Learning Annealing Schedule of Log-Linear Algorithms for Load Balancing in HetNets

Mohd. Shabbir Ali, Pierre Coucheney, and Marceau Coupechoux

***Abstract*—Load balancing among the base stations in heterogeneous networks (HetNets) is essential for their successful deployment. In this paper, we present a robust approach for load balancing by adapting log-linear learning algorithms (LLLA). A new distributed annealing learning algorithm (ALA) is proposed to learn the parameter of LLLA by adapting successive reject algorithm. ALA gives a new annealing schedule that describes the evolution of parameter τ of LLLA over a fixed horizon. The performance of this new annealing schedule is compared with commonly used annealing schedules in the literature such as linear decreasing, log decreasing, and fixed parameter. It is observed from simulations that the new annealing schedule achieves lower global cost for a fixed time horizon compared to that of other annealing schedules. For lower time horizons, ALA with linearly decreasing τ is better than ALA with a fixed vector of τ. Whereas, for higher time horizons, ALA performance is same in both the cases. Finally, we show the applicability of the proposed algorithm for load balancing.**

I. INTRODUCTION

Conventional homogeneous cellular networks are drifting towards heterogeneous networks (HetNets) due to traffic demand based deployment of small power base stations (BSs). The idea is to share the load of macro base stations with the small base stations. However, with the conventional user association technique where the user selects a BS that provides the highest received power, more users associate with macro BSs. This leads to higher load at macro BSs and at the same time under-utilisation of small BSs resources. Therefore, a critical problem in HetNets is to associate users to BSs such that the load between the BSs is balanced and small BSs resources are utilised efficiently.

An overview of load balancing approaches in the literature can be found in [1], [2]. These can be broadly classified as centralised, e.g. in [3]–[6], and decentralised optimisation approaches, see e.g. [7]–[11]. User-centric game theoretical and learning approaches are becoming popular because they provide distributive solutions, e.g. using congestion games [8], [10], evolutionary games [9] and distributed Q-learning [11].

Recently, log-linear learning algorithm (LLLA) has been proposed to achieve global optimum performance for load balancing [12]. Load balancing is modeled as a potential game whose potential function is an α-fair utility function that captures network performance. LLLA and binary LLLA (BLLLA) are proposed for complete and partial information settings, respectively. In [12], cell range expansion (CRE) technique is used where users associate with a BS that provides the highest biased received power. A CRE bias is broadcast by every BS and is typically higher for small BSs than for macro BSs. This results in an increase of the small cell coverage and thereby of the number of users associated with them.

Also in [12], the effect of parameter τ of a log-linear algorithm is presented. It is observed that with lower τ the algorithm may be stuck at a local minimum and higher τ may result in unacceptable oscillations. LLLA converges in probability to the global minimum when $\tau(t) = \frac{\kappa}{\log(1+t)}$, where κ is the depth of the deepest local minimum of potential function [13]. The value of κ is generally unknown before hand, especially in a distributed setting. Other heuristic annealing schedules such as linear decreasing τ and fixed τ are used in the literature [14]. However, if the initial τ is very high then this annealing also takes a long time for convergence, which is not suitable for load balancing. Therefore, a better annealing schedule is required that is fast and guarantees asymptotic convergence. This is the focus of this paper. We propose a new distributed annealing learning algorithm (ALA) to learn the annealing schedule of the parameter of LLLA by adapting successive reject algorithm (SRA) [15]. This annealing schedule is fast, performs better than other annealing schedules, and guarantees asymptotic convergence.

The main contributions of this paper are as follows:

Robust approach for load balancing: We present a distributed annealing learning algorithm that gives a new annealing schedule of the τ parameter of LLLA. The new annealing schedule makes LLLA robust towards the effect of different τ parameters. It makes LLLA more suitable for load balancing in HetNets that use CRE for user association.

Numerical results: We first, present a numerical example to show the behaviour of LLLA with different τ. Then, we present the performance of ALA annealing schedule, log decreasing τ, linear decreasing τ, and with the smallest τ. Then, we show that ALA outperforms other annealing schedules for a given time horizon. For lower time horizons, ALA with linearly decreasing τ is better than ALA with fixed vector of τ. Whereas, for higher time horizons, ALA performance is same in both the cases. Finally, we show the simulations of the proposed ALA for load balancing.

This paper is organised as follows. In Section II, the system model is described and the problem is formulated. In Section III, learning in a potential game framework is presented. The ALA is described in Section III. In Section IV, our approach is validated using extensive simulations. Finally, the conclusions are given in Section V.

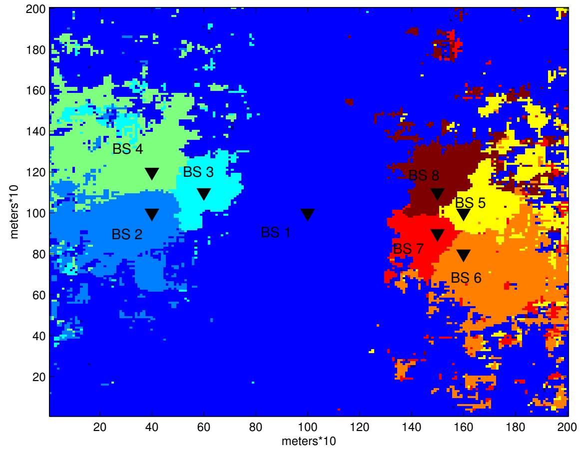


Fig. 1: Optimal coverage regions obtained using optimal CRE biases under correlated shadow fading.

II. System Model

We consider an extended system model of [12] by taking into account the auto-correlated and cross-correlated shadow fading as shown in Fig. 1. In this figure, BS i is a macro BS and others are small BSs. The users at location x arrive with a rate $\lambda(x)$ [arrivals/s/m^2] and with an average demand of $1/\mu(x)$ [bits]. The coverage regions of BSs are shown with different colours. The downlink of a cellular network (typically an LTE-Advanced network) is considered. The sets $\mathcal{B}_e$ and $\mathcal{B}_s$ denote the set of macro BSs and set of small BSs, respectively. The set of all stations is denoted $\mathcal{S} \triangleq \mathcal{B}_e \cup \mathcal{B}_s$. Every BS i maintains a CRE bias parameter $c_i \in [1, \ldots, c_{\max}]$. The CRE bias vector is denoted $\bar{\boldsymbol{c}} = [c_1, c_2, \ldots, c_{|\mathcal{S}|}]$.

1) Channel Model: The received power at location x from BS i is $P_i g_i(x)$, where P_i is the transmit power and $g_i(x)$ is the channel gain, which captures the effect of path-loss and shadowing. The effect of small-scale fading is not considered because the time for user association procedure is assumed to be much larger than the channel coherence time [7]. We consider a scenario where the locations of the BSs and of the users during their download are fixed. Therefore, the shadow fading component is a constant multiplicative factor. Formally, the channel gain model considered is [16]:

$$g_i(x) = \min\left\{1, K\left|x - x_i\right|^{-\eta} e^{\beta y_i(x)}\right\}, \tag{1}$$

where $K = \left(\frac{\lambda_w}{4\pi d_0}\right)^2$, λ_w is the wavelength, d_0 is the reference distance, x_i is the location of the BS i, $\eta \geq 2$ is the path-loss exponent, and $e^{\beta y_i(x)}$ is the shadowing component where $\beta = \frac{\log 10}{10}$ and $y_i(x)$ is a realisation of Gaussian random process of zero mean and covariance function $C_{y_i}(\Delta x)$ [17]:

$$C_{y_i}(\Delta x) = \sigma_{sh}^2 e^{-\frac{\Delta x}{D_c}}, \tag{2}$$

where σ_{sh}^2 is the variance, Δx is the displacement, and D_c is the decorrelation distance [16]. A constant cross correlation between the $y_i(x)$ and $y_j(x)$ is considered as in [18].

The SINR $\gamma_i(x)$ of a user at location x is given as:

$$\gamma_i(x) = \frac{P_i g_i(x)}{\sum_{j\in\mathcal{S}} P_j g_j(x) + N_0}, \tag{3}$$

where $N_0 = -174 + 10\log W$ is thermal noise power in dBm and W is system bandwidth in Hz.

2) CRE User Association Rule: A user association rule based on CRE and maximum transmit power is commonly used in the HetNets [1], [19]–[28]. According to this rule, a user located at x is associated to the BS i that provides the highest biased received power. The set of locations $\mathcal{D}_i(\bar{\boldsymbol{c}})$ associated to BS i is defined as:

$$\mathcal{D}_i(\bar{\boldsymbol{c}}) = \left\{x \middle| \forall j \in \mathcal{S}, P_i g_i(x) c_i \geq P_j g_j(x) c_j\right\}. \tag{4}$$

3) Traffic Model: BSs are modeled as M/G/1/PS queues as in [7], [12]. The load of BS i is given below.

$$\rho_i(\bar{\boldsymbol{c}}) = \int_{\mathcal{D}_i(\bar{\boldsymbol{c}})} \frac{\lambda(x)}{\mu(x)\nu_i(x)} 1_{\{\gamma_i(x)\geq\gamma_{\min}\}} dx, \tag{5}$$

where the data rate provided is $\nu_i(x)$ [bits/s]. A user is served only if its SINR exceeds a minimum threshold $\gamma_{\min}$. BS i is stable if and only if $0 \leq \rho_i < 1$. In this work, only stable network configurations are considered. The load vector is denoted as $\boldsymbol{\rho} = \left[\rho_1, \rho_2, \ldots, \rho_{|\mathcal{S}|}\right]$.

A. Problem Formulation and Objective Function

The goal is to minimise an $\alpha-$fairness function $\phi_\alpha(\bar{\boldsymbol{c}})$ over a feasible set $\mathcal{F}$, which are defined as [7], [12]:

$$\phi_\alpha(\bar{\boldsymbol{c}}) = \begin{cases} \sum_{i\in\mathcal{S}} \frac{(1-\rho_i(\bar{\boldsymbol{c}}))^{1-\alpha}}{\alpha-1}, & \alpha \geq 0, \alpha \neq 1, \\ -\sum_{i\in\mathcal{S}} \log\left(1 - \rho_i(\bar{\boldsymbol{c}})\right), & \alpha = 1, \end{cases} \tag{6}$$

$$\mathcal{F} = \left\{\bar{\boldsymbol{c}} \middle| 0 \leq \rho_i(\bar{\boldsymbol{c}}) < 1, c_i \in [1, c_{\max}], \forall i \in \mathcal{S}\right\}. \tag{7}$$

The function $\phi_\alpha(\bar{\boldsymbol{c}})$ is in general non-convex and even if it is convex the set $\mathcal{F}$ is non-convex because $\bar{c}$ takes discrete values. The function $\phi_\alpha(\bar{\boldsymbol{c}})$ captures various aspects of fairness and performance for the network depending on the choice of α. Minimising $\phi_0(\bar{\boldsymbol{c}})$ results in a rate-optimal policy. Minimising $\phi_1(\bar{\boldsymbol{c}})$ is equivalent to achieving proportional fairness between BSs [29]. Minimising $\phi_2(\bar{\boldsymbol{c}})$ is equivalent to minimising the average delay of the network. For more detailed discussion refer to [7]. As $\alpha \to \infty$ the minimiser of $\phi_\alpha(\bar{\boldsymbol{c}})$ tends to the min-max load vector. It is a standard result with convex objective function [7], [29], [30]. The proof for the non-convex objective function is given in [12].

III. Learning Annealing Schedule of LLLA

In this section, we first present a potential game model for user association and then describe LLLA. Next, we present the learning algorithm for annealing schedule of LLLA.

Definition 1: [User Association Game] It is defined by the tuple $\Gamma = \left\{\mathcal{S}, \{X_i\}_{i\in\mathcal{S}}, \{U_i\}_{i\in\mathcal{S}}\right\}$, $\mathcal{S}$ is a set of BSs, X_i is a set of strategies of BS i. Strategy set X_i is a discrete set of CRE bias values. Let denote a_i and a_{-i} as the strategy of player i and strategies of players except i, respectively. The cost function U_i is given below:

$$U_i(a_i, a_{-i}) = \sum_{j \in N_i} \frac{(1 - \rho_j(a_i, a_{-i}))^{1-\alpha}}{\alpha - 1}, \tag{8}$$

where N_i is the neighborhood of BS i and $\rho_j(a_i, a_{-i})$ is the load of BS j given in (5).

It is proved in [12] that the game in Definition 1 is an exact potential game with the potential function ϕ_α in (6). An exact potential game has, at least, one pure Nash equilibrium (PNE) and local optimizers of the potential function are PNEs [31]. A PNE of a game is reached when no player can benefit by changing its strategy unilaterally.

A. *Learning Algorithm for Annealing Schedule of LLLA*

Algorithm 1 Log-linear Learning Algorithm

Start with arbitrary action profile a.
while $t \geq 1$ **do**
  Set parameter τ as a function of time t.
  Randomly select a BS i.
  Compute cost $U_i\left(a_i, a_{-i}(t-1)\right)$ for all $a_i \in X_i$.
  For any $a_i \in X_i$, set $U_i\left(a_i, a_{-i}(t-1)\right) = \infty$ if $\rho_j \geq 1$ for $j \in N_i$.
  Take action $a_i(t)$ from X_i with probability $p_i^{a_i}(t)$,

$$p_i^{a_i}(t) = \frac{\exp\left(-\frac{1}{\tau} U_i\left(a_i, a_{-i}(t-1)\right)\right)}{\sum_{a_i' \in X_i} \exp\left(-\frac{1}{\tau} U_i\left(a_i', a_{-i}(t-1)\right)\right)}. \tag{9}$$

  All the other BSs must repeat their previous actions, i.e., $a_{-i}(t) = a_{-i}(t-1)$.
end while

Algorithm 2 Annealing Learning Algorithm

Initialisation: Let $A_1 = \{\tau_1, \ldots, \tau_M\}$, $\Upsilon = \frac{1}{2} + \sum_{i=2}^{M} \frac{1}{i}$.
for Phase $k = 1, 2, \ldots, M-1$ **do**

$$n_k = \lceil \frac{1}{\Upsilon} \frac{T - M}{M + 1 - k} \rceil \tag{10}$$

  and $n_0 = 0$.
  for each $i \in A_k$ **do**
    play LLLA with τ_i for $n_k - n_{k-1}$ iterations.
  end for
  Compute $A_{k+1} = A_k \setminus \arg\max_{i \in A_k} \hat{X}_{i,n_k}$.
end for

LLLA is summarised in Algorithm 1. This algorithm is similar to best response (BR) algorithm but allows deviations from the best response with a probability, which is governed by the parameter τ. It guarantees the asymptotic convergence to the optimal NE of an exact potential game [12], [32]. This means that, asymptotically, the probability that the algorithm is at the global minimum approaches to one as τ goes to zero. The effects of τ parameter on LLLA is shown using simulations in Fig. 2. For high values of τ, LLLA results into oscillations. This is due to the fact that the algorithms converge fastly in probability to the uniform distribution. As a matter of fact, it doesn't spend much time in optimal states, which is not practically desirable. For small values of τ, asymptotically, the algorithms will spend most of the time in the global optimal. However, convergence is slow in probability. This explains that the system can take a long time to escape from sub-optimal states. Contrary to best response, however, the LLLA does not get stuck into sub-optimal states.

The annealing schedule describes the evolution of parameter τ with some guarantee of convergence of LLLA. A logarithmic decreasing τ theoretically guarantees the convergence to the global minimum [13]. However, this annealing schedule is very slow and hence not practical for load balancing. A faster linear decreasing τ annealing is usually used in the literature. However, if the initial τ is very high then this annealing also takes a long time for convergence, which is not suitable for load balancing. Usually, a fixed value of τ is used for LLLA as a heuristic in the literature. However, choosing an appropriate value of τ itself is a challenge. Also, LLLA with small τ may be stuck into local minima for a long time. An annealing schedule with quick convergence with some guarantee is desirable. We propose to learn the annealing schedule for a fixed time horizon. The idea is to explore the given values of τ and reject a τ value whose average cost is the worst compared to other τ values. At the end of the time horizon, the surviving τ value is the best value to run LLLA. The details of annealing learning algorithm are described in Algorithm 2.

ALA is a successive reject algorithm originally proposed for the best arm identification of multi-arm bandit problem [15]. We adapt SRA for LLLA to identify the best τ. The multiple τ parameters are multiple arms of SRA. Every BS runs this algorithm distributively to obtain its own best τ to run LLLA. Consider a M number of τ parameters of a wide range. First, the algorithm divides the given time horizon (i.e., the T iterations) in $M-1$ phases. At the end of each phase, the algorithm dismisses the τ_i parameter with the highest empirical mean cost, $\hat{X}_{i,s} = \frac{1}{s}\sum_{t=1}^{s} X_{i,t}$, where $X_{i,t}$ is the cost obtained using (8) with τ_i at time t. During the next phase, it plays equally often each τ which has not been dismissed yet. The recommended τ is the last surviving τ. The length of the phases are carefully chosen to obtain an optimal convergence rate for finite horizon as in multi-arm bandit setting [15].

The process of dismissing undesirable τ parameters of ALA provides an annealing schedule for LLLA. The performance of this annealing schedule can be evaluated by computing the probability of improvement achieved compared to other annealing schedules. Improvement is measured in terms of the achieved potential value from different annealing schedules. Note that the ALA also guarantees asymptotic convergence to the optimal NE of the user association game because at the end of time horizon ALA runs LLLA with a single value of τ same as in the classical LLLA [12].

TABLE I: Simulation parameters.

Parameter	Variable	Value
Number of BSs	N_s	8
Transmit power of macro BS	P_{macro}	46 dBm
Transmit power of small BS	P_{small}	24 dBm
Average file size	$\frac{1}{\mu}$	0.5 Mbytes
Average traffic load density	$\frac{\lambda}{\mu}$	64 bits$/s/m^2$
System bandwidth	W	20 MHz
Carrier frequency	f_c	2.6 GHz
Noise power	N	-174+10log(W) dBm
Minimum SINR	$\gamma_{\min}$	-10 dB
Path-loss exponent	η	3.5
Reference distance	d_0	10 m
CRE bias set	c_i	$\{1, 2, \ldots, 16\}$

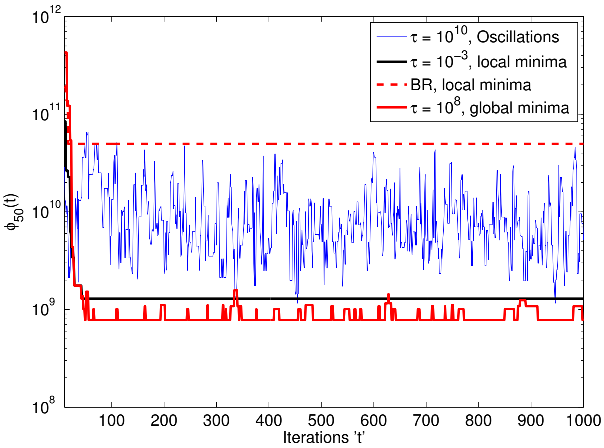


Fig. 2: Illustration of effect of different fixed τ.

IV. Simulation Results

In this section, we show simulation results of the proposed algorithm considering standard parameters as adopted in 3GPP [33]. These parameters are listed in the Table I. We consider 8 BSs located in a two-dimensional region $\mathcal{L}$. BS 1 is a macro BS that transmits with P_{macro} and the rest are small BSs that transmit with P_{small}. The user traffic varies with location across an average traffic density of 64 bits$/s/m^2$. There are two hotspots where the traffic is 5 times the average traffic. We consider shadow fading with a standard deviation of $\sigma_{sh} = 8$ dB and a decorrelation distance of $D_c = 20$ m. Cross-correlation between the shadowing components at a location is considered to be 0.5. We use the following classical Shannon formula for calculating channel capacity $\nu_i(x)$ at any location x:

$$\nu_i(x) = W \log_2 (1 + \gamma_i(x)) . \tag{11}$$

Fig. 2 shows the effect of different fixed τ on the convergence of LLLA. Oscillations are obtained for higher values of $\tau = 10^{10}$. Whereas, for lower values of $\tau = 10^{-3}$ local minima is obtained. The best response (BR) algorithms get trapped into local minima. The global minimum is obtained for $\tau = 10^8$. In practice, finding a value of τ that gives the desired performance is very challenging.

The performance of ALA with $\tau = [10^{12}, 10^{10}, 10^8, 10^2]$ is shown in Fig. 3a. In the first few iterations, a lot of oscillations are observed. This is because of the high values of τ used by ALA to explore through all the given τ parameters. As the algorithm progresses higher values of τ are rejected and the algorithm progresses towards stabilisation. At the end, the ALA runs LLLA with the surviving τ and it converges.

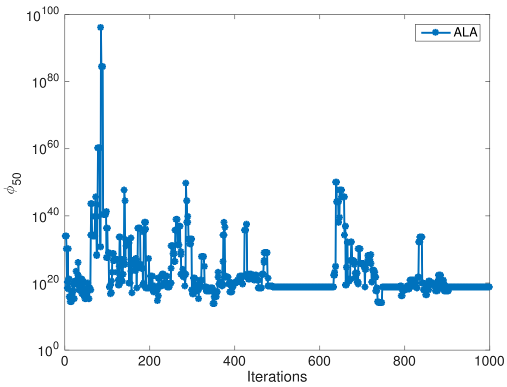
The evolution of potential function obtained from LLLA with the given smallest $\tau = 100$ is shown in Fig. 3b. LLLA converges fast to some local minima and gets stuck into it most of the time. Since, the time horizon is finite the behaviour of LLLA with small τ is same as BR because it does not have enough time to escape from the local minimum.

The performance of a linear decreasing annealing schedule is shown in Fig. 3c. Note that the initial value of τ is the highest value of the given set of τ parameters. It performs poor and results in a lot of oscillations that is not desirable. It shows that linear annealing schedule is not suitable for load balancing. We arrived at the same conclusion with the performance of log-decreasing annealing schedule shown in Fig. 3d.

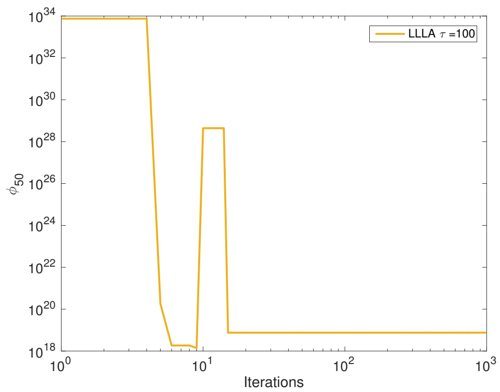
Now, we study the performance of ALA with different functions of τ. Particularly, we study algorithm A1, which is ALA with fixed $\boldsymbol{\tau} \in [10^{12}, 10^{10}, 10^8, 10^2]$, and algorithm A2, which is ALA with linearly decreasing $\boldsymbol{\tau} \in [10^{12}/t, 10^{10}/t, 10^8/t, 10^2/t]$. The probability of improvement of A1 and A2 compared to various annealing schedules is shown in Fig. 4. This probability is computed by averaging over 1000 realisations. It is observed that improvement probability of A1 and A2 compared to smallest $\tau = 100$ increases quickly as the number of iterations increases and reaches unity at 400 iterations. This shows both A1 and A2 are certainly better than LLLA with the smallest τ after 400 iterations. However, A2 is faster than A1. A1 and A2 are better than LLLA with linear and log annealing schedules. Particularly, they are better than linear and log annealing schedule more than 70% and 60% of times, respectively.

The performance comparison of A1 and A2 is shown in Fig. 5. The probability curves in the figure are computed by averaging over 500 realisations. The probability of A1 being strictly better than A2 is small in the beginning, decreases with time, and becomes zero at 400. Whereas, the probability of A2 being strictly better than A1 is higher in the beginning and then decreases to zero. The probability of A1 equal to A2 is higher, in the beginning, increases with time, and becomes unity at 400. Therefore, for low time horizons, it is better to use A2. For higher time horizons, A1 or A2 can be used.

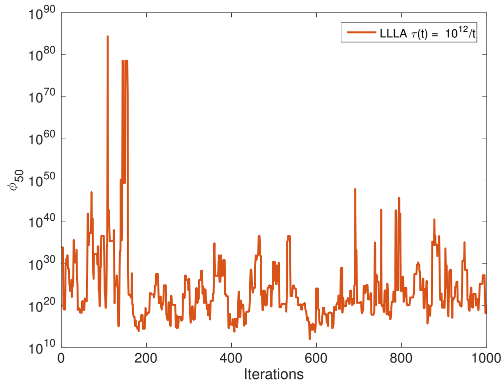
We now compare in Table II the optimal BS loads obtained using A1 for different α. As said earlier, the case with $\alpha = 0$ gives rate optimal policy, which is obtained when all biases equal to unity. This is verified by the output of ALA in Table II. This corresponds to the classical best signal association rule that results in heavy load imbalance between stations: the load of the macro BS reaches 92%, while small BSs have loads less than 11%. Min-max policy is obtained with a value of $\alpha = 50$. A min-max load vector of BSs is obtained as shown in Table II. Note that the load of macro



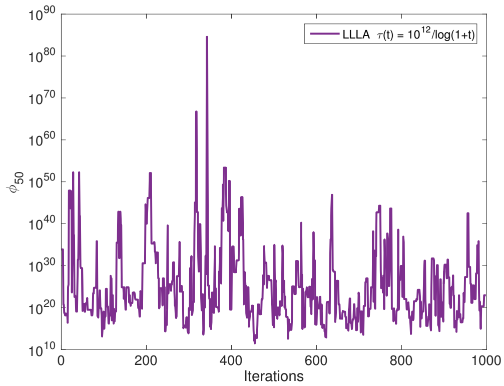
(a) ALA with $\boldsymbol{\tau} = [10^{12}, 10^{10}, 10^{8}, 10^{2}]$.



(b) LLLA with smallest $\tau = 10^2$.



(c) LLLA with linear decreasing τ.



(d) LLLA with log decreasing τ.

Fig. 3: Performance of different annealing schedules.

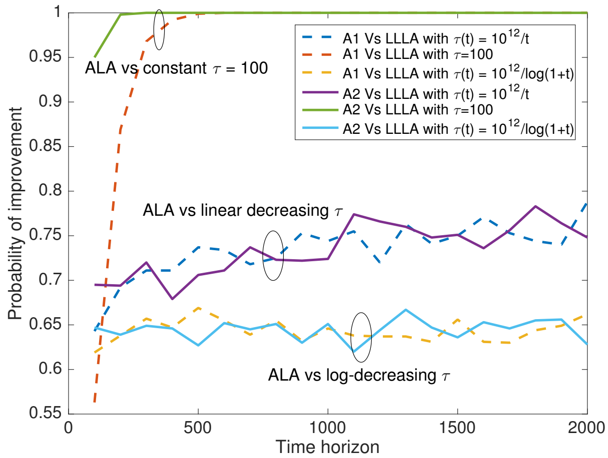


Fig. 4: Performance of A1 and A2.

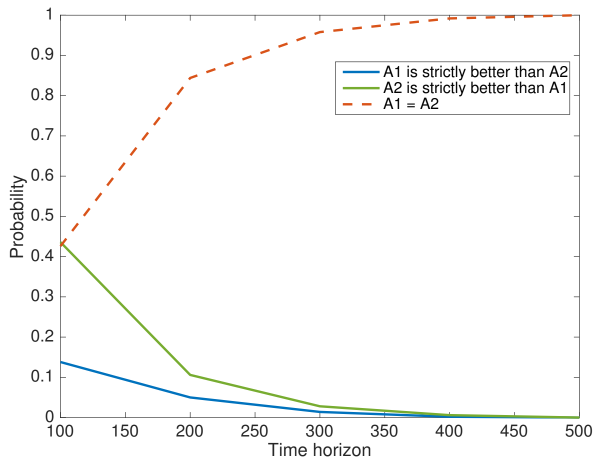


Fig. 5: Comparison of A1 and A2.

TABLE II: Comparison of optimal loads obtained using ALA for different α.

	$\alpha = 0$		$\alpha = 50$	
BS i	c_i	$\rho_i\%$	c_i	$\rho_i\%$
1	1	92	1	45
2	1	7	8	42
3	1	4	9	23
4	1	9	8	37
5	1	11	7	37
6	1	8	7	43
7	1	5	8	30
8	1	7	6	37

BS is reduced to 45%. Corresponding coverage regions for $\alpha = 50$ are shown in Fig. 1.

V. Conclusions

A new distributed annealing learning algorithm for learning the annealing schedule of LLLA over a fixed time horizon is proposed. The LLLA with the new annealing schedule is certainly better than with the fixed τ parameter. Also, the improvement probability of the proposed annealing schedule increases with the iterations when compared to linear decreasing and log decreasing annealing schedules of LLLA. For lower time horizons, ALA with linearly decreasing τ is better than ALA with fixed vector of τ. Whereas, for higher time horizons, ALA performance is same in both the cases.

References

[1] J. Andrews, S. Singh, Q. Ye, X. Lin, and H. Dhillon, "An overview of load balancing in hetnets: old myths and open problems," *IEEE Wireless Commun.*, vol. 21, pp. 18–25, April 2014.

[2] O. K. Tonguz and E. Yanmaz, "The mathematical theory of dynamic load balancing in cellular networks," *IEEE Trans. Mobile Comput.*, vol. 7, no. 12, pp. 1504–1518, 2008.

[3] K. Son, S. Chong, and G. Veciana, "Dynamic association for load balancing and interference avoidance in multi-cell networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3566–3576, 2009.

[4] B. Rengarajan and G. De Veciana, "Architecture and abstractions for environment and traffic-aware system-level coordination of wireless networks," *IEEE Trans. Netw. ACM*, vol. 19, no. 3, pp. 721–734, 2011.

[5] E. Stevens-Navarro, Y. Lin, and V. W. Wong, "An MDP-based vertical handoff decision algorithm for heterogeneous wireless networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 2, pp. 1243–1254, 2008.

[6] S.-E. Elayoubi, E. Altman, M. Haddad, and Z. Altman, "A hybrid decision approach for the association problem in heterogeneous networks," in *Proc. INFOCOM*, pp. 1–5, 2010.

[7] H. Kim, G. de Veciana, X. Yang, and M. Venkatachalam, "Distributed alpha-optimal user association and cell load balancing in wireless networks," *IEEE/ACM Trans. Netw.*, vol. 20, pp. 177–190, Feb 2012.

[8] F. Xu, C. C. Tan, Q. Li, G. Yan, and J. Wu, "Designing a practical access point association protocol," in *Proc. INFOCOM*, pp. 1–9, 2010.

[9] D. Niyato and E. Hossain, "Dynamics of network selection in heterogeneous wireless networks: An evolutionary game approach," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 2008–2017, May 2009.

[10] E. Aryafar, A. Keshavarz-Haddad, M. Wang, and M. Chiang, "RAT selection games in HetNets," in *Proc. INFOCOM*, pp. 998–1006, 2013.

[11] T. Kudo and T. Ohtsuki, "Cell range expansion using distributed Q-learning in heterogeneous networks," *EURASIP J. Wireless Commun. Netw.*, pp. 1–10, Mar. 2013.

[12] M. S. Ali, P. Coucheney, and M. Coupechoux, "Load balancing in heterogeneous networks based on distributed learning in potential games," in *International Symp. Modeling and Optimisation in Mobile, Ad Hoc, and Wireless Netw.*, 2015.

[13] B. Hajek, "Cooling schedules for optimal annealing," *Mathematics of operations research*, vol. 13, no. 2, pp. 311–329, 1988.

[14] Y. Nourani and B. Andresen, "A comparison of simulated annealing cooling strategies," *Journal of Physics A: Mathematical and General*, vol. 31, no. 41, p. 8373, 1998.

[15] J.-Y. Audibert and S. Bubeck, "Best arm identification in multi-armed bandits," in *COLT-23th Conference on Learning Theory*, pp. 13–p, 2010.

[16] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[17] M. Gudmundson, "Correlation model for shadow fading in mobile radio systems," *Electron. Lett.*, vol. 27, no. 23, pp. 2145–2146, 1991.

[18] R. Fraile, J. F. Monserrat, J. Gozálvez, and N. Cardona, "Mobile radio bi-dimensional large-scale fading modelling with site-to-site cross-correlation," *European trans. on telecommunications*, vol. 19, no. 1, pp. 101–106, 2008.

[19] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 10–21, 2011.

[20] R. Madan, J. Borran, A. Sampath, N. Bhushan, A. Khandekar, and T. Ji, "Cell association and interference coordination in heterogeneous LTE-A cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 28, pp. 1479–1489, Dec. 2010.

[21] K. Okino, T. Nakayama, C. Yamazaki, H. Sato, and Y. Kusano, "Pico cell range expansion with interference mitigation toward lte-advanced heterogeneous networks," in *ICC Workshop*, pp. 1–5, June 2011.

[22] M. Shirakabe, A. Morimoto, and N. Miki, "Performance evaluation of inter-cell interference coordination and cell range expansion in heterogeneous networks for LTE-advanced downlink," in *ISWCS*, pp. 844–848, Nov. 2011.

[23] M. Vajapeyam, A. Damnjanovic, J. Montojo, T. Ji, Y. Wei, and D. Malladi, "Downlink FTP performance of heterogeneous networks for LTE-advanced," in *ICC Workshop*, pp. 1–5, June 2011.

[24] I. Guvenc, "Capacity and fairness analysis of heterogeneous networks with range expansion and interference coordination," *IEEE Commun. Lett.*, vol. 15, pp. 1084–1087, Oct. 2011.

[25] A. Morimoto, N. Miki, H. Ishii, and D. Nishikawa, "Investigation on transmission power control in heterogeneous network employing cell range expansion for LTE-advanced uplink," in *European Wireless Conf.*, pp. 1–6, Apr. 2012.

[26] J. Oh and Y. Han, "Cell selection for range expansion with almost blank subframe in heterogeneous networks," in *Proc. PIMRC*, pp. 653–657, Sept 2012.

[27] K. Kitagawa, T. Komine, T. Yamamoto, and S. Konishi, "Performance evaluation of handover in LTE-advanced systems with pico cell range expansion," in *Proc. PIMRC*, pp. 1071–1076, Sep. 2012.

[28] M. Eguizabal and A. Hernandez, "Interference management and cell range expansion analysis for LTE picocell deployments," in *Proc. PIMRC*, pp. 1592–1597, Sept 2013.

[29] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Netw.*, vol. 8, no. 5, pp. 556–567, 2000.

[30] T. Bonald and L. Massoulié, "Impact of fairness on internet performance," in *ACM SIGMETRICS Performance Evaluation Review*, vol. 29, pp. 82–91, 2001.

[31] D. Monderer and L. S. Shapley, "Potential Games," *Games and Economic Behavior*, vol. 14, pp. 124–143, May 1996.

[32] J. R. Marden and J. S. Shamma, "Revisiting log-linear learning: Asynchrony, completeness and a payoff-based implementation," *Games and Economic Behaviour*, vol. 75, pp. 788–808, July 2012.

[33] 3GPP, "Technical specification group radio access network; evolved universal terrestrial radio access (eutra); further advancements for e-utra physical layer aspects," TR 36.814, 3GPP, Mar. 2010.