or an identifiable part of the process.

Artificial intelligence can be further defined by the application of learning that may be undertaken; machine learning and deep learning.

Machine learning uses statistical techniques to perform specific tasks, often requiring a smaller amount of data. In doing so, machine learning can be conducted by low-end systems though usually need labelling and features extraction to perform problem/task breakdown. This means that machine learning applications are faster to train, but testing may be slower to ensure the validity of results. However, these are more readily explainable as the process is understood.

Deep learning, on the other hand, uses artificial neural networks which require a more substantial amount of data to train models. This, in turn, requires high-performance GPUs (graphics processing units) but allows deep learning to process unlabeled data and solve end-to-end problems. As a result of its reliance upon large data sets, deep learning often is slower to train, however, faster to test, the biggest drawback here is what is referred to as the "black box" – while the inputs and outputs may be understood, the steps taken may not be.

B. AI TECHNOLOGY IN NETWORK ENERGY SAVING

Major vendors are currently offering AI-based energy saving solutions, and the commercial application results showed that 80% activation time has been increased and over 10% even 15% energy consumption has been saved in the entire network [12], [19].

Some leader operators have introduced intelligent capabilities in their network progress, like network operation and maintenance, network planning and optimization. According to the forecast of Analysys Mason, 80% operators hope that 40% network intelligence will be realized, while one third of them expect the rate will exceed 80% by 2025 [16].

In response to the requirement of an intelligent and self-adaptive energy saving solution, AI and big data technologies are accelerating to be exploited and introduced into network energy saving for improving the efficiency and reducing the manpower required [17], [19].

C. AI-BASED ENERGY SAVING FOR 5G

Nowadays the 5G network deployment is on the fast track around the world. Many MNOs are currently running 2G, 3G, 4G and 5G networks at the same time. The time distribution of cellular network traffic has often obvious peaks and troughs, basic functions applied to the entire cellular network is not a site-specific strategy, resulting in less efficient due to ignoring the traffic and neighboring sites patterns varied from site to site, especially in more complexity network.

That is why AI and big data technology are introduced to form a more precise energy saving strategy based on specific site traffic and other site-related conditions. The AI-based network energy saving solution can forecast the traffic load of base stations based on historical traffic load, service type, site coverage and user behaviors. Energy saving strategy can automatic configuration based on coverage identification and configuration identification by AI technology. The suitable energy saving strategy combined with different energy saving functions, including an initial relative threshold to the scenario and executable energy saving time schedule, will be enabled for the sites that are expected to have energy saving effects. Meanwhile, AI-based network energy saving solution can also ensure the balance between network power consumption and network performance based on sufficient model training.

The basic functions of an intelligent energy-saving solution driven by AI should include but not limited to the following: scene identification, traffic forecast, multi-network coordinate control, online iteration and optimization based on network performance, etc. [12]

D. ENERGY SAVING SCHEME ARCHITECTURE BASED ON AI

Fig.5 shows the overall architecture of the AI-based energy-saving network [12], [20]. Based on the intelligent application platform, the artificial intelligence algorithm is adopted to achieve the maximum balance between the system performance and the energy-saving effect, so as to achieve the network energy-saving and consumption reduction.

By shutting down the resources that are not utilized and keeping network capacity adequate yet minimum, network

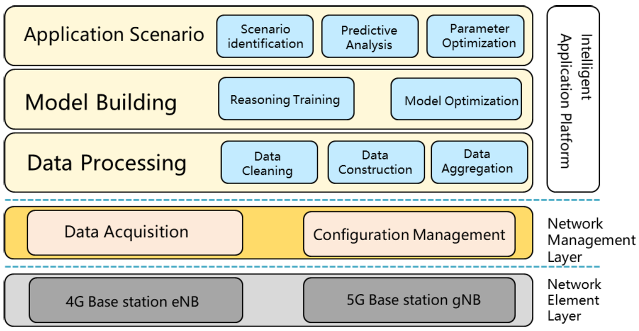


FIGURE 5. Architecture diagram of AI-based energy-saving network.

energy consumption can be optimized in line with the network traffic load forecast by AI-based.

For 5G era, enhanced AI-based energy-saving solutions can take account of the different efficiency levels of frequency bands and factors in that the power efficiency of different networks can vary. By directing users from their less power-efficient spectrum band(s) to other band(s) that are more power-efficient, more radio resources can be shut down to lower network energy consumption.

Alongside energy-saving potential, AI-based solution also can constantly monitor customer experience, network availability and data traffic to ensure there is no impact on network performance.

VI. APPLICATIONS FOR 5G ENERGY-SAVING SOLUTIONS

The applications of 5G energy-saving solutions mostly separate into two stages respectively with two modes: traditional manual mode and AI-based machine learning mode.

A. TRADITIONAL MANUAL MODE

Based on basic analysis of coverage, traffic, the traditional energy saving solution is applied for simple scenarios with unified strategy, for instance, channel shutdown will be activated when the PRB utilization is below 10% during 10pm to 6am. The strategy is configured on OMS (Operation Management System) by manpower, and the energy saving efficiency will be calculated with performance counters for energy consumption, while the KPIs are also monitored by manpower to ensure the network performance. Fig.6 illustrated a step-by-step procedure of 5G energy-saving based on manual configuration mode.

In the early commercial operation stage, the energy-saving effect of manual mode is quite good notwithstanding the lack of historical data. But as the network users and traffic-loads increase, the risk of degradation of network KPIs will increase.

With the continuous development of 5G networks and users, traditional manual mode that mainly rely on engineers' experience can no longer meet the users' and verticals requirements.

Thus, the second mode, AI-based technology with high-computational data analysis, cross-domain feature mining, and dynamic policy generation will come on the next stage.

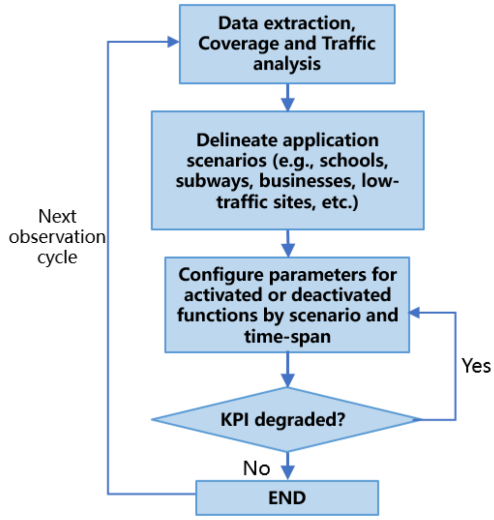


FIGURE 6. Flowchart of 5G energy-saving based on manual configuration mode.

B. AI-BASED MACHINE LEARNING MODE

Site-specific configuration, historical traffic and network KPIs will be collected. Combined with time series prediction algorithm like LSTM (long short-term memory), intelligent optimization algorithm like ant colony optimization and clustering algorithm like K-Means, the network coverage, user distribution and initial threshold will be output. With traffic forecast and evaluation radio resource utilization in real-time, cell-specific energy saving strategy will be applied to the network to ensure the balance between network performance and energy consumption.

Fig. 7 illustrates a typical flowchart of AI-based 5G energy-saving (ES) application procedure. Based on the RAN intelligent application platform, the AI algorithm is used to achieve the maximum balance between system performance and energy-saving effect to realize the optimal network energy-saving and consumption reduction by using the basic data such as configuration, performance statistics and

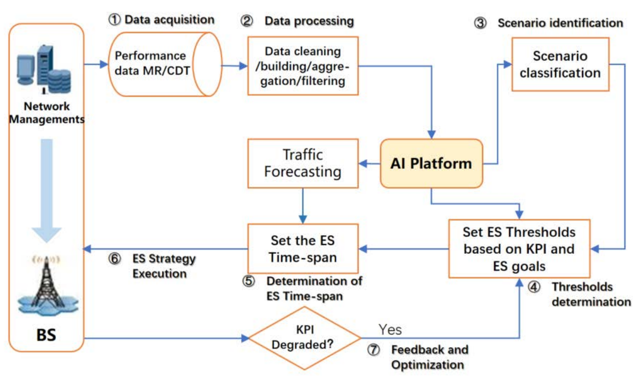


FIGURE 7. Flowchart of 5G energy-saving based on AI configuration mode.

measurement report or call detail trace (MR/CDT) of the existing network.

VII. BEST PRACTICES OF AI-BASED ENERGY-SAVING SOLUTIONS FOR 5G

A. THE PRACTICE OF SERVICE AWARENESS ENERGY SAVING FOR 5G

AI-based energy saving solution with traffic forecast does improve the energy saving efficiency of basic functions [16]–[19]. But in most multi-mode and multi-frequency cellular network, it still has some limitation that service efficiency varies from mode to mode, and/or band to band. And if all services/users are concentrated in part of the network/band, more energy consumption could be saved after idle network/band shut down or deep sleep.

AI-based service awareness in 5G network should also be taken into consideration, which exploits the differences in energy efficiency of different types of services to deliver certain services to the most energy-efficient network, helping achieve the most efficient energy usage without impact on user experience [18].

Fig.8-Fig.11 illustrates an AI-based pilot solution based on service-awareness energy-saving. The solution involves providing automation capabilities for the service management layer and resource management layer. Based on network-level AI-based intelligent energy saving policy management and site energy saving scheduling control, the mobile network energy saving solution implements network scene adaptation, one site one policy, and multi-network collaboration for intelligent base station energy saving management. This maximizes network energy saving benefits while ensuring stable network performance and achieves the optimal balance between energy consumption and KPIs.

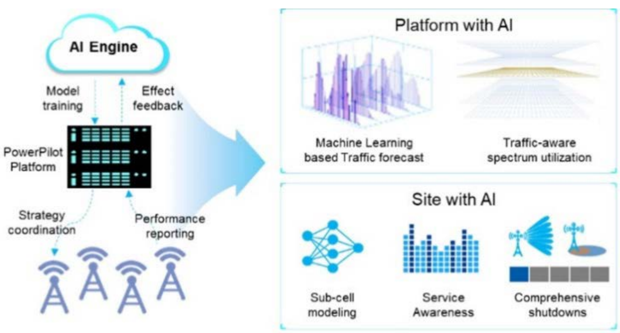


FIGURE 8. AI-based network energy-saving accurate deployment.

1) CELL-SPECIFIC INITIAL STRATEGY

Initial energy saving strategy can automatic configuration based on coverage identification and configuration identification by AI technology. The suitable energy saving strategy combined with different energy saving functions, including an initial relative threshold to the scenario and executable energy saving time schedule, will be enabled for the sites that are expected to have energy saving effects.

1) Idle or low traffic period based on historical traffic analysis
2) Energy saving threshold based on network traffic load
3) Energy saving activation time based on threshold
4) Energy saving function combination

2) TRAFFIC LOAD FORECAST

The intra-week sub-sequence split prediction method, in which each of the seven days in a week needs to be put into the series of its responding days of all the weeks, is combined with cell types, holiday factors and forecast of network traffic load. After putting all the algorithm candidates (linear regressive, ARIMA (autoregressive integrated moving average model) and LSTM, second-order exponential smoothing, etc.) into tests, the one with the best results is selected.

The result of commercial application cases showed that the prediction accuracy exceeds up to 90% for uplink/ downlink PRB utilization and RRC connected users. Fig. 9 illustrates the prediction accuracy of the commercial application case, compared the predicted value with the real traffic in DL/UL PRB usage and RRC users every 15 minutes. The predicted value matches well with the actual value in normal scenarios.

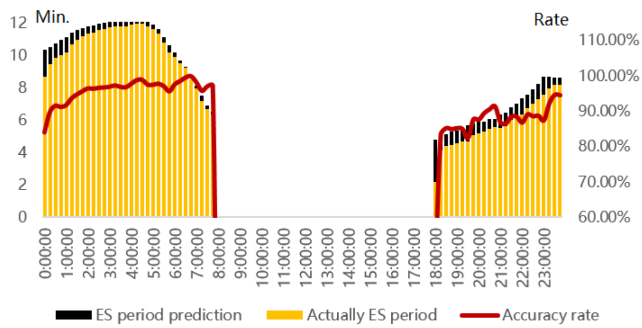


FIGURE 9. The commercial application case and the prediction accuracy.

3) SERVICE AWARENESS

Based on energy saving functions and AI-driven traffic load prediction, this pilot solution is an industry's trial project to introduce AI-powered service-awareness energy saving. By identifying service types and their energy efficiency differences, this solution can evaluate service requirements in real time and support the service with networks of higher energy efficiency to maximize energy efficiency in the entire network.

To improve the network energy efficiency based on user redistribution, there are three main steps: target network/band selection, suitable user selection and consequent user direction. Service efficiency varies from network to network, and/or band to band, the most energy-efficient network/band will be selected as target and the most suitable users will be selected. Also, energy efficiency pattern may change after user direction, leading a new round of optimization. Thus the three steps together form a closed loop.

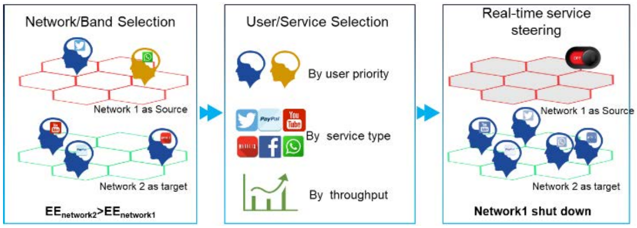


FIGURE 10. Service-awareness energy saving.

4) ONLINE ITERATION AND OPTIMIZATION OF THE THRESHOLD

In order to improve the energy saving efficiency, online iteration and optimization of the threshold can be used instead of the traditional ways which does not take site variations into consideration and results in safe but inefficiency energy saving settings.

Clustering algorithm is used to find out the optimal energy saving threshold settings with the best efficiency and shortest time. In any case when the KPI baseline is compromised, the threshold can be rolled back.

5) APPLICATION AND PERFORMANCE

This solution is put into commercial use in Chengdu city, to verify this commercial energy saving solution with service pilot. The results with Chengdu networks show that the over 35% network energy consumption of 4G/5G can be reduced without impact on the network performance or user experience.

This commercial trial in Chengdu 4G/5G network involves three phases. With only the basic energy saving function, about 9 kWh energy is saved daily per site; when the AI-driven traffic forecast is enabled simultaneously, approximately 12 kWh energy is saved daily per site. After AI-based service pilot is enabled, up to 14 kWh energy can be saved daily per site. The proportion of energy saving is increased from 16.6% to 24.5% (as illustrated in Fig. 11). During the trial, there is less impact to network performance, both the 4G network and 5G network.

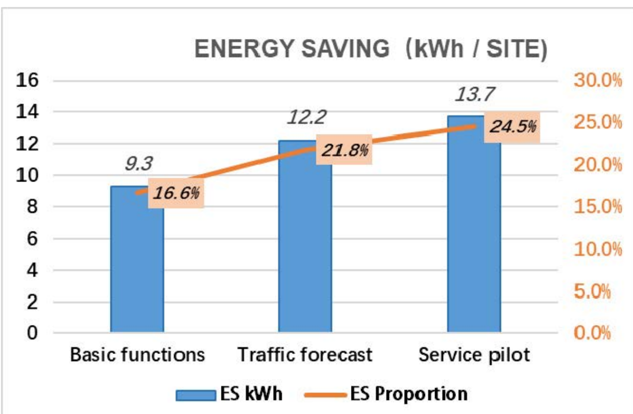


FIGURE 11. Energy saving per site daily (kWh and Proportion).

It is estimated that more than 5 million kilo-watt-hours will be saved annually per thousand stations and at least 2,500 tons of carbon emissions will be reduced if the pilot solution is deployed in the whole 4G/5G network in Chengdu.

And according to typical network configuration calculations, the energy saved by this pilot solution is twice as much as that of the conventional AI-based energy saving solutions, and it can save up to 20% of energy in a multi-mode network, thereby effectively reducing the operational expenditure.

B. THE PRACTICE OF NEW ARCHITECTURE AAU WITH BEAM PATTERN OPTIMIZATION

The cooperation and optimization of beams between cell and UE could avoid aimless radiation of base station. The practice of this type of energy saving are still on trial and ongoing research. Here is a description of the early practice results.

For antenna array of AAU, the larger size of the antenna array, the narrower beam of the channel, thus, the more focused channel energy, and the longer distance of the network coverage. With the introduction of an ultra-large-scale antenna array, the uplink and downlink coverage can be increased synchronously without increasing the transmission power. Therefore, the use of ultra-large-scale antenna arrays is an important innovation direction to improve coverage and reduce energy consumption. Through the integration and innovation of new architectures and new algorithms, both network performance and energy saving can be promoted, helping MNOs to build better and greener 5G networks.

A new pilot practice with such ultra-large-scale antenna arrays of AAU are deployed in Xiamen city, which doubled the size of the antenna array (elements from 192 to 384) and manufactured with new materials and architectures. It adopts adaptive beam optimization and intelligently adapts the user's wireless channel changes to improve the utilization efficiency of air interface resources; and adopts high-resolution beam domain noise reduction to improve the efficiency of multi-user pairing. This enables the new AAU to achieve "accurate-alignment," "fast-following," and "good-pairing," which can greatly improve user experience and cell capacity. The comparison and performance gains of this new architecture AAU are illustrated in Figure 12 and Table 1.

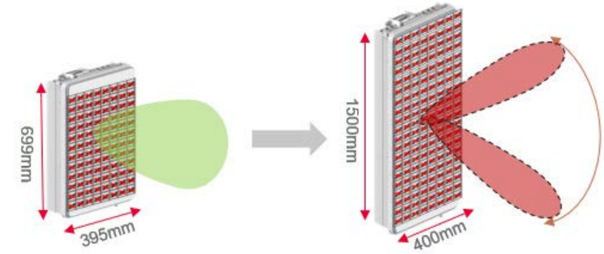


FIGURE 12. Comparison of Traditional AAU and New Architecture AAU (number of antenna elements is from 192 to 384).

Test results show that compared to the traditional solution, the new AAU can increase the cell's 30% uplink and

TABLE 1. Performance gains for new architecture AAU compared to traditional AAU.

Performance gains	Traditional AAU	New Architecture AAU
Coverage	Baseline	UL +3dB ↑ DL +3dB ↑
DL Throughput Rate at Edge Area	Baseline	30~40% ↑
UL Throughput Rate at Edge Area	Baseline	40~60% ↑
Power Consumption	Baseline	-10% ↓

downlink coverage, and the average experience rate gains for edge users can be increased by 25%. If the cell coverage index on edge area remains unchanged, the base station can be configured with lower transmission power, thereby to reduce the energy consumption of the base station. Compared with the traditional AAU, the energy consumption can be reduced by about 30%.

VIII. CONCLUSION

In this article, we discussed the energy saving principles and the practical application technologies such as the carrier shutdown, channel shutdown and symbol shutdown emerged since 4G era for mobile communication networks, which can also be leveraged to mitigate 5G energy consumption. We also elaborated some enhanced technologies, e.g., deep sleep and symbol aggregation shutdown, which have been developing in 5G era. Meanwhile, with the rapid development of AI and other emerging technologies, energy-saving solutions incorporating AI can speed the realization of efficient collaboration of multi-networks and can be introduced into all aspects of communication network operation and maintenance. As an application illustration, this article presented two best practices for AI-based smart energy saving solutions. One is the practice on AI-based service awareness energy saving for 4G/5G collaborative network, the other one is adoption of new architecture AAU with AI-based beam optimization. Both of them could help to achieve the most energy-efficient network with good performance and lower OPEX.

However, network-wide intelligence is difficult to be achieved overnight, a long-term development is required. With the continuous accumulation of data from commercial networks, the machine-learning algorithm will enable AI based power saving solution to be perfect gradually. The AI algorithm itself will be evolved iteratively to realize higher efficiency and accurate strategy adapting when network topology and traffic model changes.

Here is the thought of widen the impact:

1) Feedback to vendor industry to design more intelligent equipment to be more adaptive.

2) Impact the consumer product such as cell phone and cell phone software to more intelligent and corporate with the network in power and signal strength usage.

3) Vendor equipment design to be more integrated with AI to reduce power usage.

With the fact that 5G network deployment and its operating time is relatively short, compared to existing 2G/3G/4G networks with abundant data accumulation, how to set initial power saving strategy of 5G power saving and its reference parameters and how to perform the power saving collaboration between 4G and 5G are all undergoing studies and will be optimized along the way when more experience is accumulated after the solution been widely deployed.

Nevertheless, with the continuous innovation and evolution of 5G energy-saving technologies based on digital technologies we have sound reasons to believe that the energy efficiency of future networks will become more efficient, and the OPEX of MNOs will become much lower to achieve the NetZero goals [1], [2] for ICT industry early.

REFERENCES

[1] *Mobile Net Zero: State of the Industry on Climate Action 2021*, GSMA, London, U.K., 2021.

[2] *Setting Climate Targets: A Step by Step Guide for Mobile Network Operators to Set Science Based Targets*, GSMA, London, U.K., 2020.

[3] *GHG Emissions Trajectories for the ICT Sector Compatible With the UNFCCC Paris Agreement*, document ITU-T L.1470 (01/2020), Recommendation ITU-T, 2020.

[4] A. S. G. Andrae, "New perspectives on internet electricity use in 2030," *Eng. Appl. Sci. Lett.*, vol. 3, pp. 19–31, Jun. 2020.

[5] *Going Green: Benchmarking the Energy Efficiency of Mobile*, GSMA, London, U.K., 2021.

[6] B. Debaillie, C. Desset, and F. Louagie, "A flexible and future-proof power model for cellular base stations," in *Proc. IEEE 81st Veh. Technol. Conf. (VTC Spring)*, May 2015, pp. 1–7, doi: 10.1109/VTCSpring.2015.7145603.

[7] P. Frenger, P. Moberg, J. Malmodin, Y. Jading, and I. Godor, "Reducing energy consumption in LTE with cell DTX," in *Proc. IEEE 73rd Veh. Technol. Conf. (VTC Spring)*, May 2011, pp. 1–5, doi: 10.1109/VETECS.2011.5956235.

[8] *Physical Layer Procedures for Control (Release 16), v16.0.0*, document TSG RAN (2019). TS 38.213; NR; 3GPP, 2019.

[9] P. Frenger and R. Tano, "More capacity and less power: How 5G NR can reduce network energy consumption," in *Proc. IEEE 89th Veh. Technol. Conf. (VTC-Spring)*, Apr. 2019, pp. 1–5, doi: 10.1109/VTCSpring.2019.8746600.

[10] P. Skillermark and P. Frenger, "Enhancing energy efficiency in LTE with antenna muting," in *Proc. IEEE 75th Veh. Technol. Conf. (VTC Spring)*, May 2012, pp. 1–5, doi: 10.1109/VETECS.2012.6240248.

[11] H. Halbauer, A. Weber, D. Wiegner, and T. Wild, "Energy efficient massive MIMO array configurations," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2018, pp. 1–6, doi: 10.1109/GLOCOMW.2018.8644331.

[12] *Smart Energy Saving of 5G Base Station: Based on AI and Other Emerging Technologies to Forecast and Optimize the Management of 5G Wireless Network Energy Consumption*, document FG-AI4EE D.WG3-02, ITU-T Technical Report, 2021.

[13] (Dec. 2021). RP-213554. *New SI: Study on Network Energy Saving for NR*. 3GPP TSG RAN Meeting #94e. [Online]. Available: https://www.3gpp.org/ftp/TSG_RAN/TSG_RAN/TSGR_94e/Docs/RP-213554.zip

[14] H. H. de Garis, "The artilect war—Cosmists vs. terrans—A bitter controversy concerning whether humanity should build godlike massively intelligent machines," Etc Pubns., Pittsburgh, PA, USA, Tech. Rep., Feb. 2005.

[15] V. Mayer-Schönberger and K. Cukier, *Big Data: A Revolution That Will Transform how we Live, Work and Think*. Boston, MA, USA: Houghton Mifflin Harcourt, 2013.

[16] (2019). *AI and Automation: An Overview*. London: GSM Association. (Mar. 10, 2022). [Online]. Available: https://www.gsma.com/futurenetworks/wiki/ai-automation-an-overview/
[17] *AI in Network: Intelligent Autonomous Network Case Report*, GSMA, London, U.K., 2019.
[18] (2020). *5G Energy Efficiencies: Green is the new Black*. GSMA Intelligence. (Mar. 10, 2022). [Online]. Available: https://data.gsmaintelligence.com/api-web/v2/research-file-download?id=54165956&file=241120-5G-energy.pdf
[19] (2019). *5G-Era Mobile Network Cost Evolution*. London: GSM Association. (Mar. 9, 2022). [Online]. Available: https://www.gsma.com/futurenetworks/wiki/5g-era-mobile-network-cost-evolution/
[20] T. Rumeng, W. Tong, S. Ying, and H. Yanpu, "Intelligent energy saving solution of 5G base station based on artificial intelligence technologies," in *Proc. IEEE Int. Joint EMC/SI/PI EMC Eur. Symp.*, Jul. 2021, pp. 739–742, doi: 10.1109/EMC/SI/PI/EMCEurope52599.2021.9559261.

RUMENG TAN was born in Shucheng, Lu'an, Anhui, China, in 1973. He received the B.S. degree in electromechanical engineering from the University of Petroleum, Beijing, China, in 1995, and the M.S. degree in electromechanical engineering from Southeast University, Nanjing, China, in 2003.

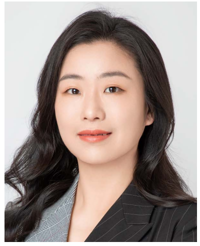
From 2003 to 2021, he was a Researcher with the Mobile Communication Institute of China Telecom Research Institute, Guangzhou, China. He is an active participant in the development of ITU-T Recommendations and Chinese ICT industrial standards, working within ITU-T SG5 and CCSA TC9, which creates and manages the EMC/EMF standards for ICT industry. He has led the drafting of ITU-T recommendations K.121, K.141, and L.Suppl.43, and published over 15 Chinese communication industrial standards. His research interests include the evolution of mobile communication networks, radio frequency electromagnetic fields, and energy saving and carbon reduction technologies and solutions in ICT industry.

Mr. Tan also serves as the Vice-Chair for the CCSA TC9 WG3 (Electromagnetic Radiation and Safety).

YING SHI was born in Wuhan, Hubei, China, in 1974. She received the B.S. degree in electrical engineering from Wuhan University, Wuhan, in 1995, the M.S. degree in electrical engineering from the Guangdong University of Technology, Guangzhou, China, in 1998, and the M.S. degree in telecommunication management from the University of Strathclyde, U.K., in 2007.

From 1998 to 2008, she was a Researcher with the Metrology Testing Center, China Telecom Research Institute, Guangzhou. From 2009 to 2021, she was a Manager and a Senior Expert of the Research and Development Management Department, China Telecom Research Institute, Guangzhou. She is an active participant in the development of ITU-T recommendations and Chinese ICT industrial standards. She has published over ten Chinese communication standards. Her research interests include environmental engineering, electromagnetic protection, and energy saving and carbon reduction technologies in ICT industry.

Ms. Shi currently serves as the Vice-Chair for CCSA TC9.



YINGYING FAN was born in Shanghai, China, in 1982. She received the B.S. and M.S. degrees in communications and signal processing from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2004 and 2007, respectively.

From 2007 to 2015, she was a Software Testing Engineer with the Shanghai Research Institute, ZTE Corporation, Shanghai, who was major in service procedure of GSM/UMTS/LTE system. Since 2015, she has been the RAN Production Planning Director of the Shanghai Research Institute, ZTE Corporation, who was major in energy saving of radio access network and network sharing. She is an active participant in the development of ITU-T recommendations, working within ITU-T SG5 for researching environmental efficiency for artificial intelligence and other emerging technologies.

WENTAO ZHU was born in Changge, Henan, China, in 1986. He received the B.S. degree in optical information science and technology and the M.S. degree in electromagnetic field and microwave technology from Beijing Jiaotong University, Beijing, China, in 2007 and 2009, respectively.

From 2009 to 2021, he was a Researcher with China Mobile Group Design Institute Company Ltd., Beijing. He is an active participant in the development of ITU-T recommendations and Chinese ICT industrial standards working. He has participated in the drafting of recommendation ITU-T K.149 and participated in the application of about 20 patents in the antenna and wireless fields. He has published over ten Chinese communication standards. His research interests include antenna system test and analysis, electromagnetic radiation test, antenna diagnosis based on big data, and massive MIMO technologies.

TONG WU worked with the EMC Certification Center, Radio Department, Department of Bioenergy and Environment, and Department of Information and Electronics, National Institute of Metrology, China. He also worked as the Secretary of the National Technical Working Group on Standardization of Assessment Methods in the fields of electricity, magnetics and electromagnetism for human irradiated bodies, undertook the technical counterpart work of IEC/TC 106. He is currently an Associate Researcher and the Deputy Director. He is mainly responsible for the measurement standards of electromagnetic environment, data center electromagnetic environment, and interference testing research. He is familiar with measurement technology, research method and standardization of electromagnetic environment, and established EMC capacity verification system.