J Supercomput (2015) 71:871–893
DOI 10.1007/s11227-014-1330-7

A learning automata-based adaptive uniform fractional guard channel algorithm

Hamid Beigy · M. R. Meybodi

Published online: 7 December 2014

Abstract In this paper, we propose an adaptive call admission algorithm based on learning automata. The proposed algorithm uses a learning automaton to specify the acceptance/rejection of incoming new calls. It is shown that the given adaptive algorithm converges to an equilibrium point which is also optimal for uniform fractional channel policy. To study the performance of the proposed call admission policy, the computer simulations are conducted. The simulation results show that the level of QoS is satisfied by the proposed algorithm and the performance of given algorithm is very close to the performance of uniform fractional guard channel policy which needs to know all parameters of input traffic. The simulation results also confirm the analysis of the steady-state behaviour.

Keywords Reinforcement learning · Learning automata · Uniform fractional guard channel policy · Adaptive uniform fractional guard channel policy

1 Introduction

Microcellular networks use channels efficiently but expected rate of handover per call is increased. As a consequence, some network performance parameters such as blocking probability of new calls (B_n) and dropping probability of hand-off calls (B_h) are affected. To have these performance parameters at reasonable level, call

H. Beigy (✉)
Department of Computer Engineering, Sharif University of Technology, Tehran, Iran
e-mail: beigy@sharif.edu

M. R. Meybodi
Department of Computer Engineering, Amirkabir University of Technology, Tehran, Iran
e-mail: mmeybodi@aut.ac.ir

admission policies are used. The call admission policy plays a very important role in the cellular networks because it directly controls B_n and B_h by putting some restrictions on allocation of channels to calls. Since the dropping probability of hand-off calls is more important than the blocking probability of new calls, call admission policies usually give the higher priority to hand-off calls. This priority is implemented through allocation of more resources (channels) to hand-off calls.

In last decades, several call admission algorithms are proposed in the literature. Fractional guard channel policy (FG) accepts new calls with a certain probability that depends on the current channel occupancy and accepts hand-off calls as long as channels are available [1]. Suppose that the given cell has C full-duplex channels. This policy accepts new calls with probability π_k when k ($0 \leq k < C$) channels are busy. Depending on the probabilities π_k, we may have different call admission policies. Guard channel policy (GC) reserves a subset of channels allocated to a cell, called guard channels, for hand-off calls (say $C - T$ channels) [2]. Whenever the channel occupancy exceeds a certain threshold T, the guard channel policy rejects new calls until the channel occupancy goes below the threshold. The guard channel policy accepts hand-off calls as long as channels are available. It has been shown that there is an optimal threshold T^* at which the blocking probability of new calls is minimized subject to the hard constraint on the dropping probability of hand-off calls and an algorithm for finding such optimal threshold is given in [3]. The GC policy reserves an integral number of guard channels for hand-off calls. If performance parameter B_h is considered, the guard channel policy gives very good performance, but performance parameter B_n is degraded to great extent. To have more control on the blocking probability of new calls and the dropping probability of hand-off calls, limited fractional guard channel policy (LFG) is introduced [1]. The LFG can be obtained from FG policy by setting $\pi_k = 1$ (for $0 \leq k < T$), $\pi_T = \pi$ and $\pi_k = 0$ (for $T < k < C$). There is an optimal threshold T^* and an optimal value of π^* for which blocking probability of new calls is minimized subject to the hard constraint on dropping probability of hand-off calls and an algorithm for finding these optimal parameters is given in [1]. Uniform fractional guard channel policy (UFC) is also a restricted version of FG, which accepts new calls with probability of π independent of channel occupancy [4]. There is an optimal value of π^* for which blocking probability of new calls is minimized subject to the hard constraint on dropping probability of hand-off calls and an algorithm for finding the optimal value of π is given in [4]. It was shown that, the UFC policy performs better than GC policy under the low hand-off traffic conditions [4]. There are some call admission policies which allow either hand-off or new calls to be queued until free channels are obtained in the cell [5,6]. There also some call admission control algorithms that consider the multi-services networks [7–9].

All the above-mentioned call admission policies are static and assume that all parameters of traffic are known in advance. These policies are useful when input traffic is a stationary process with known parameters. Since the parameters of input traffic are unknown and possibly time varying, adaptive version of these policies must be used. In [10], a learning automata (LA)-based algorithm is given that adjusts π for uniform fractional channel policy. In [11], a LA-based algorithm is given which adjusts T and π for limited fractional guard channel policy. It was shown that the learning automaton finds the optimal values of T and π. In [12], a cellular learning automata-based algorithm is given which adjusts T for guard channel policy. In [13], two LA-based algorithms are given which adjust T for guard channel policy and their steady-state behaviour studied. These algorithms reserve channels for guard channels explicitly. The three last algorithms reserve channels to service calls while the algorithm presented in this paper does not reserve any channel. There are also some alternative approaches for managing channels in the cellular networks such as game theory and hidden Markov chains [14,15].

In this paper, we propose a learning automata-based adaptive uniform fractional guard channel algorithm (AUFC). The proposed algorithm uses a learning automaton to accept/reject new calls and a pre-specified level of dropping probability of hand-off calls is used to penalize/reward the action selected by the learning automaton. This algorithm accepts new calls as long as the dropping probability of hand-off calls is below the pre-specified threshold. Then, we study the steady-state behaviour of the proposed algorithm. It is shown that the given algorithm converges to an equilibrium point which is also optimal for UFC policy which knows all traffic parameters. The proposed algorithm unlike the algorithms given in [13] does not reserve any channel for guard channels explicitly. The simulation results show that, the performance of the proposed algorithm is very close to the performance of the UFC policy, which needs to know all traffic parameters in high hand-off traffic conditions and maintains the level of QoS in the system. The main advantage of the proposed algorithm is its steady-state behaviour. The value of π found by the proposed algorithm is equal to the optimal value of π found by the algorithm given in [4].

The rest of this paper is organized as follows: The proposed algorithm for finding the optimal value of parameter π and its steady-state behaviour are given in Sect. 2. The computer simulations is given in Sects. 3 and 4 concludes the paper.

2 Adaptive uniform fractional guard channel algorithm

The UFC policy uses admission probability π to accept new calls and accepts hand-off calls as long as channels are available [4]. To study the blocking performance of UFC policy, we assume that the network is homogeneous in which the arrival processes of new and hand-off calls are Poisson with rates λ_n and λ_h, respectively, the channel-holding time for both types of calls are exponentially distributed with mean μ^{-1}, and the time interval between two calls from a mobile host is much greater than the mean call holding time. The above assumptions have been found to be reasonable as long as the number of mobile hosts in a cell is much greater than the number of channels allocated to that cell. Since the network is homogeneous, we can examine the performance of a single network cell in isolation. We define the state of a particular cell at time t to be the number of busy channels in that cell and is represented by $c(t)$. $\{c(t)|t \geq 0\}$ is a continuous-time Markov chain (birth–death process shown in Fig. 1) with states $0, 1, \ldots, C$.

At state $0 \leq n < C$, new calls are accepted with probability $0 \leq \pi \leq 1$ and hand-off calls are accepted with probability 1. Both types of calls are blocked in state C. Thus, the state-dependent arrival rate in the birth–death process is equal to $[a + (1 - a)\pi]\lambda$,

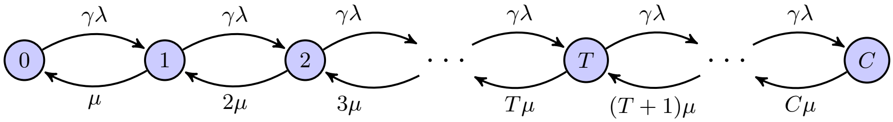

Fig. 1 Markov chain model of cell

where $\lambda = \lambda_n + \lambda_h$ and $\alpha = \lambda_h/\lambda$. Define the steady-state probability

$$P_n = \lim_{t\to\infty} \mathrm{Prob}[c(t) = n] \qquad n = 0, 1, \ldots, C. \tag{1}$$

The equilibrium equations for the steady-state probabilities P_n $(n = 0, 1, \ldots, C)$ are equal to

$$\gamma\lambda P_{n-1} = n\mu P_n,$$

where $\gamma = [a+(1-a)\pi]$. Solving the equilibrium equations, the following expression can be derived for P_n $(n = 0, 1, \ldots, C)$.

$$P_n = \frac{(\rho\gamma)^n}{n!} P_0, \tag{2}$$

where $\rho = \lambda/\mu$ and P_0 is the probability that all channels are free and obtained using equation $\sum_{n=0}^{C} P_n = 1$. The value of P_0 is calculated by the following expression:

$$P_0 = \left[\sum_{n=0}^{C} \frac{(\rho\gamma)^n}{n!}\right]^{-1} \tag{3}$$

The dropping probability of hand-off calls, denoted by $B_h(C, \pi)$, and the blocking probability of new calls, denoted by $B_n(C, \pi)$, are given by the following expressions:

$$B_h(C, \pi) = P_C = \frac{(\rho\gamma)^C}{C!} P_0. \tag{4}$$

$$B_n(C, \pi) = \sum_{n=0}^{C-1} (1-\pi) P_n + P_C = 1 - \pi\,[1 - B_h(C, \pi)] \tag{5}$$

The UFC policy is static and assumes that all parameters of traffic are known in advance. Static policies are useful when input traffic is a stationary process with known parameters. Since the parameters of input traffic are unknown and possibly time varying, adaptive version of these policies must be used. In other hand, learning automata (LA) are, by design, “simple agents for doing simple things”. Learning automata have been used successfully in many applications such as call admission control and channel assignment in cellular mobile networks [12,13,16], telephone and data network routing [17,18], solving NP-Complete problems [19,20], capacity assignment [21], optimization [22], neural network engineering [23,24], modelling the learning process [25], and discovering and tracking of spatiotemporal patterns in noisy sequences of events [26] to mention a few. A learning automaton (LA) has finite set of actions and at each stage chooses one of them. The choice of an action (α) depends on the state of LA represented by an action probability vector. For each action chosen by the LA, the environment gives a reinforcement signal with unknown probability distribution. Then, upon receiving the reinforcement signal (β), the LA updates its action probability vector by employing a learning algorithm [27]. Let α_i be the action chosen at time k as a sample realization from probability distribution $p(k)$. In linear reward–ϵpenalty algorithm ($L_{R-\epsilon P}$) scheme the recurrence equation for updating p is defined as

$$p_j(k+1) = \begin{cases} p_j(k) + a \times [1 - p_j(k)] & \text{if } i = j \text{ and environment gives reward} \\ p_j(k) - a \times p_j(k) & \text{if } i \neq j \text{ and environment gives reward} \\ p_j(k) \times (1-b) & \text{if } i = j \text{ and environment gives penalty} \\ \frac{b}{r-1} + p_j(k)(1-b) & \text{if } i \neq j \text{ and environment gives penalty} \end{cases} \tag{6}$$

Parameters $0 < b \ll a < 1$ represent *step lengths* and r is the number of actions for learning automaton; a and b determine the amount of increase and decreases of the action probabilities, respectively. If a equals b, the recurrence Eq. (6) is called linear reward–penalty (L_{R-P}) algorithm.

In this section, we first propose a learning automaton-based adaptive uniform fractional guard channel algorithm (AUFC) shown in Algorithm 1 and then study its steady-state behaviour. The proposed adaptive algorithm uses a learning automaton to determine the admission probability, π, when the parameters a and ρ (or equivalently λ_h, λ_n and μ) are unknown or probably time varying. The proposed algorithm adjusts parameter π as the network operates. This algorithm, which uses one reward–penalty type learning automaton with two actions in each cell, can be described as follows: The action set of this automaton corresponds to {*ACCEPT, REJECT*}. The automaton associated to each cell determines the probability of acceptance of new calls (π). Since initially the values of a and ρ are unknown, the probability of selecting these actions are set to 0.5. When a hand-off call arrives, it is accepted as long as there is a free channel. If there is no free channel, the hand-off call is dropped. When a new call arrives to a particular cell, the learning automaton associated to that cell chooses one of its actions. Let π be the probability of selecting action *ACCEPT*. Thus, the learning automaton accepts new calls with probability π as long as there is a free channel and rejects new calls with probability $1 - \pi$. If action *ACCEPT* is selected by automaton and the cell has at least one free channel, the incoming call is accepted and the selected action is rewarded because the selected action is successful. If there is no free channel to be allocated to the arrived new call, the call is blocked and action *ACCEPT* is penalized because the selected action failed in the environment. When the automaton selects action *REJECT*, the adaptive UFC computes an estimation of the dropping probability of hand-off calls ($\hat{B}_h$) and uses it to decide whether or not to accept the incoming new call. If the current estimate of dropping probability of hand-off calls is less than the given threshold p_h and there is a free channel, then the new call is accepted and the action *REJECT* is penalized because the selected action failed in the environment; otherwise, the new call is rejected and the action *REJECT* is rewarded because the appropriate action is selected. The learning automaton then uses this reinforcement signal to update the probability of accepting the new calls (π). This type of reinforcement signal increases the channel utilization because its main goal is to use channels as much as possible while the level of QoS is preserved.

The reward/penalty function measures goodness of actions applied to the environment. This function is a part of environment and affects both the steady-state behaviour and convergence speed of the algorithm. Actually, the reward/penalty function specifies the penalty probability of each action denoted by c and for all actions are shown by $c = (c_1, c_2)$, where c_k is the probability that action α_k will be penalized. This relation will be given in Lemma 1. This function also has the effect on the speed of convergence as it determines the value of increase/decrease of the probability of selection actions in the action probability vector p as given by Eq. (6).

Algorithm 1 Adaptive uniform fractional guard channel policy.

Require: *call* is the incoming call, C is the number of channels allocated to the cell, and p_h is the bound for dropping probability of hand-off calls.

procedure ADAPTIVEUFCCALLADMISION($call, C, p_h$)
    if *call* is a new call **then**
        Learning automaton chooses an action.
        if Learning automaton choose the ACCEPT action **then**
            if $c(t) < C$ **then**
                Accept *call* and reward the $ACCEPT$ action
            else
                Reject *call* and penalize the $ACCEPT$ action
            end if
        else ▷ Learning automaton chooses action $REJECT$ action
            Reject *call*
            if $(\hat{B}_h < p_h)$ **and** $(c(t) < C)$ **then**
                Accept *call* and penalize the $REJECT$ action
            else
                Reject *call* and reward the $REJECT$ action
            end if
        end if
    end if
end procedure

2.1 Behaviour of the adaptive UFC

In the rest of this section, we study the steady-state behaviour of the proposed algorithm. We show that, when the adaptive UFC algorithm uses the L_{R-P} reinforcement scheme, a unique value for π is found by the learning automaton, which is also optimal for the UFC algorithm. To study the behaviour of the adaptive UFC algorithm, we first model the environment for the learning automaton and then study the steady-state behaviour.

Lemma 1 *Let $p = (p_1, p_2)$ be the action probability vector of learning automata and $p_1 = \pi$ be the probability of accepting new calls. Then, the steady-state behaviour of the adaptive UFC algorithm can be shown by a triple $\langle \underline{\alpha}, \underline{\beta}, \underline{C} \rangle$, where $\underline{\alpha} = \{ACCEPT, REJECT\}$ shows the set of actions of automaton, $\underline{\beta} = \{0, 1\}$ represents the set of inputs for automaton and $\underline{C}(p) = \{c_1(p), c_2(p)\}$ is the set of penalty probabilities, where $c_1(p)$ and $c_2(p)$ are given by the following expressions.*

$$c_1(p) = P_C = \frac{(\rho\gamma)^C}{C!} P_0 \tag{7}$$

$$c_2(p) = \text{Prob}\left[\hat{B}_h < p_h\right] = \frac{1}{\sqrt{2\pi}\sigma_b} \int_{-\infty}^{p_h} e^{-\frac{1}{2}\left(\frac{x-\mu_b}{\sigma_b}\right)^2} dx \tag{8}$$

where μ_b and σ_b^2 are mean and variance of $\hat{B}_h$, respectively.

The above Lemma gave the model of the environment and in the following Lemma, we study the properties of the environment.

Lemma 2 *When $\rho < C$, the environment given in Lemma 1 has the following characteristics. Let us to write p for $p(n)$ and $c_i(p)$ for $c_i(n)$.*

1. *$c_i(p)$ (for $i = 1, 2$) is continuous function in p.*
2. *$c_i(p)$ (for $i = 1, 2$) are continuously differentiable in all their arguments.*
3. *$c_i(p)$ and $\frac{\partial c_i(p)}{\partial p_i}$ (for $i = 1, 2$) are Lipschitz function of all their arguments.*
4. *The derivative of $c_i(p)$ (for $i = 1, 2$) have the following features.*

$$\frac{\partial c_i(p)}{\partial p_i} > 0, \tag{9}$$

$$\frac{\partial c_1(p)}{\partial p_2} \ll \frac{\partial c_2(p)}{\partial p_2}, \tag{10}$$

$$\frac{\partial c_2(p)}{\partial p_1} \ll \frac{\partial c_1(p)}{\partial p_1}. \tag{11}$$

The process $\{p(n)\}_{n\geq 0}$ defined by the adaptive UFC algorithm is a homogeneous Markov process. The following theorem is concerned with its convergence behaviour and its proof given in [28].

Theorem 1 *The Markov process $\{p(n)\}_{n\geq 0}$ is ergodic and converges in distribution as $n \to \infty$ to a unique stationary probability $\bar{p}$ independent of the initial distribution of $\bar{p}$.*

In what follows, the steady-state behaviour of the adaptive UFC algorithm will be analysed. Define the average penalty rate of action α_i as $f_i\left(p(n)\right) = c_i\left(p(n)\right) p_i(n)$, $p^* = \left(p_1^*, p_2^*\right)$ and $p_1^* + p_2^* = 1$. In the following lemma, it is shown that there is a unique p^* for which the average penalty rates for both actions become equal.

Lemma 3 *For the adaptive UFC algorithm, there exists a unique p^* such that*

$$f\left(p^*\right) = f_2\left(p^*\right) - f_1\left(p^*\right) = 0. \tag{12}$$

Let $\Delta p_1(n) = p_1(n+1) - p_1(n)$, thus the conditional expectation and variance of Δp_1 are expressed as

$$\begin{aligned} \mathrm{E}\left[\frac{\Delta p_1(n)}{a} | p(n) = p\right] &= f_2(p) - f_1(p) \\ &= w(p) \end{aligned} \tag{13}$$

Let

$$S(p) = \mathrm{E}\left[\frac{\Delta p_1^2(n)}{a^2} | p(n) = p\right].$$

Thus, the variance of Δp_1 is equal to

$$\tilde{S}(p) = S(p) - w^2(p).$$

Since $\{p(n)\}_{n \geq 0}$ is ergodic and converges in distribution to a unique stationary probability $\bar{p}$, thus in steady state, we obtain $\mathrm{E}\left[\Delta \bar{p}_i\right] = 0$ or $\mathrm{E}\left[w(\bar{p})\right] = 0$. The zero of $\mathrm{E}\left[w(\bar{p})\right]$ is p^* and, in general, $\mathrm{E}\left[w(\bar{p})\right] = 0$ need not yield p^*. However, if the learning parameter a is chosen to be sufficiently small, then the difference between these two values may be made small, as indicated by the following theorem.

Lemma 4 *Let $p(0)$ be the initial action probability vector of the adaptive UFC algorithm with stationary measure $\bar{p}$, then*

$$\begin{aligned} \mathrm{E}\left[p_i(n) - p_i^*\right]^2 &\leq Ka, \\ &= O(a), \end{aligned}$$

where $K > 0$ denotes a constant.

Since $\left(\mathrm{E}\left[p_i(n) - p^*\right]\right)^2 \leq \mathrm{E}\left[p_i(n) - p^*\right]^2$, $\mathrm{E}\left[p_i(n)\right] - p_i^* = O\left(\sqrt{a}\right)$ and $p_i(n) \to p_i^*$ with probability 1 as $a \to 0$. In steady state, $\mathrm{E}\left[\Delta \bar{p}_i\right] = 0$ or $\mathrm{E}\left[w\left(p(n)\right)\right] \to 0$ as $n \to \infty$. This implies that $\mathrm{E}\left[f_2(\bar{p})\right] = \mathrm{E}\left[f_1(\bar{p})\right]$. From this equation and Lemma 3, we can write

$$\mathrm{E}\left[f_2(\bar{p})\right] - f_2(p^*) = \mathrm{E}\left[f_1(\bar{p})\right] - f_1(p^*). \tag{14}$$

Since $f_i(.)$ is a Lipschitz function with Lipschitz bound α, we have

$$\begin{aligned} \left|\mathrm{E}\left[f_i(\bar{p}) - f_i(p^*)\right]\right| &\leq \alpha \left|\mathrm{E}\left[\bar{p}_i - p_i^*\right]\right|, \\ &\leq \alpha \left|\mathrm{E}\left[\bar{p}_i\right] - p_i^*\right|, \\ &= O(\sqrt{a}). \end{aligned}$$

Thus, we have $\left|\mathrm{E}\left[f_i(\bar{p}) - f_i(p^*)\right]\right| = O(\sqrt{a})$. Hence for small values of the parameter a, it can be concluded that asymptotic behaviour of the adaptive UFC algorithm can be approximated by $f(p^*) = 0$.

Lemma 5 *Let $p(0)$ be the initial action probability vector of the adaptive UFC algorithm with stationary measure $\bar{p}$ and define $z_i(n) = \frac{p_i(n)-p_i^*}{\sqrt{a}}$ and $z(n) = z_1(n)$, then the following relations are held. These relations verify the condition of lemma 2.1 (pp. 156) [29], whose application results in the normal approximation.*

$$\mathrm{E}\left[\Delta z(n)|z(n)\right] = aw'(p^*)z(n) + o(a), \tag{15}$$

$$\mathrm{E}\left[\Delta z^2(n)|z(n)\right] = a\tilde{S}(p^*) + o(a) \tag{16}$$

$$\mathrm{E}\left[|\Delta z(n)|^3|z(n)\right] = o(a), \tag{17}$$

where $o(a)$ denotes a random variable such that $\mathrm{E}\left[\frac{o(a)}{a}\right] \to 0$ *as* $a \to 0$.

The following Theorem states that the steady-state probability of accepting new calls found by the proposed algorithm is equal to the optimal value found by the algorithm given in [4].

Theorem 2 *Let $p(0)$ be the initial action probability vector of the adaptive UFC algorithm with stationary measure $\bar{p}$ and Lemma 4 holds, then*

$$z_i(n) \sim N\left(0, \sigma^2\right),$$

where $\sigma^2 = \frac{\tilde{S}(p^*)}{2|w'(p^*)|}$ as $a \to 0$ and $na \to \infty$.

Theorem 3 *The equilibrium probability of learning automaton in the adaptive UFC algorithm, $p^* = (\pi^*, 1-\pi^*)$, minimizes the blocking probability of new calls subject to the hard constraint on the dropping probability of hand-off calls ($B_h(C, \pi) \le p_h$).*

3 Simulation results

In this section, through simulation we compare the performance of the uniform fractional guard channel algorithm [4], an adaptive uniform fractional guard channel algorithm [10] and the proposed algorithm. To compare these algorithms, computer simulations are conducted. Simulation are conducted based on the single cell of homogeneous cellular network system. In such network, each cell has eight full-duplex channels ($C = 8$). Figures 2 and 3 compare the dropping probability of hand-off calls and the blocking probability of new calls for different hand-off traffic. In these simulations, new call arrival rate is fixed to 30 calls per minute ($\lambda_n = 30$), channel-holding time is set to 6 s ($\mu^{-1} = 6$), and the hand-off call traffic is varied between two calls per minute and 20 calls per minute. The results shown in these figures are obtained by averaging 10 runs from 2,000,000 seconds simulation of each algorithm. The level of QoS for the dropping probability of hand-off calls is set to 0.01. The optimal value of π of uniform fractional channel policy is obtained by algorithm given [4]. Figure 2 shows the dropping probability of hand-off calls of the proposed algorithm and the algorithm given in [10] for different rates of hand-off traffic. As this figure shows,

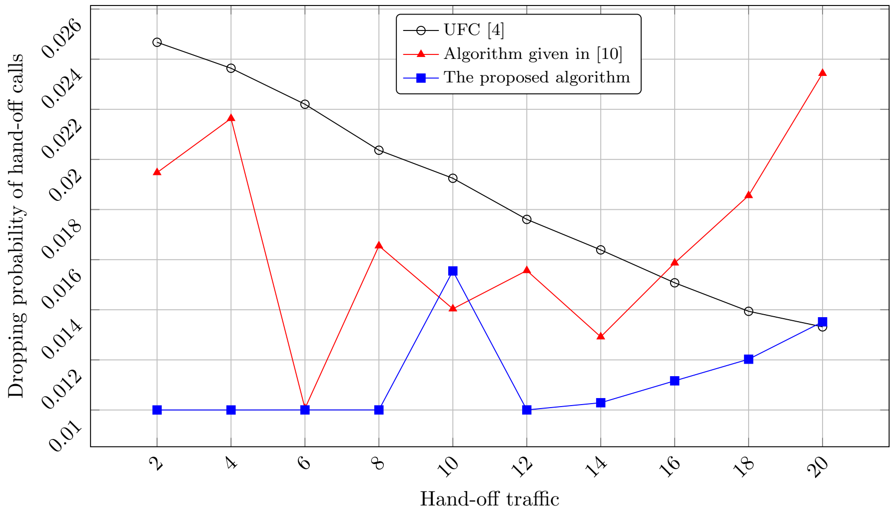

Fig. 2 Dropping probability of the proposed algorithm for different hand-off traffic

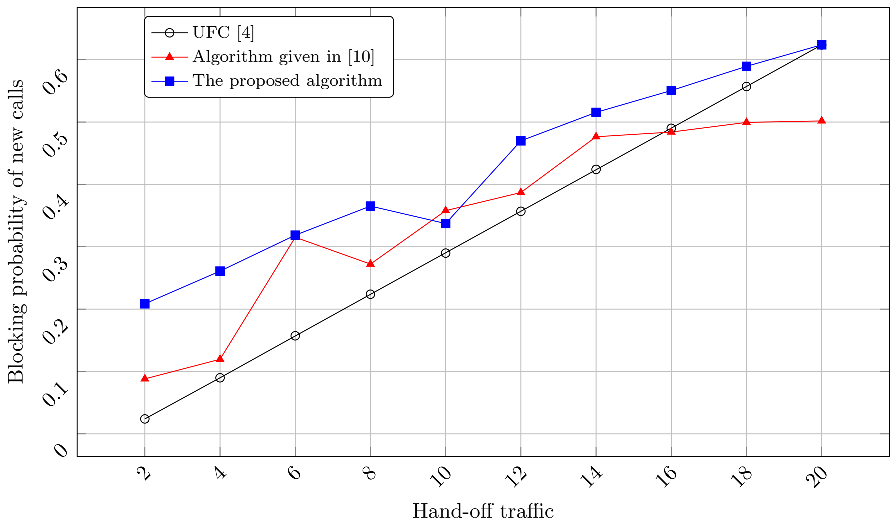

Fig. 3 Blocking probability of the proposed algorithm for different hand-off traffic

the ability of the proposed algorithm for maintaining the QoS of hand-off traffic for different hand-off load is higher than the algorithm given in [10], although there are some variation and the proposed algorithm cannot maintain the hard constraint.

Figure 3 compares the blocking probability of new calls for the proposed algorithm and the algorithm given in [10] for different rates of hand-off traffic. As this figure shows, the blocking probability of new calls for the proposed algorithm is higher than the blocking probability of new calls the algorithm given in [10]. This is due to the higher priority given to hand-off calls to maintain the level of QoS specified by p_h.

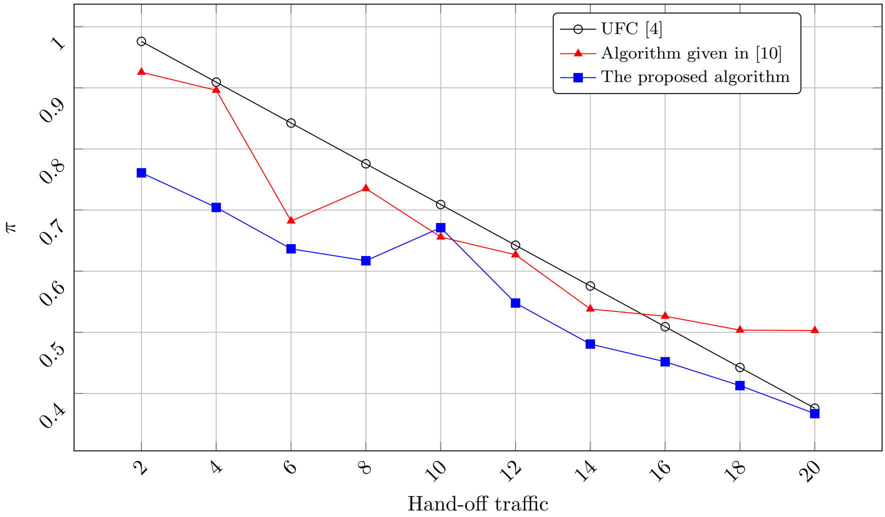

Fig. 4 New calls acceptance probability of algorithms for different hand-off traffic

By carefully inspecting Figs. 2 and 3, it is evident that for some range of input traffics, the performance of the proposed policy is close to the performance of the uniform fractional guard channel policy and performs better than the algorithm given in [10]. Since in the low hand-off traffic conditions, the UFC policy does not maintain the upper bound on the dropping probability of hand-off calls, the blocking probability of new calls for the proposed algorithm is greater than the blocking probability of new calls for UFC. When the hand-off traffic becomes high, the UFC policy maintains the upper bound on the dropping probability of hand-off calls and the performance of UFC policy and the proposed algorithm is very close. In such situations, the probability of accepting new calls converges to the optimal value found by the algorithm given in [4].

Figure 4 shows the new call acceptance probability of the above-mentioned algorithms for different rates of hand-off traffic when the other parameters of the cell are fixed. As this figure shows, the value of π is a decreasing function of the hand-off traffic and the proposed algorithm gives a smaller value of π when hand-off traffic becomes higher. This is due to the higher priority given to hand-off calls.

Figure 5 shows the blocking probability of new calls and dropping probability of hand-off calls of the proposed algorithm for different rates of hand-off traffic when the other parameters of the cell are fixed. As this figure shows, the proposed algorithm is able to maintain the QoS of hand-off traffic for different hand-off load. The blocking probability of new calls will be increased as the hand-off traffic increase, because the number of channels allocated to the cell is fixed.

Figure 6 shows the evolution of the performance parameters for the adaptive uniform fractional guard channel policy. By carefully inspecting Fig. 6, it is evident that the admission probability, π, oscillates and converges to a value. The converged value is also the optimal value obtained by the algorithm given [4].

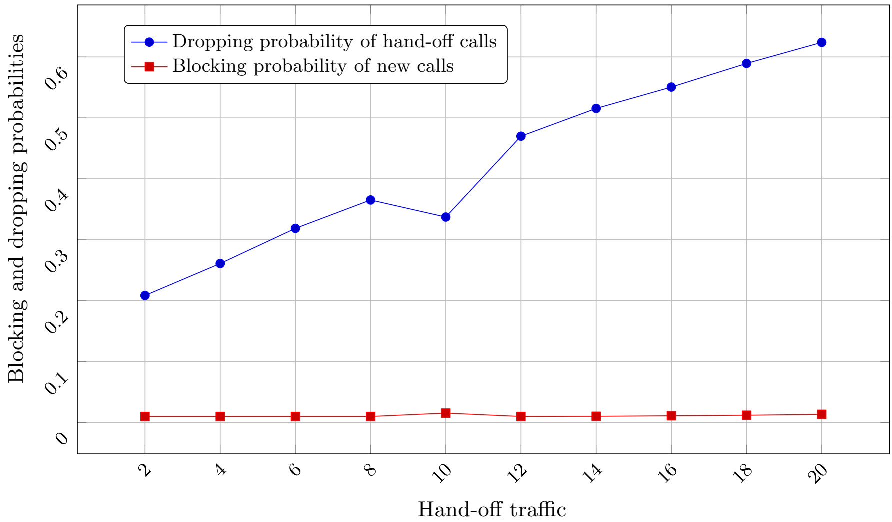

Fig. 5 Performance of the proposed algorithm for different hand-off traffic

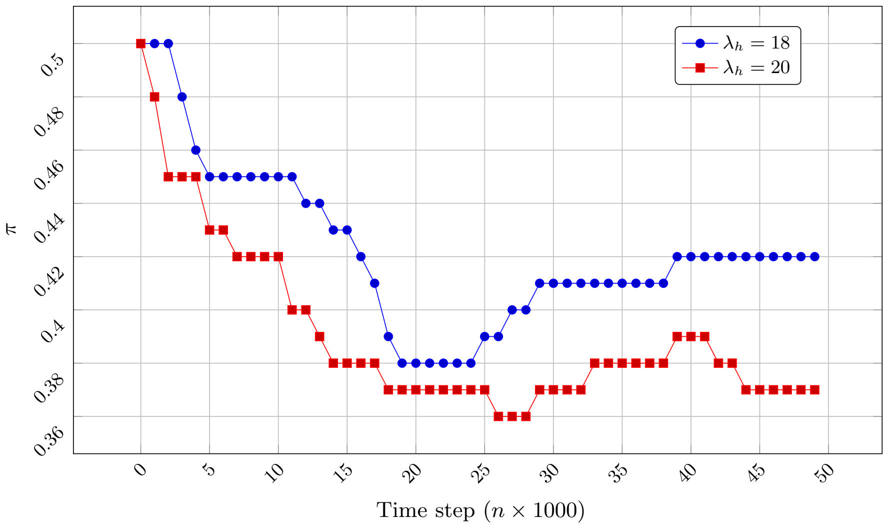

Fig. 6 The proposed algorithm for different hand-off traffic

Figure 7 shows the performance of the proposed algorithm in a non-stationary environment. In this figure, the algorithm starts with the following traffic setting: $\lambda_n = 30$, $\lambda_h = 20$, and $\mu^{-1} = 6$. Then after 1,650 s, the traffic is changed to $\lambda_n = 30$, $\lambda_h = 10$, and $\mu^{-1} = 6$. This figure shows that the proposed algorithm is able to track the changes in the environment without knowing the parameters of the environment.

Table 1 compares the the value found by the proposed algorithm and the algorithm given in [4]. This table shows that the two values are very close while the proposed

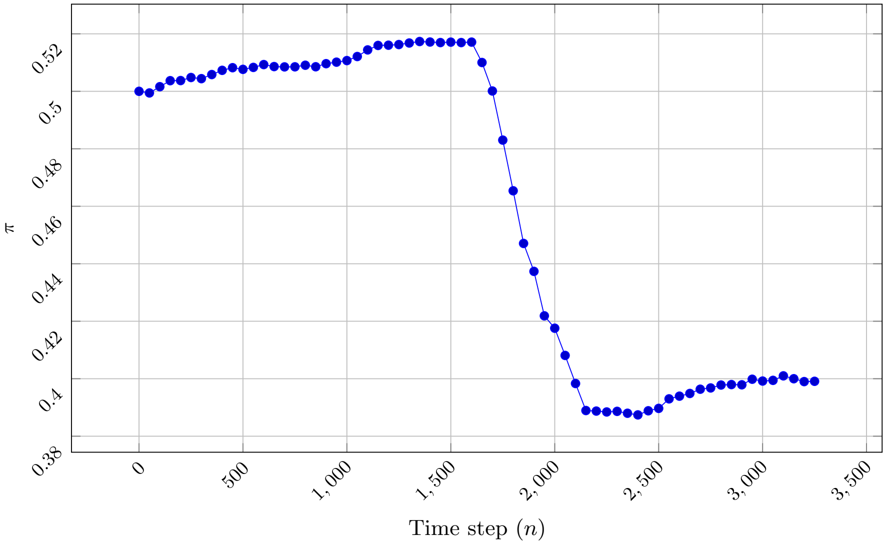

Fig. 7 The tracking capability of the proposed algorithm in non-stationary traffic

Table 1 The optimal value of π and converged value of π

λ_n	λ_h	μ^{-1}	π^*	π
20	5	6	0.6882	0.6795
20	5	4	0.3755	0.3722
15	10	6	0.9758	0.9625
15	10	4	0.5006	0.49101

algorithm gives a higher priority (lower value of π) to the hand-off calls. This results was obtained also for other traffic parameters as mentioned before. The results presented in this table also confirm the results obtained in Theorem 3. The difference between π^* and π is due to the finite time simulation.

4 Conclusions

In this paper, we first propose a new learning automata-based adaptive uniform fractional channel algorithm. The proposed algorithm uses a learning automaton to determine the admission probability, π, when the traffic parameters are unknown or probably time varying. This algorithm adjusts parameter π as network operates. Then, we studied the steady-state behaviour of the proposed algorithm. It was shown that the proposed algorithm converges to an equilibrium point which is also optimal for UFC policy.The simulation results show that the level of QoS is satisfied by the proposed algorithm and the performance of given algorithm is very close to the performance of uniform fractional guard channel policy which needs to know all parameters of input traffic. The simulation results also show that in non-stationary environment, the proposed algorithm is able to adapt it as the environment changes. The rate of convergence is also investigated by the simulation. This rate depends on some parameters such as the reward/penalty function and learning rate parameters. As a future work, we will investigate the speed of the convergence analytically.

Acknowledgments The authors would like to thank the anonymous reviewers for their valuable comments and suggestions which improved the paper.

5 Appendix: Proof of Theorems and Lemmas

In this appendix, we give the proof of some lemmas and theorems given in this paper.

5.1 Proof of Lemma 1

Before we begin to prove the lemma, we introduce some definitions and notations. To count how many calls are arrived, we introduce concept of local time for each type of calls. The local time for each type of calls starts with 0 and incremented by 1 when a call of given type is arrived. Let us to define n^n and n^h as the local times for new and hand-off calls, respectively. Then, we define two sequences of random variables n^n_m ($n^n_1 < n^n_2 < \cdots$) and n^h_m ($n^h_1 < n^h_2 < \cdots$), where n^n_m (n^h_m) is the global time when the m^{th} new (hand-off) call is arrived.

The proof for penalty probability of $c_1(p)$ is trivial, because action ACCEPT is penalized when all allocated channels are busy. Since the probability of all channels being busy is equal to P_C, then $c_1(p)$ is equal to P_C. To find expression for $c_2(p)$, we define X_n as the indicator of dropping of a hand-off call at the hand-off local time n, where $X_n = 1$ if a hand-off call arrives at hand-off local time $n^h = n$ and dropped and $X_n = 0$ if a hand-off call arrives at hand-off local time $n^h = n$ and accepted. Since in interval $[n, n + 1]$, it is possible that $M \geq 0$ new calls to be accepted or $N \geq 0$ calls to be completed, then the state of the Markov chain describing cell at hand-off local time $n + 1$ is independent of its state at the hand-off local time n when $N + M > 0$. Although there is an exception $N + M = 0$, which we ignore in our analysis due to the violation of Markov chain properties. Therefore, $X_1, X_2, \ldots, X_n$ are independent identically distributed (i.i.d) random variables with the following first- and second-order statistics.

$$\mathrm{E}\left[X_n\right] = \sum_{k=0}^{C} k P_k = \rho\gamma\left[1 - P_C\right]. \tag{18}$$

$$\mathrm{Var}\left[X_n\right] = \mathrm{E}\left[X_k^2\right] - \left(\mathrm{E}\left[X_k\right]\right)^2 = \rho\gamma\left[1 - P_C\right]\left[1 + \rho\gamma P_C\right] - (\rho\gamma)^2 P_{C-1}. \tag{19}$$

Using the central limit theorem $\bar{X}_n = \hat{B}_h = \frac{1}{n}\sum_{k=0}^{n} X_k$ is a random variable with normal distribution ($\hat{B}_h \sim N(\mu_b, \sigma_b)$) with the following mean and variance [30].

$$\mu_b = \mathrm{E}\left[\hat{B}_h\right] = \mathrm{E}\left[\bar{X}_n\right] = \mathrm{E}\left[X_n\right] = \rho\gamma\left[1 - P_C\right]. \tag{20}$$

$$\begin{aligned}\sigma_b &= \mathrm{Var}\left[\hat{B}_h\right] = \mathrm{Var}\left[\bar{X}_n\right] = \frac{\mathrm{Var}\left[X_n\right]}{n},\\ &= \frac{\rho\gamma\left[1 - P_C\right]\left[1 + \rho\gamma P_C\right] - (\rho\gamma)^2 P_{C-1}}{n}.\end{aligned} \tag{21}$$

Thus, the value of penalty probability of $c_2(p)$ is equal to

$$\begin{aligned}c_2(p) &= \mathrm{Prob}\left[\hat{B}_h < p_h\right],\\ &= \frac{1}{\sqrt{2\pi}\sigma_b}\int_{-\infty}^{p_h} e^{-\frac{1}{2}\left(\frac{x-\mu_b}{\sigma_b}\right)^2} dx\end{aligned}$$

which completes the proof of this lemma. □

5.2 Proof of Lemma 2

The proofs for items one through three are trivial using Eq. (7) and we only give the proof of item 4. From Eq. (7) and when $\rho < C$, we have

$$\frac{\partial c_1(p)}{\partial p_1} = (1-a)P_C\left[\frac{C}{\gamma} - \rho(1 - P_C)\right] > 0, \tag{22}$$

$$\frac{\partial c_1(p)}{\partial p_2} = -(1-a)P_C\left[\frac{C}{\gamma} - \rho(1 - P_C)\right] < 0. \tag{23}$$

Using Eq. (7), we obtain

$$\begin{aligned}\frac{\partial c_2(p)}{\partial p_2} &= \frac{1}{\sigma_b\sqrt{2\pi}}\Bigg[e^{\frac{-1}{2}\left(\frac{p_h-\mu_b}{\sigma_b}\right)^2}\left\{\left(\frac{\mu - p_h}{\sigma_b}\right)\frac{\partial\sigma_b}{\partial p_2} - \frac{\partial\mu_b}{\partial p_2}\right\}\\ &\quad - \frac{2}{\sigma_b}\frac{\partial\sigma_b}{\partial p_2}\int_{-\infty}^{p_h} e^{\frac{-1}{2}\left(\frac{x-\mu_b}{\sigma_b}\right)^2} dx\Bigg],\\ \frac{\partial c_2(p)}{\partial p_1} &= \frac{-1}{\sigma_b\sqrt{2\pi}}\Bigg[e^{\frac{-1}{2}\left(\frac{p_h-\mu_b}{\sigma_b}\right)^2}\left\{\left(\frac{\mu - p_h}{\sigma_b}\right)\frac{\partial\sigma_b}{\partial p_1} - \frac{\partial\mu_b}{\partial p_1}\right\}\\ &\quad - \frac{2}{\sigma_b}\frac{\partial\sigma_b}{\partial p_1}\int_{-\infty}^{p_h} e^{\frac{-1}{2}\left(\frac{x-\mu_b}{\sigma_b}\right)^2} dx\Bigg].\end{aligned}$$

Increasing p_2, decreases the probability of accepting new calls and hence the number of busy channels decreased. Therefore, the dropping probability of hand-off calls is decreased or $c_2(p) = \mathrm{Prob}\left[\hat{B}_h < p_h\right]$ is increased. Thus, we have

$$\frac{\partial c_2(p)}{\partial p_2} > 0 \tag{24}$$

$$\frac{\partial c_2(p)}{\partial p_1} < 0. \tag{25}$$

However, by choosing the proper value for parameters, condition $\frac{\partial c_2(p)}{\partial p_2} > 0$ is also satisfied. From Eqs. (22) and (24), Eq. (9) is concluded, from Eqs. (22) and (25), Eq. (10) is concluded and from Eqs. (23) and (24), Eq. (11) is concluded. This completes the proof of this lemma. □

5.3 Proof of Lemma 3

Consider $f(p)$ at its two end points

$$f(p) = \begin{cases} c_2(0,1) & p_1 = 0 \\ -c_1(1,0) & p_1 = 1. \end{cases} \tag{26}$$

Since $f(p)$ is a continuous function of p_1 and p_2, there exists at least a p^* such that $f(p^*) = 0$. For proving the uniqueness of p^*, the derivative of $f(p)$ with respect to p_1 is computed and then using Lemma 2, we obtain

$$\begin{aligned} \frac{\partial f(p)}{\partial p_1} &= \frac{\partial c_2(p)}{\partial p_1} - (1 + p_1)(c_1 + c_2), \\ &< 0. \end{aligned}$$

Since the derivative of $f(p)$ with respect to p_1 is negative, $f(p)$ is a strictly decreasing function of p_1. Thus there exists one and only one point p^* for which function $f(p)$ crosses the horizontal line and hence the lemma. □

5.4 Proof of Lemma 4

Define $p^2 = p^T p$ for vector p. Let $p = p_1$ and

$$g(p) = \begin{cases} \frac{w(p)}{p^* - p} & p \neq p^* \\ -\left.\frac{\partial w(p)}{\partial p}\right|_{p=p^*} & p = p^* \end{cases} \tag{27}$$

Since $w(p) < 0$ when $p > p^*$ and $w(p) > 0$ when $p < p^*$, $g(p)$ is positive and continuous in interval [0, 1]. Hence, there exists a $R > 0$ such that $g(p) \geq R$. Thus, we have

$$\begin{aligned} \left[p^* - p(n)\right] w(p(n)) &= \left[p^* - p(n)\right]^2 g(p(n)), \\ &\geq R\left[p^* - p(n)\right]^2. \end{aligned} \tag{28}$$

for all probability p, then computing

$$\left[p(n+1)-p^*\right]^2=\left[p(n)-p^*\right]^2+2\left[p(n)-p^*\right]\Delta p(n)+\Delta p^2(n)$$

and taking expectation on both sides, cancelling $\mathrm{E}\left[p(n)-p^*\right]^2$ and dividing by $2a$, we obtain

$$\mathrm{E}\left[\left\{p(n)-p^*\right\}\frac{\Delta p(n)}{a}\right]+\frac{a}{2}\mathrm{E}\left[\frac{\Delta p^2(n)}{a^2}\right]=0,$$

or

$$\mathrm{E}\left[\left\{p(n)-p^*\right\}w\left(p(n)\right)\right]+\frac{a}{2}\mathrm{E}\left[\tilde{S}\left(p(n)\right)\right]=0.$$

Since, we have only bounded variables, $\tilde{S}\left(p(n)\right)$ is also bounded; thus, there exists a $K>0$ such that $\mathrm{E}\left[\tilde{S}\left(p(n)\right)\right]\leq K$. Hence, we obtain

$$\begin{aligned}\mathrm{E}\left[\left\{p^*-p(n)\right\}w\left(p(n)\right)\right]&=\frac{a}{2}\mathrm{E}\left[\tilde{S}\left(p(n)\right)\right],\\&\leq Ka.\end{aligned}$$

Using this Eq. (28), we obtain

$$\begin{aligned}\mathrm{E}\left[p^*-p(n)\right]^2&\leq K\mathrm{E}\left[\left\{p^*-p(n)\right\}w(p(n))\right],\\&\leq Ka,\\&=O(a).\end{aligned}$$

and hence the lemma. □

5.5 Proof of Lemma 5

To prove Eq. (15), let us to define

$$\zeta=\frac{\mathrm{E}\left[\Delta z(n)|z(n)\right]}{\sqrt{a}}=\frac{\mathrm{E}\left[\Delta p(n)|z(n)\right]}{a}=w\left(p(n)\right)-w\left(p^*\right). \tag{29}$$

Since $w(.)$ is Lipshitz with bound β, we have $|w\left(p(n)\right)-w\left(p^*\right)|\leq K|p(n)-p^*|$, where $K>0$ is a constant. Using this Eq. (29), we obtain

$$|\zeta|\leq K|p(n)-p^*|\leq K\sqrt{a}|z(n)|. \tag{30}$$

Let

$$h(\lambda)=w(x+\lambda(y-x))$$

where $\lambda \in [0, 1]$. It follows that

$$\begin{aligned} h'(\lambda) &= \frac{\partial h(\lambda)}{\partial \lambda}, \\ &= w'(x + \lambda(y - x))(y - x). \end{aligned} \tag{31}$$

Since $h'(.)$ is continuous, we have

$$w(y) - w(x) = h(1) - h(0) = \int_0^1 w'(x + \lambda(y - x))[y - x]d.\lambda \tag{32}$$

Subtracting $w'(x)(y - x)$ from both sides of the above equation, we obtain

$$w(y) - w(x) - w'(x)[y - x] = \int_0^1 \left[w'(x + \lambda(y - x)) - w'(x)\right][y - x]d\lambda$$

Since $w(.)$ is Lipschitz with bound β, we obtain

$$w(y) - w(x) - w'(x)(y - x) \leq \frac{\beta}{2}|y - x|^2.$$

Substituting y with $p(n)$ and x with p^* in the above equation, we obtain

$$\begin{aligned} w(p(n)) - w(p^*) - w'(p^*)(p(n) - p^*) &\leq K|p(n) - p^*|^2, \\ w(p(n)) - w(p^*) - \sqrt{a}w'(p^*)z(n) &\leq Ka|z(n)|^2. \end{aligned} \tag{33}$$

Using this Eqs. (29) and (30) and Lemma 4, we obtain

$$\begin{aligned} |\zeta - \sqrt{a}w'(p^*)z(n)| &\leq Ka|z(n)|^2, \\ &\leq K|p(n) - p^*|^2, \\ &\leq Ka. \end{aligned} \tag{34}$$

Multiplying both sides of the above equation by $\sqrt{a}$, we obtain

$$\left|\sqrt{a}\zeta - aw'(p^*)z(n)\right| \leq Ka^{3/2},$$

or

$$\left|\mathrm{E}[\Delta z(n)|z(n)] - aw'(p^*)z(n)\right| \leq K\sqrt{a},$$

which implies Eq. (15). To derive Eq. (16), let us to define

$$\eta = \frac{\mathrm{E}\left[\Delta z^2(n)|z(n)\right]}{a} = S(p(n)) = \tilde{S}(p(n)) + \zeta^2.$$

By subtracting $\tilde{S}(p(n))$ from both sides of the above equation, we obtain

$$\begin{aligned}|\eta - \tilde{S}(p^*)| &= |\tilde{S}(p(n)) + \zeta^2 - \tilde{S}(p^*)| \\ &\leq |\tilde{S}(p(n)) - \tilde{S}(p^*)| + |\zeta^2|.\end{aligned} \tag{35}$$

Since $\tilde{S}(.)$ is Lipschitz, we have

$$|\tilde{S}(p(n)) - \tilde{S}(p^*)| \leq K|p(n) - p^*|.$$

Substituting this Eq. (30) into Eq. (35), we obtain

$$|\eta - \tilde{S}(p^*)| \leq K|p(n) - p^*| + K|p(n) - p^*|^2$$

Using Lemma 4, we have $\mathrm{E}\left[p(n) - p^*\right]^2 \leq Ka$ and $\mathrm{E}\left[p(n) - p^*\right] \leq K\sqrt{a}$. Thus, we obtain $|\eta - \tilde{S}(p^*)| = o(a)$. Hence, as a consequence, we have $\mathrm{E}|\eta - \tilde{S}(p^*)| \to 0$ as $a \to 0$, which confirms Eq. (16). Equation (17) follows by observing that

$$\mathrm{E}\left[\left|\frac{\Delta p(n)}{a}\right|^3 \middle| p(n) = p\right] = \xi(p) < \xi < \infty$$

Substituting Eq. (17) into the above equation, we obtain

$$\begin{aligned}\mathrm{E}\left[\left|\frac{\Delta p(n)}{a}\right|^3 \middle| p(n)\right] &< \xi \\ \mathrm{E}\left[\left|\frac{|\Delta z(n)|^3}{a^{3/2}}\right| p(n)\right] &< \xi \\ \mathrm{E}\left[|\Delta z(n)|^3 \middle| p(n)\right] &< \xi a^{3/2}\end{aligned} \tag{36}$$

where $\xi a^{3/2} \to 0$ as $a \to 0$. This completes the proof of this lemma. □

5.6 Proof of Theorem 2

Let $h(u) = \mathrm{E}\left[e^{iuz(n)}\right]$ be the characteristic function of $z(n)$. Then using the third-order taylor's expansion of e^{iu} for real u, we obtain

$$\begin{aligned}\mathrm{E}\left[e^{iuz(n)} \middle| z(n)\right] = 1 + iu\mathrm{E}[\Delta z(n)|z(n)] - \frac{u^2}{2}\mathrm{E}\left[\Delta z^2(n) \middle| z(n)\right] \\ +k|u|^3\mathrm{E}\left[|\Delta z(n)|^3 \middle| z(n)\right],\end{aligned}$$

where $k \leq 1/6$; thus

$$\begin{aligned} h(u) &= \mathrm{E}\left[e^{iuz(n+1)}\right], \\ &= \mathrm{E}\left[e^{iuz(n)}\mathrm{E}\left(e^{iu\Delta z(n)}\Big| z(n)\right)\right], \\ &= h(u) + iu\mathrm{E}\left[e^{iuz(n)}\mathrm{E}\left\{\Delta z(n)|z(n)\right\}\right], \\ &- \frac{u^2}{2}\mathrm{E}\left[e^{iuz(n)}\mathrm{E}\left\{\Delta z^2(n)\Big| z(n)\right\}\right] + k|u|^3\mathrm{E}\left[ke^{iuz(n)}\mathrm{E}\left\{|\Delta z(n)|^3\Big| z(n)\right\}\right]. \end{aligned} \tag{37}$$

Cancelling $h(u)$ and dividing by u, results

$$\begin{aligned} &i\mathrm{E}\left[e^{iuz(n)}\mathrm{E}\left\{\Delta z(n)|z(n)\right\}\right] - \frac{u}{2}\mathrm{E}\left[e^{iuz(n)}\mathrm{E}\left\{\Delta z^2(n)\Big| z(n)\right\}\right] \\ &\quad + k|u|^2\mathrm{E}\left[ke^{iuz(n)}\mathrm{E}\left\{|\Delta z(n)|^3\Big| z(n)\right\}\right] = 0. \end{aligned}$$

Thus, using estimates of Lemma 5, we have

$$\begin{aligned} &iaw'(p^*)\mathrm{E}\left[e^{iuz(n)}z(n)\right] - \frac{u}{2}\tilde{S}(p^*)\mathrm{E}\left[e^{iuz(n)}\right] + \mathrm{E}\left[o(a)\right] \\ &\quad + u\mathrm{E}\left[o(a)\right] + u^2\mathrm{E}\left[o(a)\right] = 0. \end{aligned}$$

From Eqs. (15) and (17), it is evident that $\mathrm{E}[|z(n)|] < \infty$ when a is small or

$$aw'(p^*)\frac{dh(u)}{du} - \frac{u}{2}\tilde{S}(p^*)h(u) + \mathrm{E}\left[o(a)\right] + u\mathrm{E}\left[o(a)\right] + u^2\mathrm{E}\left[o(a)\right] = 0.$$

Dividing the above equation by $aw'(p^*)$ and using fact $w'(p^*) < 0$, we obtain

$$\frac{dh(u)}{du} + u\frac{\tilde{S}(p^*)}{2\,|w'(p^*)|}h(u) + \epsilon(u) = 0,$$

where

$$\varphi = \sup_u \frac{|\epsilon(u)|}{1+u^2} \to 0,$$

as $a \to 0$. Since $h(0) = 1$, it follows that

$$h(u) = e^{-\frac{(u\sigma)^2}{2}}\left(1 - \int_0^u e^{\frac{(ux)^2}{2}}dx\right),$$

where $\sigma^2 = \frac{\tilde{S}(p^*)}{2|w'(p^*)|}$. But we have

$$\left| \int_0^{|u|} e^{\frac{(ux)^2}{2}} \epsilon(u) dx \right| \leq \varphi \int_0^{|u|} e^{\frac{(ux)^2}{2}} \left(1 + x^2\right) dx \to 0,$$

as $a \to 0$; thus

$$h(u) \to e^{-\frac{(u\sigma)^2}{2}}.$$

Then using the facts that each characteristic function determines the distribution uniquely and $h(u)$ is characteristic function of $N(0, \sigma^2)$, thus we obtain

$$z(n) \sim N(0, \sigma^2),$$

and hence the theorem. □

5.7 Proof of Theorem 3

In the equilibrium state, the average penalty rates for both actions are equal or $f_1(p^*) = f_2(p^*)$, which results $c_1\pi^* = c_2(1 - \pi^*)$. Thus we have

$$\pi^* = \frac{\delta}{\delta + P_C}, \tag{38}$$

where $\delta = \text{Prob}\left[\hat{B}_h < p_h\right]$. Thus average number of blocked new calls, $\bar{N}_n$, is equal to

$$\begin{aligned} \bar{N}_n &= \lambda_n \left[1 - \pi^*(1 - P_C)\right], \\ &= \lambda_n (1 + \delta) \frac{P_C}{P_C + \delta}. \end{aligned} \tag{39}$$

Computing derivative of $\bar{N}_n$ with respect to δ results

$$\begin{aligned} \frac{\partial \bar{N}_n}{\partial \delta} &= -\lambda_n \frac{P_C(1 - P_C)}{(P_C + \delta)^2}, \\ &< 0. \end{aligned} \tag{40}$$

Thus $\bar{N}_n$ is a strictly decreasing function of δ. Since the adaptive UFC algorithm gives the higher priority to the hand-off calls, it attempts to minimize the dropping probability of hand-off calls. Using this fact and Eq. (40), it is evident that $\bar{N}_n$ is minimized which results in minimization of the blocking probability of new calls and hence the theorem. □

References

1. Ramjee R, Towsley D, Nagarajan R (1997) On optimal call admission control in cellular networks. Wirel Netw 3:29–41
2. Hong D, Rappaport S (1986) Traffic modelling and performance analysis for cellular mobile radio telephone systems with prioritized and non-prioritized handoffs procedure. IEEE Trans Veh Technol 35:77–92
3. Haring G, Marie R, Puigjaner R, Trivedi K (2001) Loss formulas and their application to optimization for cellular networks. IEEE Trans Veh Technol 50:664–673
4. Beigy H, Meybodi MR (2004) A new fractional channel policy. J High Speed Netw 13:25–36
5. Yoon CH, Kwan C (1993) Performance of personal portable radio telephone systems with and without guard channels. IEEE J Sel Areas Commun 11:911–917
6. Guern R (1988) Queuing-blocking system with two arrival streams and guard channels. IEEE Trans Commun 36:153–163
7. Li B, Li L, Li B, Sivalingam KM, Cao X-R (2004) Call admission control for voice/data integrated cellular networks: performance analysis and comparative study. IEEE J Sel Areas Commun 22:706–718
8. Chen X, Li B, Fang Y (2005) A dynamic multiple-threshold bandwidth reservation (DMTBR) scheme for QoS provisioning in multimedia wireless networks. IEEE Trans Wirel Commun 4:583–592
9. Beigy H, Meybodi MR (2005) A general call admission policy for next generation wireless networks. Comput Commun 28:1798–1813
10. Beigy H, Meybodi MR (2004) Adaptive uniform fractional channel algorithms. Iran J Electr Comput Eng, 3:47–53
11. Beigy H, Meybodi MR (2005) An adaptive call admission algorithm for cellular networks. Electr Comput Eng 31:132–151
12. Beigy H, Meybodi MR (2008) Asynchronous cellular learning automata. Automatica 44:1350–1357
13. Beigy H, Meybodi MR (2011) Learning automata based dynamic guard channel algorithms. J Comput Electr Eng 37(4):601–613
14. Baccarelli E, Cusani R (1996) Recursive Kalman-type optimal estimation and detection of hidden markov chains. Signal Process 51:55–64
15. Baccarelli E, Biagi M (2003) Optimized power allocation and signal shaping for interference-limited multi-antenna ad hoc networks, vol 2775 of Springer lecture notes in computer science, Springer, pp 138–152
16. Beigy H, Meybodi MR (2009) Cellular learning automata based dynamic channel assignment algorithms. Int J Comput Intell Appl 8(3):287–314
17. Srikantakumar PR, Narendra KS (1982) A learning model for routing in telephone networks. SIAM J Control Optim 20:34–57
18. Nedzelnitsky OV, Narendra KS (1987) Nonstationary models of learning automata routing in data communication networks. IEEE Trans Syst Man Cybern. SMC–17:1004–1015
19. Oommen BJ, de St Croix EV (1996) Graph partitioning using learning automata. IEEE Trans Comput 45:195–208
20. Beigy H, Meybodi MR (2006) Utilizing distributed learning automata to solve stochastic shortest path problems. Int J Uncertain Fuzziness Knowl Based Syst 14:591–615
21. Oommen BJ, Roberts TD (2000) Continuous learning automata solutions to the capacity assignment problem. IEEE Trans Comput 49:608–620
22. Moradabadi B, Beigy H (2014) A new real-coded Bayesian optimization algorithm based on a team of learning automata for continuous optimization. Genetic programming and evolvable machines 15:169–193
23. Meybodi MR, Beigy H (2001) Neural network engineering using learning automata: determining of desired size of three layer feedforward neural networks. J Fac Eng 34:1–26
24. Beigy H, Meybodi MR (2001) Backpropagation algorithm adaptation parameters using learning automata. Int J Neural Syst 11:219–228
25. Oommen BJ, Hashem MK (2013) Modeling the learning process of the teacher in a tutorial-like system using learning automata. IEEE Trans Syst Man Cybern Part B Cybern 43(6):2020–2031
26. Yazidi A, Granmo OC, Oommen BJ (2013) Learning automaton based on-line discovery and tracking of spatio-temporal event patterns. IEEE Trans Syst Man Cybern Part B Cybern 43(3):1118–1130
27. Narendra KS, Thathachar KS (1989) Learning automata: an Introduction. Printice, New York

28. Srikantakumar P (1980) Learning models and adaptive routing in telephone and data communication networks. PhD thesis, department of electrical engineering, University of Yale, USA
29. Norman MF (1972) Markovian process and learning models. Academic Press, New York
30. Mood AM, Grabill FA, Bobes DC (1963) Introduction to the theory of statistis. McGraw-Hill