

Optimal energy saving for cellular networks Using TACT-Short fading

Selvalakshmi.B¹, M.Ilayaraja²

1PG Student, II-M.E., communication systems.

*2Associate professor, Department of Electronics and communication Engineering
Dhanalakshmi srinivasan engg college, perambalur. (Affiliated to Anna University, Chennai)*

Abstract: Recent works have valid the chance of increase the energy potency in radio access networks (RANs) done by dynamically turning on/off some base stations (BSs). According to this paper, to extend the analysis over Base Stations switching operations that ought to match up with traffic load variations. . we have a tendency to first of all formulate the traffic variations as a Markov Decision Process. After that minimize the energy consumption of RANs, and to design a reinforcement learning framework primarily based BSs switching operation scheme. Furthermore, to speed up the continued learning method, a transfer actor-critic algorithmic program (TACT), that utilizes the transferred learning experience in historical periods or neighbouring regions, is planned and demonstrably converges. The planned considerateness algorithmic program contributes to a performance jumpstart and demonstrates the feasibility of great energy potency improvement at the expense of tolerable delay performance.

Index Terms: Radio access networks, base stations, green communications, energy saving, reinforcement learning, transfer learning, actor-critic algorithm.

I. Introduction:

Wireless cellular networks are growing chop-chop within the previous few decades. The subscriber range and traffic volume in cellular networks have explosively increased. The Base station (BS) transmits common management signals and information signals to mobile users (MUs). Network designing, cell size and capability area unit typically mounted supported the estimation of peak traffic load. For a cellular network in an exceedingly town, the traffic load within the day time is comparatively significant in workplace areas and lightweight in residential areas, whereas the alternative things happen within the evening. The massive range of BSs contribute a significant portion of the energy consumption of cellular networks. once a SB is in its operating mode, the energy consumption of process circuits and cooling system takes up concerning 60percent of the overall consumption the data and communication technology (ICT) business accounts for twenty-four to 100% of the world's overall power consumption presently, over 80% of the ability consumption takes place within the radio access networks (RANs), particularly the bottom stations (BSs). Luca Chiaraviglio et al showed the likelihood of energy saving by simulations. And projected a way to dynamically regulate the operating standing of Bs, counting on the anticipated traffic hundreds. However, to dependably predict the traffic masses remains quite difficult, that makes these works suffering in sensible applications. On the opposite hand, and conferred dynamic Bs switch algorithms with the traffic masses a previous and preliminarily proved the effectiveness of energy saving. Besides, it's conjointly found that turning on/off a number of the BSs can directly have an effect on the associated BS of a mobile terminal (MT). Moreover, decisions of user associations successively result in the traffic load variations of BSs. Hence, any 2 consecutive Bss witch operations are related with one another and current Bs switch operation also will additional influence the general energy consumption within the long haul. In alternative words, the expected energy saving theme should be farsighted whereas minimizing the energy consumption. It ought to concern its result on each the present and future system performance to deliver a visionary BSs witch operation answer. The authors in conferred a partly farsighted energy saving theme which mixes Bs switch operation and user association, by giving a heuristic answer on the premise of a stationary traffic load profile. During this paper, we tend to try and solve this drawback from a unique perspective rather than predicting the quantity of traffic masses, we tend to apply a Markov call method to model the traffic load variations. Afterwards, the answer to the developed MDP drawback is earned by creating use of actor-critic algorithmic program a reinforcement learning (RL) approach actor-critic algorithm, a reinforcement learning (RL) approach, one advantage of that is that there's no necessity to possess a previous information temporal and spacial connection within the traffic hundreds and hurrying the on-going learning method in regions of interest as delineated in Fig one. As a result, the training frame work of Bs shift operation is more increased by in cooperating the concept of TL into the classical actor-critic rule namely the Transfer Actor-Critic rule during

this paper. Firstly, we have a tendency to show that the training framework is possible to save lots of the energy consumption in RANs while not the information of traffic hundreds a previous. Moreover, the performance of the training framework approaches that of the progressive theme (SOTA) that is assumed to own totally information of traffic hundreds. These preliminary results have already been bestowed in second, we have a tendency to extend the concept of TL to the traditional RL rules and show that the planned tactfulness rule outperforms the classical AC algorithm with a performance jumpstart. Thirdly, this paper details the convergence analysis of the tactfulness rule and thereby contributes to the final literature in RL field, particularly the final AC rule.

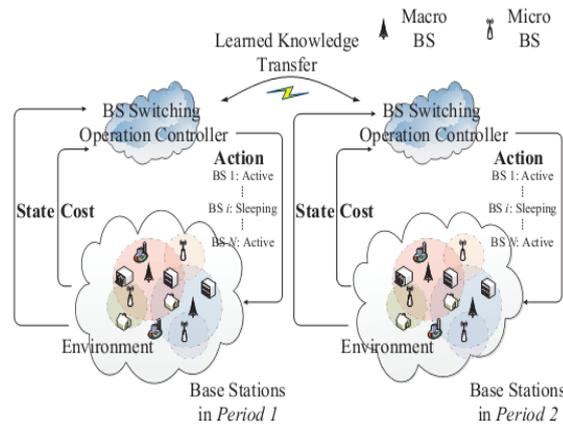
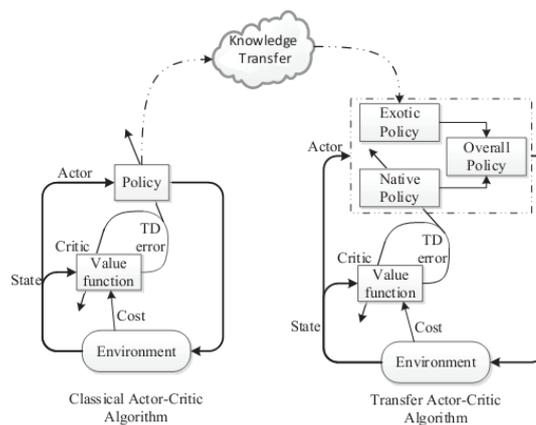


Fig. 1. Transfer learning for reinforcement learning in BS switching operation scenario.

II. Stochastic Bs Switching Operation In Reinforcement Learning Framework

A. Markov decision process

Markov decision processes (MDPs), named once Andrey Markov, offer a mathematical framework for modeling higher cognitive process in things wherever outcomes are partially random and partially underneath the management of a choice maker. MDPs are helpful for learning a large vary of optimization issues resolved via dynamic programming and reinforcement learning. More exactly, a Andre Markov decision method could be a distinct time random method. At when step, the method is in some state, and also the top dog might opt for any action that's on the market in state. The method responds at following time step by every which way getting in a brand new state, and giving the choice maker a corresponding reward.



B. The actor-critic learning framework for energy saving

There are some well-known strategies to unravel the MDP problems like dynamic programming .sadly, these strategies heavily rely upon previous information of the environmental dynamics. However, it's difficult to grasp the long run traffic hundreds exactly ahead. Therefore, during this paper, we tend to use reinforcement learning approaches to unravel the MDP drawback while not requiring the information of traffic hundreds a previous and specifically adopt the actor-critic algorithmic rule. Because the name implies, the actor-critic algorithmic rule encompasses 3components: actor, critic, and surroundings. At a given state, the actor

selects AN action in an exceedingly random means then executes it. This execution transforms the state of surroundings to a replacement one with bound chance, and feeds back the value to the actor. Then, the critic criticizes the action dead by the actor and updates the worth perform through a time distinction (TD) error. when the criticism, the actor can update the policy to like the action with a smaller value, and the other way around. The algorithmic rule repeats the higher than procedure till convergence. the explanations to adopt actor-critic algorithmic rule are three-folded: (i) since it generates the action directly from the hold on policy, it needs very little computation to pick out Associate in Nursing action to perform; (ii) it will learn Associate in Nursing expressly random policy which can be helpful in non-Markov traffic variation setting of RANs (iii) it one by one updates the worth perform and policy. As a result, it'd be additional simply to implement the policy information transfer in Section IV, compared to alternative critic-only algorithms like ϵ -greed and Q-learning. We style Associate in actor-critic learning framework for energy saving theme as illustrated in Fig. 3.

(i) Action selection: Beforehand, let's assume that the system is at the start of stage k . Meanwhile, the traffic load state is $s(k)$. Thereafter, the controller must choose Associate in Nursing action consistent with a random strategy, the aim of that is to enhance performance whereas expressly equalisation 2 competitory objectives: a) finding out a far better BS change operation (exploration) and b) taking as very little price as doable (exploitation). As a result, the controller not solely performs an honest BS change operation supported its past expertise.

III. Transfer Actor –Critic Algorithm For Stochastic Bs Switching operation:

A. Motivation and formulation of transfer actor-critic algorithm

The previous section addresses the methodology to take advantage of the classical AC rule to conduct the Bs switch operation, culminating in a good energy saving strategy within the finish. During this section, we present the implies that the controller utilizes the information of learned ways throughout historical periods or neighboring regions to be within the groove of finding the optimum Bs switch operations. Basically, the policy, say $p(s,a)$, that finally determines the strategy $\pi(s, a)$ in one learning task, indicates the tendency of action a to be chosen in state s . once the training method converges, the tendency to decide on a particular action a during a specific state is relatively larger than that of alternative actions. In alternative words, it implies that if the Bs switch operation is conducted supported one learned strategy, the energy saving within the whole system tends to be optimized within the end of the day. Hence, if the information of this policy $p(s, a)$ is transferred to a different task, the information transferred from amount one to amount two (target task) among a similar region of interest in Fig. 1, the controller within the target task will create an endeavor by taking a similar action a once the traffic masses get state s .

Block diagram:

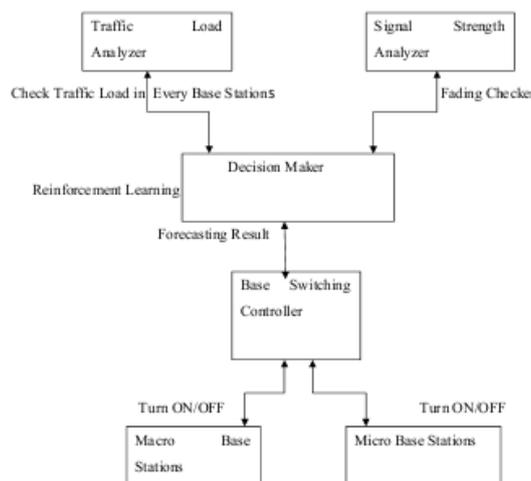


Fig:4.1 Architecture of classical transfer actor critic algorithm

Compared to learning from the scratch, the controller would possibly directly create the wisest alternative at the terribly starting. However, in spite of the similarities between the supply task and also the target task, there would possibly still exist some variations. for instance, the system would possibly get a similar state in 2 totally different tasks, whereas the traffic masses within the supply task can be sometimes beyond that within the target one Hence, rather than staying on the chosen action a in supply task, the controller in target task will create a additional aggressive option to flip additional BSs into sleeping

mode, therefore saving additional energy consumption. Consequently, during this case, the transferred policy guides during a negative manner. During this section, we present the means the controller utilizes the information of learned methods throughout historical periods or neighboring regions to be within the groove of finding the optimum BS switch operations. The controller has tried to settle on this action and nurtured its own learning expertise. Taking the higher than issues under consideration, we tend to propose a replacement policy update technique, named Transferred Actor-Critic algorithmic rule (TACT) as Fig. 2. Within the tact algorithmic rule, the policy (i.e., p_0) to pick an action is split as a native one p_n and an exotic one p_e . while not loss of generality, let's assume that at stage k , the traffic load state is $s(k)$ and also the chosen action could be $a(k)$. Consequently, the policy p_0

$$p_0^{(k+1)}(s^{(k)}, a^{(k)}) = \left[(1 - \zeta(\nu_2(s^{(k)}, a^{(k)}, k))) p_n^{(k+1)}(s^{(k)}, a^{(k)}) + \zeta(\nu_2(s^{(k)}, a^{(k)}, k)) p_e(s^{(k)}, a^{(k)}) \right]_{-p_1}^{p_1}$$

is updated as

Besides that, $p_n(s, a)$ still updates itself per the classical actor-critic formula, particularly(11).Initially, the exotic policy $p_e(s, a)$ dominates within the overall strategy. Hence, once the atmosphere enters a state s , the presence of $p_e(s, a)$ contributes to settle on the action, which could be best to s within the supply task. Consequently, the projected policy update methodology results in a potential performance jumpstart. On the opposite hand, since $\zeta(\bullet) \in (0, 1)$ is that the transfer rate and $\zeta(k) \rightarrow 0$ as $k \rightarrow \infty$, the impact of the transferred exotic policy $p_e(s,a)$ unceasingly decreases. Therefore, the controller can't solely benefit of the learned experience within the supply task, however conjointly fleetly get obviate the negative tips Besides that, $p_n(s, a)$ still updates itself per the classical actor-critic formula, particularly(11).Initially, the exotic policy $p_e(s, a)$ dominates within the overall strategy. Hence, once the atmosphere enters a state s , the presence of $p_e(s, a)$ contributes to settle on the action, which could be best to s within the supply task. Consequently, the projected policy update methodology results in a potential performance jumpstart. On the opposite hand, since $\zeta(\bullet) \in (0, 1)$ is that the transfer rate and $\zeta(k) \rightarrow 0$ as $k \rightarrow \infty$, the impact of the transferred exotic policy $p_e(s,a)$ unceasingly decreases. Therefore, the controller can't solely benefit of the learned experience within the supply task, however conjointly fleetly get obviate the negative tips 1fig

5.9 OUTPUT:

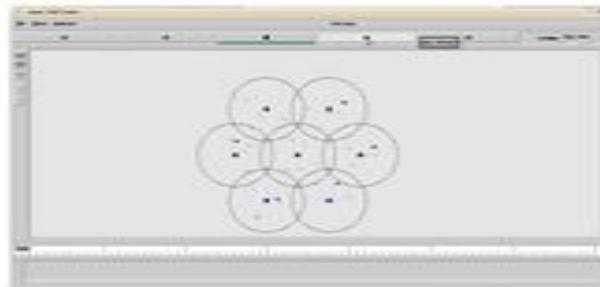


Fig.5.10 all the base stations working in low traffic



Fig.5.11 base station switched to one with low traffic

Fig:1 output for BS switching using tact algorithm.

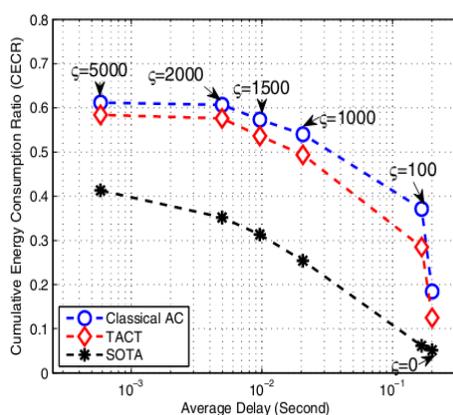


Fig.2 . Performance tradeoff between energy and delay under different delay equivalent cost scenarios.

of consideration theme over classical AC theme versus Kull back-Leibler divergence. The bars appreciate the left coordinate axis mirror the CECR improvement whereas the line corresponding the correct coordinate axis represents the Kull back-Leibler divergence

IV. Conclusion

In this paper, we've developed a learning framework for Bs energy saving. We have a tendency to specifically developed the Bs shift operations beneath variable traffic hundreds as a Markov call method. Besides, we have a tendency to adopt the actor-critic technique, a reinforcement learning rule, to present the Bs shift resolution to decrease the general energy consumption. After-wards, to totally exploit the temporal relevance in traffic hundreds, we have a tendency to propose a transfer actor-critic rule to boost the ways by taking advantage of learned data from historical periods. Our projected rule demonstrably converges given sure restrictions that arise throughout the training method, and therefore the in depth simulation results manifest the effectiveness and lustiness of our energy saving schemes beneath numerous sensible configurations. Just like the simulated temporal data transfer, our projected consideration approach is probably viable to be applied in special situations to realize a performance improvement. Sadly, he mapping of data are going to be typically less simple within the latter case, attributable to the underlying Bs geographical preparation variations. Therefore, we have a tendency to square measure dedicated to handle the connected pregnant however tougher problems over spacial data transfer within the future.

References

- [1]. J. Wu, S. Rangan, and H. Zhang, Eds., Green Communications: Theoretical Fundamentals, Algorithms and Applications ,1st ed. CRC Press,2012.
- [2]. H. Zhang, A. Gladisch, M. Pickavet, Z.Tao, and W. Mohr, "Energy efficiency in communications," IEEE Commun.Mag. , vol. 48, no. 11,pp. 48–49, Nov. 2010.
- [3]. M.Marsan, L.Chiaraviglio, D. Ciullo, and M.Meo, "Optimal energy savings in cellular networks," in Proc. 2009 IEEE ICC Workshops.
- [4]. China Mobile Research Institute, "C-RAN: road towards green radio access network," Tech. Rep., 2010.
- [5]. G. P.Fettweis and E. Zimmermann, "ICT energy consumption-trends and challenges," in Proc. 2008 WPMC ,vol.4.
- [6]. K. Son, H. Kim, Y. Yi, and B. Krishnamachari, "Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks," IEEE J. Sel. Areas Commun. , vol. 29, no. 8, pp. 1525–1536, Sept. 2011.
- [7]. C. Peng, S.-B. Lee, S. Lu, H. Luo, and H. Li, "Traffic-driven power savings in operational 3G cellular networks," in Proc. 2011 ACM Mobicom.
- [8]. Z. Niu, "TANGO: traffic-aware network planning and green operation," IEEE Wireless Commun., vol. 18, no. 5, pp. 25–29, Oct. 2011.
- [9]. Chiaraviglio,D.Ciullo,M.Meo,M.Marsan,andI.Torino,"Energy-aware UMTS access networks," in Proc. 2008
- [10]. Z. Niu, Y. Wu, J. Gong, and Z. Yang, "Cell zooming for cost-efficient green cellular networks," IEEE Commun. Mag., vol. 48, no. 11, pp. 74–79, Nov. 2010.
- [11]. R. Li, Z. Zhao, Y. Wei, X. Zhou, and H. Zhang, "GM-PAB: a grid-based energy saving scheme with predicted traffic load guidance for cellular networks," in Proc. 2012 IEEE ICC.
- [12]. E. Oh and B.Krishnamachari, "Energy savings through dynamic base station switching in cellular wireless access networks," in Proc. 2010 IEEE Globecom.
- [13]. S. Zhou, J. Gong, Z. Yang, Z. Niu, and P. Yang, "Green mobile access network with dynamic base station energy saving," in Proc. 2009 ACM Mobicom.
- [14]. V. Konda and V. Borkar, "Actor-critic-type learning algorithms for Markov decision processes," SIAM J. Contr. Optim., vol. 38, no. 1, pp.94–123, 1999.

- [15]. V. Konda and J. Tsitsiklis, "Actor-critic algorithms," *SIAM J. Contr. Optim.* , vol. 42, no. 4, pp. 1143–1166, 2000.
- [16]. R. Sutton and A. Barto, *Reinforcement Learning: An Introduction*.Cam-bridge University Press, 1998. Available: <http://webdocs.cs.ualberta.ca/~sutton/book/ebook/>
- [17]. H. Berenji and D. Vengerov, "A convergent actor-critic-based FRL algorithm with application to power management of wireless transmitters," *IEEE Trans. Fuzzy Syst.* , vol. 11, no. 4, pp. 478–485, Aug. 2003.
- [18]. F. Woergoetter and B. Porr, "Reinforcement learning," vol. 3, no. 3, p.1448, 2008.
- [19]. X. Zhou, Z. Zhao, R. Li, Y. Zhou, and H. Zhang, "The predictability of cellular networks traffic," in *Proc. 2012 IEEE ISCT*.