Advanced Sleep Modes to comply with delay constraints in energy efficient 5G networks

Advanced Sleep Modes to comply with delay constraints in energy efficient 5G networks

Michela Meo, Daniela Renga, Zunera Umar
Department of Electronics and Telecommunications
Politecnico di Torino
Turin, Italy
{firstname.lastname}@polito.it

***Abstract*—The staggering growth of mobile traffic fostered by the extensive spreading of 5G technology and massive Internet of Things (IoT) applications is leading to network densification, entailing a boost in network power consumption, with consequent higher operational cost for Mobile Network Operators (MNOs) and raising sustainability issues. To reduce energy consumption when the traffic is low, new BSs feature Advanced Sleep Modes (ASM) that allow to reduce the network energy consumption by gradually deactivating the BSs into progressively deeper sleep modes with lower power consumption. However, the deep sleep modes cause high reactivation delays that may jeopardize the Quality of Service.**

In this paper, focusing on the periods in which traffic is very low, we extensively investigate the potentiality of ASMs based operation in terms of the trade-off between energy saving and delay under different 5G scenarios and traffic loads. By observing that optimal configuration settings depend on the scenario and on the load, we design a framework based on a stochastic model to perform dynamic tuning of the configuration settings that adapts in real time the parameters to the actual traffic load and scenario.

***Index Terms*—Sleep Mode, 5G networks, Energy Efficiency, IoT**

I. INTRODUCTION

In recent years, we have witnessed a staggering growth of mobile traffic, posing significant challenges to Mobile Network Operators (MNOs) in terms of network energy demand and operational costs. This trend is bound to be further boosted in the next years. According to the Cisco forecast, almost 30 billion of networked devices are expected by 2023, and the global mobile data traffic will reach almost 80 exabytes per month by the same year [1]. In this context, the wide spreading of 5G technology, which enables mobile networks to provide incredibly huge bandwidth capacity and ubiquitous and fast network access, is further pushing the process of mobile network densification. The enhancements that 5G technology is introducing in terms of higher bandwidth availability and ultra-low latency are facilitating the deployment of massive Internet of Things (IoT) applications, from IoT services focused on environmental monitoring, to smart mobility and autonomous vehicles, machine-to-machine (M2M) communications for factory automation in Industry 4.0, Smart Grid management, Smart Farming, just to cite some examples. By 2023, IoT devices will represent half of all networked devices, and over 10% of devices and connections will have 5G capability [1].

Considering that up to 80% of the total network consumption is accounted for by the access segment [2], a rapid growth of the energy demand to operate cellular networks is entailed by this boost of traffic volumes. Indeed, MNOs are already facing a substantial rise of the energy bill due to power supply [3]. Among the possible solutions to cope with these challenges, Base Station (BS) Sleep Modes (SMs) can be used to reduce mobile network consumption by deactivating unneeded radio resources during periods in which the traffic demand is low. However, the reactivation time from the sleep mode introduces additional undesirable set-up delays and, consequently, the risk to compromise quality of service. This has slowed down the adoption of sleep modes. Recently, the introduction of Advanced Sleep Modes (ASMs), which allow to progressively deactivate the BS to deeper sleep modes that correspond to lower power levels but longer reactivation times [4], are making it possible to find convenient trade-offs between energy saving and reactivation delay. Clearly, ASMs can be applied in scenarios, like those expected in 5G deployments, in which the cell layout is dense and a few macro-cells provide full coverage and are always active, while several small cells can be activated on demand, only when needed.

While ASMs are very promising, they are defined through a large set of parameters whose setting is not straightforward and might compromise the effectiveness of the overall approach. To the best of our knowledge, no other work in the literature analyses the effect on the system performance of varying the ASM configuration settings and its potential in energy efficient 5G networks. This paper focuses on BS operation with ASMs and investigates the impact of the parameters' setting on the trade-off between energy saving and delay. The main objective of the work is twofold. On the one side, we show that good parameters' settings strongly depend on both the scenario and the load, so that an adaptive definition of the parameters of the ASMs is desired. On the other side, we propose a simple stochastic model that can be used in real time to dynamically define, in any given scenario and under any load, a proper parameters' setting.

II. RELATED WORK

Several studies in the literature focus on investigating the benefits yielded by the application of Resource on Demand techniques based on SMs, showing the effectiveness of these approaches in improving the energy efficiency of mobile networks [5]–[7]. The interest on SM based strategies is expected to further increase with the widespread diffusion of 5G technologies [6], [8], [9]. Indeed, the 5G paradigm entails a staggering network densification, leading to increasingly higher energy consumption [9], [10]. Furthermore, with the current penetration of Renewable Energy sources to power BSs, SM based techniques allow to more efficiently cope with the intermittent and unpredictable production of Renewable Energy [5], [7], [11]–[14].

Nevertheless, SM based approaches may raise concerns in relation to the risk of jeopardizing Quality of Service, due to the time needed to switch on BS components in case of a new service request.

Typically, traditional SM approaches completely deactivate the BS or put it in a deep sleep state, with negligible power consumption but at the price of a high delay not compatible with 5G tight delay constraints. Hence, our study exploits the flexibility provided by Advanced Sleep Modes (ASMs), that allow to progressively deactivate the BS components, to gradually achieve lower power levels [4]. With respect to standard SM techniques based on switching BSs off, the time required under ASM operation to reactivate the BS is lower, becoming larger as the power level decreases while entering deeper SM levels. While ASM approaches allow to trade off energy saving and additional reactivation delay, the ASM operation is typically considered under standard configuration settings, that do not allow to fully take advantage of the ASM potentiality. In this work, we develop a framework for adapting the ASM operation and its parameters' setting to the actual load and scenario.

III. 5G SCENARIOS

To investigate the Advanced Sleep Mode BS operation, we consider three representative 5G scenarios. Clearly, the scenarios are considered under low traffic conditions, since ASMs apply during low traffic periods only. Moreover, we always assume that ASMs apply to BSs that can be activated on demand; i.e., BSs that overlap with a set of macro-cells providing full coverage, so that when a UE receives a call or issues a request for service, the signaling occurs on the macro-cell and the macro-cell activates the BS in sleep mode.

- ***Signaling (SIG) scenario:*** the BS has no traffic to deliver, but it has to periodically send synchronization and control signals. The request arrival rate, denoted by λ, is constant and the average transmission time of signaling information is 1 ms. Focusing on periods of low traffic demand, no other type of traffic data is assumed.
- ***Massive IoT (MIoT) scenario:*** from time to time, the BS has to forward a few data collected from a Wireless Sensor Network. Inter-arrival times are exponentially distributed and the average service time is equal to 10 ms. Signaling data exchange is assumed to be negligible.

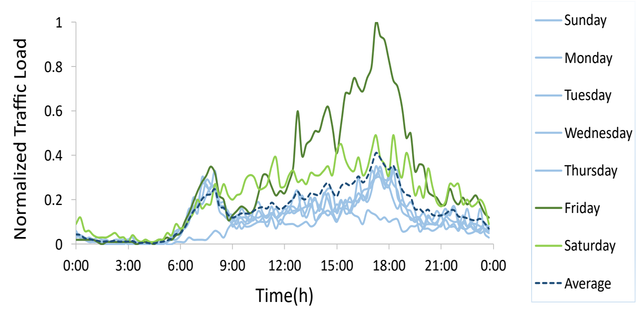


Fig. 1: Daily mobile traffic patterns in different days of the week in a big city.

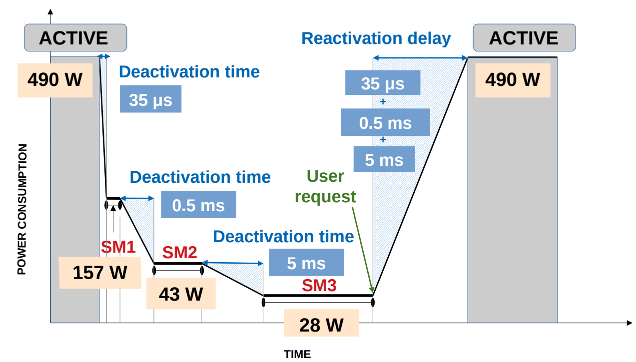


Fig. 2: Advanced Sleep Mode operation.

- ***Standard 5G traffic (S5GT) scenario:*** the BS handles (low) mobile user traffic. Requests inter-arrival times are exponentially distributed and the average service time is equal to 30 s. Even in this case traffic load due to signaling is assumed to be negligible.

IV. ADVANCED SLEEP MODES

The potentiality of ASMs for energy saving is remarkable, since mobile access networks are typically overprovisioned and subject to traffic patterns with high intra-day variability, and long periods of low traffic during which a lot of energy is wasted. As an example, Fig. 1 shows the normalized average daily traffic pattern registered in different days of the week in a touristic area of a large city in Italy. The potentiality for energy saving comes from the traffic being negligible during the night time for all the days of the week.

We consider the Advanced Sleep Modes paradigm as described in [4]. The paradigm was first introduced by IMEC [15] within the Earth project [16] and is currently being integrated in the standardization process to define the New Radio for 5G networks [17], [18].

A. Multiple Sleep Mode operation

According to the ASM approach, whose operation is depicted in Fig. 2, during idle periods the BS can be gradually put into progressively deeper SMs, each featuring a different level of power consumption. When no data need to be transmitted, specific subsets of BS components are progressively deactivated, starting from those featuring the smallest activation delay and the lowest consumption, like the Power Amplifier and some processing components. BS components power consumption is quantified using IMEC Power tool [19] and thereby three SM levels are envisioned with similar time of transition, characterized by the following values of *reactivation time*, i.e., the time required to transit from SMi to the immediately preceding state SM*(i-1)*, or to the active state in case of *i=1*: 35.5 μs (SM1); 0.5 ms (SM2); 5 ms (SM3)[1]. The same values hold for the respective *deactivation time*. As shown in Fig. 2, when the BS enters SMi, a minimum time must elapse before entering SM$(i+1)$, this time is denoted by SMi_{min} and it can be set to arbitrary values. Note that in the figure both power levels and deactivation/reactivation times are not drawn to scale for sake of readability.

When the BS is out of operation during SMs, it retains wake-up functionality even in the deepest SM, since it can be reactivated by the backhaul, that always remains active [4]. If a new data transmission or signaling request arrives during any SM, the BS must be immediately reactivated and the request is temporarily buffered for the time required to transit from the current SM to the fully active (or signaling) state. Conversely, if the new request occurs during the deactivation slope toward SMi, the deactivation process heading to SMi must be completed before starting the reactivation to the fully active state. Requests arriving during a reactivation period are buffered for the remaining time required to fully wake up the BS.

B. Base Station consumption model

The BS power consumption follows the model proposed in [4] and exhibits the following consumption levels:

TABLE I: Power consumption levels.

Active state	Signaling	Idle	SM1	SM2	SM3
490 W	490 W	328 W	157 W	42.9 W	28.5 W

Considering the investigated scenarios, mostly featuring relatively low loads, when the BS is transmitting data a lower power consumption (490 W) is assumed with respect to the full load power (750 W). During an activation/deactivation slope, the power consumption corresponds to the power consumption of the level from which the BS is awaking or from which it is deactivating, respectively. The BS is in Active state during data transmission, whereas the Signaling state corresponds to the transmission of synchronization and control signals (primary and secondary synchronization signals - PSS and SSS, respectively - and physical broadcast channel signals - PBSH) [20]. The Idle state is assumed when the BS is neither transmitting data nor signaling, and the ASM operation is not active.

[1]There exists also a sleep mode 4 that is not considered here, as well as in other related work, because its activation time, equal to 0.5 s, is too long to be compliant with the strict 5G latency constraints

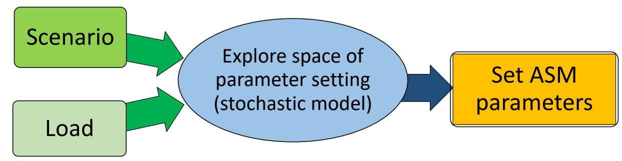


Fig. 3: Sketch of the Tailored SM Operation.

C. ASM in operation

We consider two different operation modes for the ASMs that differ for the setting of the minimum time that the BS must spend in each SM before entering the deeper SM level:

1) **Standard SM Operation (SO)**: the duration of each SM is set according to the standard values proposed in [4], which are set as double the value of the activation time from the corresponding SM; the parameters' setting is the same for the scenarios and whatever the value of the load is.
2) **Tailored SM Operation (TO)**: the duration of the SMs is set based on the scenario and on the expected load, so as to trade-off consumption and delay. As sketched in Fig. 3, the space of possible parameter settings, which is huge, is explored by means of a simple stochastic model that estimates power consumption and delay and whose solution can be very fast.

V. Stochastic model

In this section we propose the stochastic model that can be used in real time to decide, given the scenario and the estimated load, the proper parameters' setting.

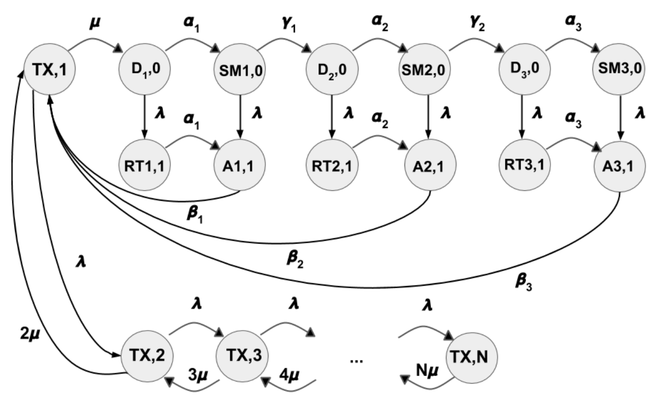


Fig. 4: Markovian model. The values of transition rates are reported in Table II.

A. Model definition

A stochastic model based on a Continuous Time Markov Chains (CTMC) is designed to represent the BS operation. The state variable is defined by:

$$\bar{z} = (S, R)$$

where S indicates the BS state and R represents the number of current requests in the system, either being served or waiting to be served because buffered during reactivation periods. The BS state S can be: Active (A), Signaling (SI), Idle (I), Sleeping (SMi with $i = 1, 2, 3$), or a transition state. There are three types of transition states: i) *deactivation* states D_i (with i=1,2,3) represent the deactivation phases during which the BS enters SMi from the previous SM$(i-1)$ or, if i=1, from the active/signaling state; ii) *activation* states A_i represent the phases of reactivation from SMi, and iii) *residual transition* states, denoted by RT_i, represent special states visited when a new arrival occurs during a deactivation phase (towards SMi) and some residual time is required before the BS completes the entrance to SMi and can start to wake up by entering a reactivation transition state. Note that in our model we represent with exponential random variables even the time spent in those states whose duration is actually deterministic, like the activation/deactivation states and the minimum SM duration before entering a deeper SM level. The average duration of the minimum time spent in each SM before entering a deeper sleep state can either be set to the values defined in the Standard SM Operation, or it can be freely set to any combination of values, like in the Tailored SM Operation, given the absence of any physical constraint that may limit it. In the CTMC, the transition rates are derived based on the mean request inter-arrival time, the deactivation/reactivation times (denoted by t_D and t_A, respectively), the duration of SMi_{min}, and the average service time (denoted by t_s) that depends on the specific considered 5G scenario. Fig. 4 depicts the state transition diagram of the Markov chain. Notice that we neglect the possibility that, while reactivation is in progress, an additional arrival occurs. This simplifies the CTMC and has a negligible impact on the results, since, being the load very low, these events are rare. The values of the transition rates are reported in Table II.

TABLE II: Transition rates.

arrival	λ
service completion	$\mu = t_s{}^{-1}$
deactivation to SM_i	$\alpha_i = t_{D_i}{}^{-1}$
reactivation from SM_1	$\beta_1 = t_{A_1}^{-1}$
reactivation from SM_2	$\beta_2 = (t_{A_1} + t_{A_2})^{-1}$
reactivation from SM_3	$\beta_3 = (t_{A_1} + t_{A_2} + t_{A_3})^{-1}$
residual deactivation from SM_i	$\gamma_i = SMi_{min}^{-1}$

Given the input settings, the model is employed to derive the steady state probability of being in each state under the corresponding parameter configuration.

B. Performance metrics

The metrics to evaluate the system performance so as to decide the parameters setting are reported below.

a) Average reactivation delay

it is the average delay experienced under ASM operation by a request arriving at the system during any SM state.

Let us denote by $\pi(\overline{z})$ the steady state probability of state $\overline{z}$. The average reactivation delay, denoted by $E[d]$, can be derived from the expected first passage time to the active state, i.e. the average time before the active state is visited for the first time, starting from a state corresponding either to deactivation or sleep mode. Since all these states correspond to no request already queued, i.e., $\overline{z} = (S, R)$ with $R = 0$, for the convenience of notation in what follows we will drop the variable R from the indication of the state. $E[d]$ is defined as:

$$E[d] = \frac{\sum_{j=1}^{3} \left[\pi(D_j) \cdot fpt(RT_j, TX) + \pi(SM_j) \cdot fpt(A_j, TX)\right]}{\sum_{j=1}^{3} \left[\pi(D_j) + \pi(SM_j)\right]} \quad (1)$$

where $fpt(RT_j, TX)$ is the first passage time from state RT_j to the transmission state TX, that stands for either active state A or signaling state SI, depending on the considered scenario. Notice that when a request arrives in state D_j the CTMC moves to state RT_j and triggers the reactivation. Similarly, $fpt(A_j, TX)$ represent the first passage time from A_j to TX, and state A_j is entered whenever an arrival occurs in state SM_j.

b) Average power consumption

it is the average power consumed by the BS. It is denoted by $E[P]$ and can be computed as:

$$E[P] = \sum_{\forall \overline{z}} \pi(\overline{z}) \cdot P_{\overline{z}} \quad (2)$$

where $P_{\overline{z}}$ is the power consumption at state $\overline{z}$, according to Table I.

c) Power saving

it is the fraction of power that can be saved under ASM operation with respect to the baseline case in which the BS is always active. The average power saving, denoted by PS, is derived as:

$$PS = \frac{P_B - E[P]}{P_B} \quad (3)$$

where P_B is the average power consumed in the baseline case, which corresponds to a BS that is either fully active when traffic or signaling data are exchanged or it enters idle state when no service requests occur.

VI. Results

We first discuss the impact of SM operation on energy saving and delay in 5G scenarios under standard settings, showing results obtained via simulation. Afterwards, we focus on the TO, according to which the ASM parameters are set so as to properly trade-off reactivation delay and energy consumption and we observe that the TO can significantly improve the performance that can be achieve under standard parameter setting. Finally, we present the validation of the stochastic model that can be used to dynamically adapt in real-time the BS strategy configuration settings to the traffic load variations over time.

A. Standard Operation

We discuss the system performance obtained from simulations having a duration of 6 hours. Fig. 5 reports the average reactivation delay observed in the Signaling (SIG), Massive IoT (MI), and Standard 5G traffic (S5GT) scenarios (featuring

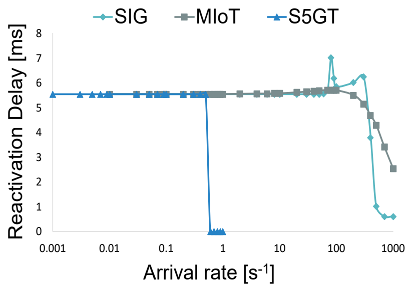


Fig. 5: Reactivation delay for different arrival rates, with proposed 5G scenarios.

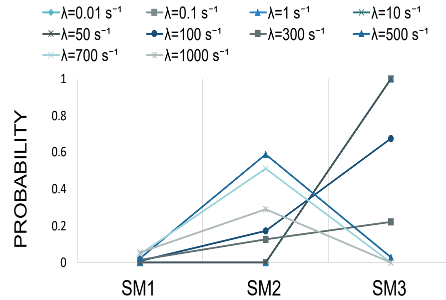


Fig. 6: Probability of arriving during each SMs for different arrival rates, in the SIG scenario.

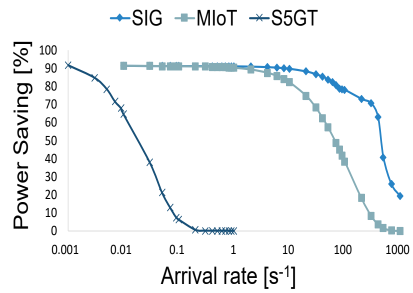


Fig. 7: Energy saving for different arrival rates, with proposed 5G scenarios.

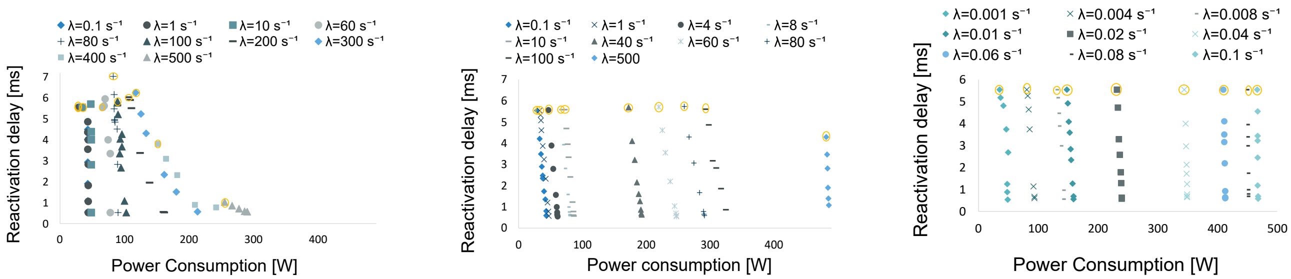
service time 1 ms, 10 ms, and 30 s, respectively) under several request arrival rates in the range [0.01,1000] s^-1 for SIG and MIoT, and [0.001,1] s^-1 for S5GT. For low values of the arrival rate, the delay is about 5.5 ms in all scenarios. In SIG, as λ becomes higher than 60 s^{-1}, the delay increases up to a peak of 7 ms, and then sharply decreases, showing a second (lower) peak at slightly higher arrival rates. Conversely, the delay in MIoT achieves a smaller peak of 5.7 ms for similar λ values. The reason for this behavior is explained in Fig. 6, that shows the probability, denoted by p_{A_i}, that a new arrival occurs either during SMi or during the deactivation toward SMi in the SIG scenario, for different values of λ. For low arrival rates, the probability of a new arrival occurring in SM3 is dominating, leading to high value of the delay. As λ grows larger, we observe an increasing probability of arriving during SM2, a state in which the reactivation delay is lower, and this determines a drop in latency. The delay peak of 7 ms observed in the SIG scenario is motivated by an increased probability of arriving during the transition between SM2 and SM3, that is not observed in the MIoT case (data not reported for the sake of brevity) and that is due to the need for completing the transition from SM2 to SM3 before starting the reactivation. Similar considerations hold for the second peak observed as λ further increases, corresponding to more frequent arrivals during the transition from SM1 to SM2. In S5GT a constant delay is observed at comparatively lower values of λ with respect to SIG and MIoT referring to activation from SM3. A sharp decrease is observed starting from $\lambda = 0.3\ s^{-1}$, with delay rapidly vanishing.

The power saving obtained in the three scenarios is depicted in Fig. 7 for different arrival rates. As λ grows, power saving tends to gradually decrease up to a sharp descent towards no power saving. In MIoT the saving reduction occurs at lower arrival rates than in SIG, due to the higher service time. Indeed, given the same λ, a 10-fold higher service time leads to longer periods in which the BS is active and no energy can be saved. Similarly, the decrease of energy saving in S5GT starts at remarkably lower values of λ, due to the considerably higher service time that makes the BS idle time lower. The SIG curve temporarily reduces its steepness in correspondence of the delay peaks observed in Fig. 5, due to lower consumption during the longer period required to reactivate the BS, as already explained.

B. Tailored SM operation

We now present the results obtained via simulation under the Tailored SM operation (TO), where we improve the trade-off between delay and energy consumption with respect to the standard operation by acting on the ASM parameter settings. In order to compare different parameter settings, since the set of parameters that define the ASM operation is large and the space of possible combinations of values is huge, we show only a subset of possible settings obtained as follows. Given the arrival rate, a range of different values for the minimum time in SM2, namely $SM2_{min}$, is tested combined with the standard value of $SM1_{min}$. The range of possible values of $SM2_{min}$ is defined as $f_S \cdot SM2^*_{min}$, with $SM2^*_{min}$ given by the standard and f_S in a range that depends on the arrival rate and on the scenario. In particular, $f_S \in [1.05, 1.4]$ for the SIG and MIoT scenarios and $f_S \in [3 \cdot 10^3, 3 \cdot 10^6]$ for the S5GT case, resulting in $SM2_{min}$ lasting up to 50 minutes. Obviously, other settings are possible, but these ones show the potential advantages of TO over the standard setting.

Fig. 8 shows the trade-off between the average reactivation delay and the power consumption in SIG, MIoT and S5GT. Each marker corresponds to a different arrival rate; the symbols circled in orange represent the delay and power consumption resulting from the Standard Operation configuration settings. Observe that for each value of λ, the delay can vary remarkably depending on the parameter settings and in most cases can be significantly smaller than the value observed under the standard setting. Conversely, the power consumption variation is mostly negligible, especially in MIoT and S5GT. Only in SIG, for relatively high values of arrival rate, a lower



(a) SIG (b) MIoT (c) S5GT

Fig. 8: Reactivation delay and average power consumption for different combinations of SM2 durations and variable arrival rates, in SIG, MIoT and S5GT.

reactivation delay can be obtained at the price of increasing the power consumption, starting from $\lambda = 200\ s^{-1}$. The presented results highlight how, by properly setting the parameters (in this case, by varying the minimum duration of SM2), the reactivation delay can be dramatically decreased in most cases without significantly affecting the power saving or, at most, at the price of a limited increase of power consumption. Hence, this proves that by properly setting the parameters it is possible to significantly improve the energy-delay trade-off.

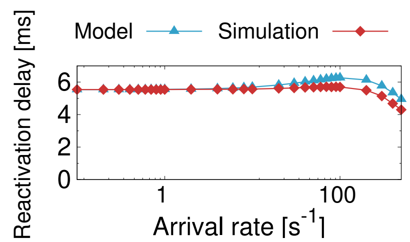
C. Model validation

In order to validate the stochastic model, we compare its results with those obtained by simulation under the same ASM configuration settings. Indeed, in the development of the stochastic model, we introduced the assumption that the deactivation/reactivation times and the minimum time in the SMs are exponentially distributed, while they are constant. With these comparisons we intend to validate the accuracy of the estimations obtained through the model.

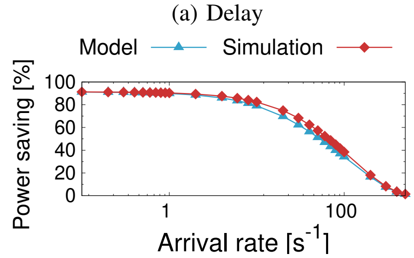
Fig. 9 shows the reactivation delay and the power saving for different arrival rates as derived from the stochastic model (blue curves) and via simulation (red curves) in MIoT scenario. The values of the delay obtained from the model match those registered under simulation for low arrival rates. As λ increases, the model tends to slightly overestimate the delay with respect to the simulation. Similarly, Fig. 9b shows that the power saving curves from the model and from the simulation overlap for low values of λ, whereas, as the arrival rate becomes larger, the power saving predicted by the model tends to only slightly differ with respect to the simulation. Nevertheless, the general trend of delay and power saving from the model is consistent with the simulation results, and the estimation obtained with the model can be used to decide the desired trade-off. Similar results, not shown here for the sake of brevity, are obtained with the SIG and S5GT scenarios.

D. Real-time energy efficient delay modulation

We now examine the capability of the model to reliably predict the delay and power saving when the TO strategy is applied. We consider the MIoT and S5GT scenarios. Fig. 10 depicts, for both the model or the simulation, the reactivation delay versus power consumption for various values of λ, under different parameter settings (obtained, as before, by changing the value



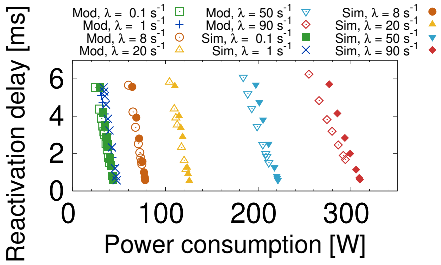
(a) Delay



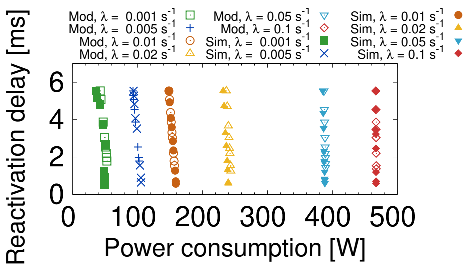
(b) Power saving

Fig. 9: Reactivation delay and average power saving derived from the model and via simulation under standard settings of SM duration, in MIoT.

of $SM2_{min}$). Each point corresponds to a specific configuration setting, with each symbol representing a different arrival rate. Values from the model are represented by empty symbols, whereas solid symbols indicate simulated results. The model is very accurate in estimating the simulated power consumption. Only in MIoT, starting from λ=50 s^{-1}, the power estimated by the model tends to slightly differ from the simulation as the arrival rate increases, whereas the delay shows only very limited deviation. Despite possible small differences between the power consumption and reactivation delay estimated by the model and the values obtained via simulation, the general trend of power consumption and delay is similar in the two approaches. Furthermore, our results show that the proposed stochastic model is quite representative of the BS performance in terms of reactivation delay under TO. Hence, the presented model can be effectively used to dynamically reduce the reactivation delay in typical 5G scenarios, characterized by strict latency constraints. By properly modulating in real time the SM configuration settings as the arrival rate varies over time, it becomes possible to constantly trade off system performance and energy saving. This allows to seamlessly tune



(a) MIoT



(b) S5GT

Fig. 10: Reactivation delay and average power consumption derived from the model (Mod) and via simulation (Sim) under multiple combinations of SM settings (different SM2 durations), in MIoT and S5GT scenarios.

the SM settings during the actual system operation.

VII. Conclusion

The substantial raise of mobile traffic in 5G networks entails a huge growth of the energy demand and related operational cost for MNOs. Approaches based on Advanced Sleep Modes allow to save energy by gradually deactivating the BSs into progressively lower power levels when the traffic is low. This paper investigates the impact of ASM techniques on energy saving and reactivation delay, considering various 5G scenarios that are representative of Massive IoT data transmission, standard 5G traffic exchange, and a scenario envisioning only Signaling transmission. Our results show that the application of ASM based operation strategies under standard settings typically adopted in the literature does not allow to reduce the activation delay in all cases. Parameter setting should be adapted to the scenario and the load. Indeed, by properly setting the parameters that define the time spent at each SM level, the reactivation delay can be remarkably reduced under any (low) load by up to more than 90%, without impairing energy saving.

Finally, a stochastic model, representing the ASM operation of a BS, is proposed and validated. The model can be used to dynamically trade off reactivation delay and power consumption in real time, based on the actual traffic load. Indeed, by being fast to solve, it is possible to use it to explore the parameter setting space in such a short time to take decisions and tune the parameters in real-time. As future work, we are planning to investigate the application of the presented ASM based strategies in more complex scenarios with multiple BSs. In addition, we are planning to verify whether it is feasible to integrate an optimization procedure in the proposed framework for real time operation.

References

[1] Cisco in *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2018-2023 White Paper*, February 2020.

[2] A. Gati, F. E. Salem, A. M. Galindo Serrano, D. Marquet, S. Le Masson, T. Rivera, D.-T. Phan-Huy, Z. Altman, J.-B. Landre, O. Simon, *et al.*, "Key technologies to accelerate the ict green evolution–an operator's point of view," *arXiv preprint arXiv:1903.09627*, 2019.

[3] D. Renga and M. Meo, "Dimensioning renewable energy systems to power mobile networks," *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 2, pp. 366–380, 2019.

[4] F. E. Salem, A. Gati, Z. Altman, and T. Chahed, "Advanced Sleep Modes and Their Impact on Flow-Level Performance of 5G Networks," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, pp. 1–7, Sep. 2017.

[5] J. Wu, Y. Zhang, M. Zukerman, and E. K. Yung, "Energy-efficient base-stations sleep-mode techniques in green cellular networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 17, no. 2, pp. 803–826, 2015.

[6] F. Han, S. Zhao, L. Zhang, and J. Wu, "Survey of strategies for switching off base stations in heterogeneous networks for greener 5g systems," *IEEE Access*, vol. 4, pp. 4959–4973, 2016.

[7] Sriram Prasanth T, Sai Srujan Kumar M, and Shankar T, "A survey on techniques related to base station sleeping in green communication and comp analysis," in *2016 IEEE International Conference on Engineering and Technology (ICETECH)*, pp. 1059–1067, 2016.

[8] F. Elsherif, E. K. P. Chong, and J. Kim, "Energy-efficient base station control framework for 5g cellular networks based on markov decision process," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 9267–9279, 2019.

[9] H. Ye, Z. Ju, B. Wu, J. Pei, and S. Fu, "Joint base station cooperative transmission and on-off mechanism in internet of things networks," in *2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech)*, pp. 336–339, 2017.

[10] Z. Jian, W. Muqing, and Z. Min, "Energy-efficient switching on/off strategies analysis for dense cellular networks with partial conventional base-stations," *IEEE Access*, vol. 8, pp. 9133–9145, 2020.

[11] D. Renga, H. Al Haj Hassan, M. Meo, and L. Nuaymi, "Energy management and base station on/off switching in green mobile networks for offering ancillary services," *IEEE Transactions on Green Communications and Networking*, vol. 2, no. 3, pp. 868–880, 2018.

[12] G. Vallero, D. Renga, M. Meo, and M. A. Marsan, "Greener ran operation through machine learning," *IEEE Transactions on Network and Service Management*, vol. 16, no. 3, pp. 896–908, 2019.

[13] M. Meo, D. Renga, and M. T. Sarti, "Household users cooperation to reduce cost in green mobile networks," in *2018 IEEE International Telecommunications Energy Conference (INTELEC)*, pp. 1–8, 2018.

[14] H. Al Haj Hassan, D. Renga, M. Meo, and L. Nuaymi, "A novel energy model for renewable energy-enabled cellular networks providing ancillary services to the smart grid," *IEEE Transactions on Green Communications and Networking*, vol. 3, no. 2, pp. 381–396, 2019.

[15] "Interuniversity microelectronics centre (imec)." Available at https://www.imec-int.com (2020/08/12).

[16] M. Gruber, O. Blume, D. Ferling, D. Zeller, M. A. Imran, and E. C. Strinati, "EARTH — Energy Aware Radio and Network Technologies," in *2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1–5, 2009.

[17] "On Requirements and Design of SS Burst Set and SS Block Index Indication," in *3GPP TSG-RAN WG1 Meeting #88 R1-1703092*, (Athens, Greece), 2017.

[18] "NR Synchronization Complexity and Periodicity," in *3GPP TSG-RAN WG1 Meeting #88 R1-1702122*, (Athens, Greece), 2017.

[19] B. Debaillie, C. Desset, and F. Louagie, "A flexible and future-proof power model for cellular base stations," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, pp. 1–7, 2015.

[20] E. Dahlman, S. Parkvall, and J. Sköld, "Chapter 5 - nr overview," in *5G NR: the Next Generation Wireless Access Technology* (E. Dahlman, S. Parkvall, and J. Sköld, eds.), pp. 57 – 71, Academic Press, 2018.