Research Article
Volume 9:3, 2020
DOI: 10.37421/Jtsm.2020.9.303

Telecommunications System & Management

ISSN: 2167-0919

Open Access

Smart cell selection method for Femtocell Networks

Indrajit Kumar Paul[1*], Mitali Halder[2]

[1]MN HetRAN RD&Pz F&M AP GW Vrf, Nokia Networks Karnatka ,India

[2]Dept. of computer Science Presidency University Karnatka, India

Abstract

In a femtocell network , which is configured as open access, a user from a neighboring cell preferably from a different type of cell eg.(Macro, Pico or micro cell), can make handover to the femtocell network through handover for better coverage and enhance there channel capabilities for better user experience. To avoid any disruption of service for users, which can happen because of ping-pong HO (handover) it is mandatory to have a effective cell selection method should be in place. In Traditional approach this cell selection method uses RSSI /RSRP value obtained by measurement report, cell load, channel quality etc. to make decision for cell selection for HO. However problem with traditional based approach is that present measured performance does not necessarily reflect the future performance, thus the need for some kind of smart cell selection that can predict the horizon. Subsequently, we present in this paper a reinforcement learning (RL), i.e Q-learning algorithm, as a generic solution for the cell selection problem in a femtocell network.

Keywords: cell selection • Femtocell • handover • AI, Reinforcement learning • radio resource optimization • Q- learning.

Introduction

With femtocell network, Mobile network operators provide a wide range of interesting solutions that maintains a good indoor coverage as well as capacity, with BS (basestation) which are of low cost and with low power consumption. When a user who is latched to Macro base station found a open access femtocell network, user may leave its macro network and switch to a femtocell network, if in user neighbor there are large number of neighboring femtocells available, which is the case in general, then user has lot of choice for HO, so in this situation it is having worth to find a method of cell selection during HO, so that user can be benefitted with capacity and also to eliminate redundant HO, which in turn will save resource.

In literature different metrics for handover has been proposed like SINR[1], RSRP/RSSI[2], BER[3] etc. as per previous work like [4] capacity based handover performs better than SINR based handover.

Problem with previous work capacity based handover like [4] is that at we can only be aware of capacity at present time t, but we cannot guarantee that this capacity will be maintained in the future, because the measure of the gained capacity at time t does not reflect the behaviour of the target cell at time t + delta t. there is a possibility that the capacity may degrade or increase because of any reason like channel condition or user location like (cell edge or cell center). However there may be situation when we can observe huge degradation of capacity in situation like when user is at cell edge, and this can lead to ping-pong HO situation, which are not desirable as it will impact user experience, also this redundant HO situation like ping-pong HO situation will lead to unnecessary signaling load ad HO needs lot of signaling as well. Hence it is very important to have HO procedure as less frequently as possible, for which we need to ensure a HO decision which give better capacity and that too for a sufficient long duration. To achieve this requirement we propose a Q-learning algorithm (Reinforcement learning) based cell selection method for a open access femtocell n/w, which can make sure a cell selection which is more robust in terms of capacity and duration, it will be achieved by the algorithm by analyzing the historical behavior of different cell and predicting the behavior of cell in upcoming future.

In the classical Reinforcement Learning frame work, [5] agent communicates with environment and learn the optimal action. On every step agent takes a action and get a reward or penalty based on policy, which define how good and bad the action is and get a new state, and these values are used to update the Q table, which in turn helps the RL model to know which is the optimal action it should take to get maximum rewards. Q-learning directly learns the optimal policy. because the estimate of q-value is updated on the basis of 'the estimate from the maximum estimate of possible next actions', regardless of which action you took.

In most real-world scenarios, the model is not available. You have to implicitly infer the model from the observations. Methods designed to solve these scenarios are called model- free methods , Since our objective is that the user from Macro Base station wants to maximize its capacity for longer duration , the solution of this problem comes with model free frame work of Reinforcement Learning, as in model free method , since the Monte-Carlo methods also known as Model free method require only knowledge base (history/past experiences) sample sequences of (states, actions and rewards) from the interaction with the environment, and no actual model of the environment, hence user(agent) takes advantage of the RL algorithm to estimate the efficiency of neighboring femtocells based on their past behavior.

*Address for Correspondence: Indrajit Kumar Paul, MN HetRAN RD&Pz F&M AP GW Vrf, Nokia Networks Karnatka ,India, Email: Indrajit.paul@nokia.com

Received August 14, 2020; Accepted August 28, 2020; Published September 4, 2020

System Model And Problem Statement

System Model:

We have considered a Macro base station which is having Femtocell as its neighbour, where macro users nMU are connected to macro base station and femto user nFU are connected to Femtocell base station. All base stations and UE (user equipment) are with one antenn, also we assume that the femtocell networks are deployed as open access mode for better utilization of spectrum. Since TDD sync is assumed to be perfect hence assuming TDD for femtocell and macro link.

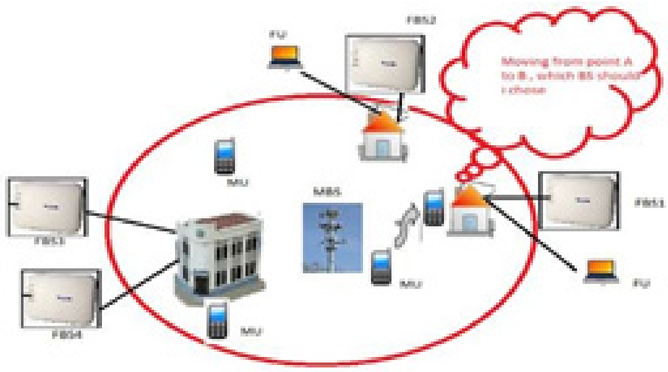

Figure 1. System model

ProblemStatement

In a large-scale femtocell n/w, which is highly dense, the main problem which operator face is of handover and cell selection is the most crucial step in handover process, since a bad selection of cell will lead to multiple redundant handover, which in turn will result is bad user experience and unnecessary wastage of radio resource. Hence need arises for a smart method of cell selection, which can predict best cell for cell selection based on past experience. To achieve this objective we are proposing to use RL frame work, since RL help to find optimal policy for the agent to take action which give maximum reward , as in this case we want cell to be selected for macro user which gives maximum capacity to user.

In this study the RL model free algorithm which we are suggesting is Q-learning, In Q-learning, you do max-update. After taking action 'a' from state's', you get an immediate reward of r and land to state s'. From s', you take the most greedy action, i.e., action with highest q-value, since Q learning method is a online method and this property of Q learning makes it a perfect solution for our cell selection problem, as we need to perform the cell selection in real time.

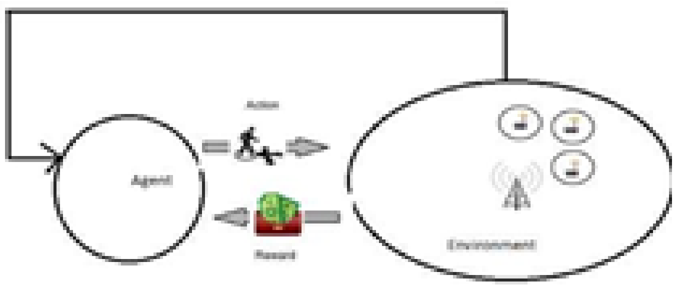

Figure 2. RL framework

Proposed Solutions

General Solution:

In a femtocell network , as user has the capability of searching neighbor cells and it can also distinguish between accessible femtocell base station with non accessible ones. A macro user can try to enhance its capacity by making handover from serving cell (MBS / FBS) to femtocell base station. To make this handover macro user should make decision of cell selection either select best FBS or remain connected to MBS. Solution can be in 3 major steps:

1. Based on LMMSE estimator macro user collects channel gain information between macro and FBS base station.
2. With the help of Q-learning RL algorithm, macro user can predict best cell.
3. Macro user join serving cell

Q-learning Algorithm

In Q-learning, agent do max-update. After taking action 'a' from state's', you get an immediate reward of r and land to state s'. From s', you take the most greedy action, i.e., action with highest q-value. basically in Q-learning agent tries to find a optimal policy which gives maximum Q value (see figure 2.)

Lets define all parameter related to Q-learning:

- Environment : it means every thing except agent , basically agent observe the environment , in our problem environment is femtocell base station & macro base station in macro user neighbour list.
- Agent: it is the basically the one who takes action, in our case agent is macro user who wants to perform a handover for which he need to perform cell selection.
- State: it is the situation of environment which is currently at, in our case it is current serving cell of macro user. state is defined as S = {s = 1, 2, ..., N + 1} where N represent number neighbour femtocell is present.
- State: it is the situation of environment which is currently at, in our case it is current serving cell of macro user. state is defined as S = {s = 1, 2, ..., N + 1}, where N represent number neighbour femtocell is present.
- Action: when an agent takes a decision , it is termed as action , in our scenario it is to identify the target cell ,the macro user can remain connected to macro base station or chose any femtocell basestation from its neighbour cell. Action is defined as A = {a = 1, 2, ..., N + 1}
- Reward: it is a parameter, which defines how good the action was , which agent has taken.in our problem reward is the gain which macro user has got after performing cell selection if gain is good agent will get +ve reward and if it bad then agent will be penalized with -ve reward.

The objective of this Q-learning based algorithm is to obtain a policy that help us achieve maximum discounted cumulative reward.

$$\max \pi \sum_{t=0}^{\infty} \gamma t \; r^{t}$$

where γ is the discount factor ($0 < \gamma < 1$) and rt is the received reward at time t. For $\gamma = 0$ future rewards have no impact on the state value, while for γ close to 1, future actions are considered as important as the immediate rewards. For a given policy Pi we define a Q- value as:

$$Q(s, a) = R(s, a) + \gamma \sum_{v \in c} P_{s,v}(a) Q(v,b)$$

- R(s, a) is the expected reward agent will get when the environment is in the state s and the agent executes the action a:
- Ps,v(a) is the transition probability from the initial states to the new state v as a result of the action a.
- Q(v, b) is the Q-function of the next state- action pairs.

With the help of Bellman optimal equation, Value of a state under an optimal policy must be equal to the expected total return for the best action from that state. At each state, greedily pick the best action that has the maximum q- function.

Proposed Solution:

Solution which we are proposing is , single agent multiple state RL frame work , where we are considering that one macro user is doing HO scenario and it is having multiple option of cell where it can make HO { in case we want to consider a scenario where multiple HO is happening simultaneously , in that scenario we will use multiple agent multiple state RL frame work }, hence macro user need to take a decision which cell to select to achieve maximum capacity, this selection problem can be solved by the use of single agent and multiple state RL model. In our scenario Q(s.a) is our system where r(s,a) is the expected reward which our agent in our case its macro user , will achieve after performing action and moving to each state action pair (s,a).

$$Q(s,a) = E[r(s,a)]$$

In our scenario, as we have a very dynamic Environment, with so many neighbor cell, we will not be having the exact information that the reward received is from which cell. Therefore we define a q function $Q_t(s,a)$ which represent Q(s,a) at time t, basically it takes average of all rewards which has been received in time t.

Reward function: Now lets define our reward function r. Since our ultimate objective is to maximise and maintain the capacity for our macro user, which is performing handover and joins a new cell. There can be 2 possible rewards.

- if MU chooses the MBS as a serving cell, r is expressed as (1).
- if MU chooses to join one of the FBSs in its NCL, r is expressed as (2)

Action selection strategy: To learn Q function Q(s,a) various approach can be taken , we are proposing ϵ-greedy approach for this cell selection .

Policy improvement is done by constructing an improved policy p' as the ϵ -greedy maximisation with respect to $Q_p(s,a)$. So, p'(a|s) is:

$$\pi'(a|s) = \begin{Bmatrix} 1 - \epsilon + \frac{\epsilon}{|A(s)|} & a = \arg\max_{a'} Q_\pi(s,a') \\ \frac{\epsilon}{|A(s)|} & otherwise \end{Bmatrix}$$

Where |A(s)| is the action space, i.e. total possible actions in any given state and ϵ is a hyperparameter

that controls the trade-off between exploration and exploitation.

- If ϵ is too small, then actions are biased to be more greedy
- If ϵ is too large, then actions explore more.

To select the best cell, in exploitation phase i.e 1-ϵ proportion of trials, where as when value of ϵ is between 0 & 1, we are exploring and hence we will perform more trials.

Performance Evaluation

Now we will evaluate the performance of proposed method for one macro user, who is trying to perform handover to select an optimal cell in open access femtocell network, in this simulation we are considering a person walking is a macro user at a speed of 1 m/s, considering Macro user is initially attach to Macro base station, the considered path loss are similar to [6], simulation

TABLE I SIMULATION PARAMETERS

Parameter	Value
Carrier Frequency	2000 MHz
System Bandwidth	10 MHz
Number of Paths	6
Time Sampling	1/1000 s
Macrocell Size	1000 m
Femtocell Size	30 m
Transmit Power at MBS	46 dBm
Transmit Power at FBS	20 dBm
FBSs Distribution	random
Transmit Power at User Equipment(UE)	23 dBm
Noise Power at each UE	-174 dBm
Noise Figure at UE	9 dB
Total Number of MUs: NMU	10
Number of MUs Performing a Handover	1
Number of FUs at each FBS: NFU	Poisson distribution (π = 2)
Total Number of FBSs: F	30
Outer Wall Loss	15 dB
α	0.5
γ	0.9
ϵ	0.1

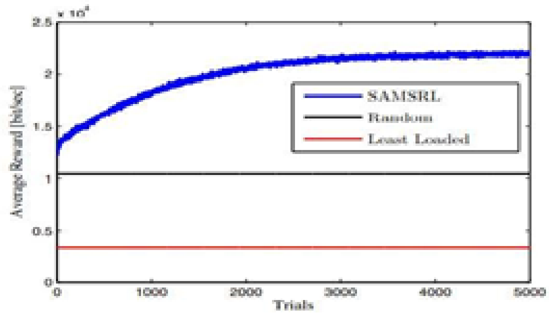

Figure 3. Comparison of SAMSRL, random and Least Loaded

Simulation result

We are comparing different cell selection method found in literature i.e least load & capacity based method, with our proposed model i.e Single Agent multiple state RL(SAMSRL) method, to evaluate the performance of our model, as shown in fig.2 we had compared the performance of our model, with other cell selection method i.e LL & random [5], for this comparison we have 1st define the load in the cell, by declaring number of users in cell, then for each algorithm we have executed our model for 2000 times and performed average 2000 handover, when plotting the capacity received by macro user . We clearly see that the model proposed by us, perform better than legacy Least loaded and Random algorithm [5].

When only one handover is carried out: after performing single handover for a macro user, we have compared the capacity gain achieved by macro user with proposed model wrt. gain achieved by conventional method like [4]

We can see that initially conventional method outperformed the proposed model, but it is for very small duration and as discussed earlier our objective is to find a cell , which gives maximum capacity for longer duration , which is achieved by our

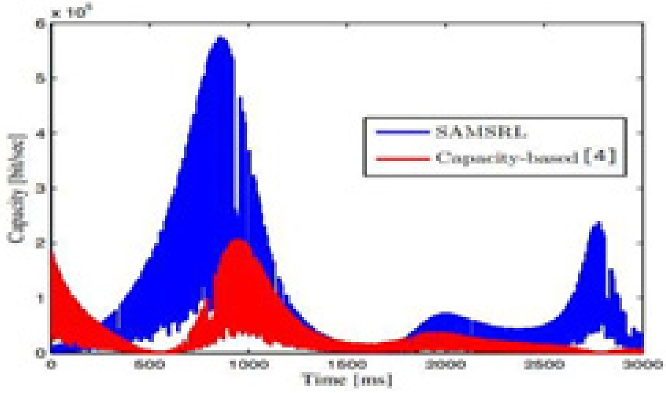

Figure 4. capacity based comparison

With the above comparison its clear that our SAMSRL model is performing better than any conventional method of cell selection.

When more than one handover is carried out: now we extend our simulation testing, to evaluate if our SAMSRL method has done any impacts on handover, since our final objective was to reduce handover, which will in turn help us to reduce signaling on our base station and help to improve user experience. for better clarify we have given comparison of conventional (capacity based [4]) and SAMSRL in different figures. While testing in our simulator, we have set the condition that when ever threshold is goes below 103 bits/sec macro user will trigger handover.

From fig. 5(a) & 5(b) we can see that in conventional method handover triggered twice, this is due to reason that cell selection happened with conventional method , after some duration channel quality of the cell degraded, where as in proposed SAMSRL method no handover triggered, which signifies that the cell selected by proposed model channel quality remains good for longer duration as compared to cell selected by conventional (capacity b

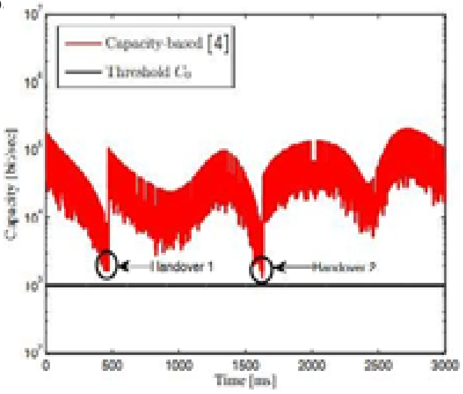

Figure 5(a). number of handovers in capacity based[4]

To be more sure for this behavior that with proposed model number of handover happening is less than the conventional (capacity based [4]) method, we have perform 40 runs of scenario and verified the average number of handover happened with both

conventic d SAMSRL model .

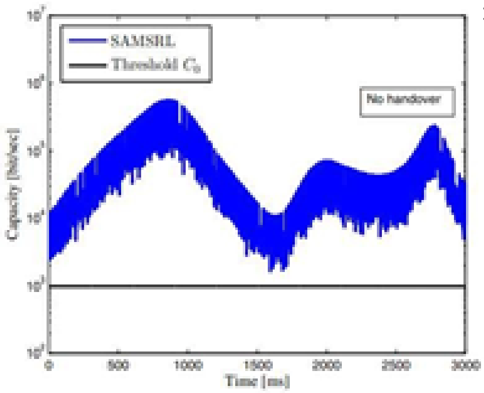

Figure 5(b). number of handover in SAMSRL

in Fig. 6 result has been shown , where we can clearly see that there is less number of HO happened with proposed model as compare to conventional method(Capacity based [4]).However with this long run we can see handover happened with proposed

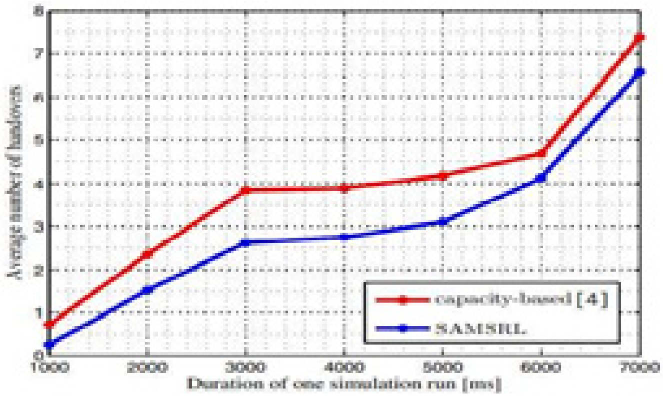

Figure 6. Average number of handover happened with different method

Summary of findings

from the above simulation results, we hereby conclude that the results achieved from proposed model (SAMSRL) is almost two folds better than conventional method i.e (Least load , random & capacity-based):

- Improvement in capacity, which result in channel stability.
- Reduction of signalling, since less handover will be required with the proposed model.

Conclusion

In this research paper we propose a smart cell selection model for open access femtocell network. This model is based on Q- learning algorithm of Single Agent Multiple State RL frame work for cell selection to make sure that the macro user will get higher capacity, after performing handover from serving cell to its neighboring femtocell network. Based on the result we have achieved with simulator, it is clear that the proposed model SAMSRL will enhance user experience, by reducing ping

-pong handover situation and also enhance resource utilization, since it will reduce signaling. Multiple Agent Multiple state RL i.e (MAMSRL) can be used to model when multiple user decide to perform handover simultaneously and this scenario we can consider as a future research work.

References

1. E. Frech and C. Mesquida, "Cellular Models and Hand-Off Criteria," VETEC 1 (1989) : 128-135.

2. J. Holtzman and A. Sampath, "Adaptive Averaging Methodology for Handoffs in Cellular Systems," IEEE Trans. Veh. Technol 44 (1995) : 59–66

3. K. Cornett and S. Wicker, "Bit Error Rate Estimation Techniques for Digital Land Mobile Radios," VETEC (1991) : 543-548.

4. H. Mahmoud, I. Gu andvenc, and F. Watanabe, "Performance of Open Access Femtocell Networks with Different Cell-Selection Methods," VTC Spring (2010) : 1-5

5. L. Pack Kaelbling, M. Littman, and A. Moore, "Reinforcement Learning: A Survey," Journal of Artificial Intelligence Research 4 (1996) : 237–285.

6. FemtoForum, "Interference Management in OFDMA Femtocellsr" (2010)

How to cite this article: Indrajit Kumar Paul. "Smart cell selection method for Femtocell Networks" J Telecommun Syst Manage 9(2020): ** DOI: 10.37421/Jtsm.2020.9.303