Wireless Pers Commun (2011) 57:73–87
DOI 10.1007/s11277-010-0007-7

Cognitive Radio with Reinforcement Learning Applied to Multicast Downlink Transmission with Power Adjustment

Mengfei Yang · David Grace

Published online: 22 April 2010

Abstract This paper shows how channel assignment in multicast terrestrial communication systems with distributed channel occupancy detection can be improved using intelligence based on reinforcement learning and transmitter power adjustment. It is shown how such schemes greatly reduce the number of reassignments and improve the dropping probability, at the expense of increased blocking. It is found that using different minimum quality of service threshold percentages can partly control and improve the performance, in place of the more traditional SINR threshold levels. The paper also shows how a power adjustment technique is developed which significantly reduces the level of overlap between adjacent base stations, and further reduces interference and transmitter power.

Keywords Cognitive radio · Reinforcement learning · Multicast · Distributed sensing · Power adjustment

1 Introduction

Cognitive radio has been suggested as a new way to implement efficient reuse of the pooled radio spectrum assigned to multiple wireless communication systems, by exploiting a wide variety of intelligent behavior [1]. Cognitive radio detects the unused portion of spectrum, and then uses this band taking the required steps to avoid interference with future licensed users [2]. Radios can monitor the spectrum and choose frequencies that minimize interference to existing communication activity, in order to optimize the usage, especially the reuse of spectrum [3]. Multicasting, especially on the downlink, will be an important feature of future pooled spectrum systems instead of unicast, and channel assignment schemes are one of the important tasks for controlling the efficiency of multicasting spectrum usage. Generally,

M. Yang (✉) · D. Grace
Communication Research Group, Department of Electronics, University of York, York YO10 5DD, UK
e-mail: my505@ohm.york.ac.uk

D. Grace
e-mail: dg@ohm.york.ac.uk

channel assignment can be divided into fixed channel assignment (FCA), dynamic channel assignment (DCA), which is often used for cognitive radio, and hybrid channel assignment (HCA) [4–7]. Here we use a user population based DCA optimization method to determine the most suitable channels, which uses the statistics of the SINR obtained from multiple users to select the best channel for different base station combinations. The scheme used here selects a channel based on serving the different percentages of a user population across the coverage area [8]. Distributed detection is used here, which infers information from the environment or users instead of exchanging information between BSs.

Reinforcement learning (RL) is a method to describe the behavior of an agent that learns through trial-and-error interactions with a dynamic environment so as to maximize some notion of long-term reward [9]. It will learn the information based on the external environment and previous states, which then influences the current activation [1]. The weight is used to show the influence from the previous users or the factors based on circumstance, which are then updated on each activation. In this paper, we implement this computational method by using weights to give the positive or negative feedback for each base station on each channel.

Cognitive radios have the capability to adjust their transmission parameters to achieve better performance [2]. The transmission parameters that may be adjusted to improve communication quality include: operating frequency, modulation scheme and transmit power [10,11]. In this paper, the transmit power will be adjusted to reduce overlap between neighboring base stations (which operate on different channels) thereby saving power and further reducing the interference.

Therefore the main aim here is to apply reinforcement learning to the channel assignment process of multicast communication systems, which operate with downlink transmitter power adjustment at the base station. This will exploit information from randomly distributed users, providing distributed detection. The assignment will use a threshold based on a quality of service guarantee across differing percentages of this user population. Previous research has adopted a SINR cdf to control the performance of a distributed channel assignment scheme with reinforcement learning [9].

This paper is organized as follows: in Sect. 2, the model of the multicast scenario and related system parameters are briefly overviewed. Section 3 introduces the distributed channel assignment scheme and their characteristics, followed by the reinforcement learning rules. The results arising from different user populations influencing the distributed reinforcement learning schemes are then analyzed and discussed in Sect. 4, followed by the power adjustment applied in the system. Finally conclusions are presented in Sect. 5.

2 Coexistence Scenario and System Parameters

To evaluate performance we examine a specific percentage of users being served by each base station simultaneously. A simplified test scenario is used, which is based on a number of channels with terrestrial base stations randomly distributed over the service area. The area of interest is defined by a square coverage area with sides of 10 km with test users located in random positions on a 100 m grid. Due to the random base station locations, the area of influence from each base station is unpredictable. A number of channels in this paper simulate a multi-channel terrestrial communication system, with the findings generally applicable to a frequency bands below 6 GHz. The numbers of base stations and channels can be changed flexibly with different requirements of the terrestrial system. Here 1,000 users are randomly distributed over the coverage area, and a subset of users are randomly chosen from this set of users.

The received power for user P_s is the significant factor for obtaining SINR (which will be described later). The propagation model implemented here is Okumura-Hata. In urban areas, P_s has the relationship with transmitted power P_t and path loss L_{dB} as below:

$$P_s = P_t / L_{dB} \tag{1}$$

and L_{dB} is calculated by

$$L_{dB} = A + B \log R - E \tag{2}$$

There are three equations to calculate the parameters for (2)

$$A = 69.55 + 26.16 \log f_c - 13.82 \log h_b$$
$$B = 44.9 - 6.55 \log h_b$$
$$E = 3.2(\log(11.75 h_m))^2 - 4.97 \text{ for large cities}$$

where f_c is the carrier frequency, h_b and h_m are the base station antenna height and mobile station antenna height above the local terrain height. Where f_c is the carrier frequency, h_b and h_m are the base station antenna height and mobile station antenna height above the local terrain height, and R is the distance between transmitter and receiver. According to (1), the received power at a user decreases with the increasing distance from the base station (BS). To determine the performance of the system, it is necessary to determine the Signal-to-Interference-plus-Noise Ratio (SINR).

$$SINR = \frac{P_s}{P_n + \sum P_i} \tag{3}$$

where P_s is the wanted signal power and P_n is the noise power, $\sum P_i$ is the summation of all the interfering base stations on the same channel. Both P_s and P_i are determined using (1). We assume that the base stations, that operate on different distributed channels will not influence the SINR value of other base stations, i.e. adjacent channel interference is negligible. In order to ensure a sufficiently high SINR value, a threshold is used as a method of evaluating the appropriate interference between each user and finally determining the performance.

The noise power used to form part of SINR can be expressed by [12]:

$$P_n = 10 \log(FkTB) \tag{4}$$

where F is the noise factor, k is Boltzmann's constant, T is the ambient temperature and B is the bandwidth. Other parameters of the terrestrial system are shown in Table 1.

3 Channel Assignment Schemes with Reinforcement Learning

An example of a multicast scenario is shown below in Fig. 1. The randomly located users connect to the nearest base station. A certain minimum percentage of the users have to receive SINR for each base station to be successfully assigned on the channel. So the user population here represents different proportions of the users which satisfy the SINR requirements. For example, SINR values of more than 95% of the users should be greater than 4.3 dB. We choose 90 and 95% of the user populations satisfying this minimum performance criterion in this paper, which is typical in terms of coverage based quality of service.

The distributed schemes here do not require a central controller; instead they infer information from the environment or users instead of exchanging information between BSs. Although

Table 1 System parameters

Parameter	Value
Service area	10km × 10km
Transmitter height	15 m
Transmit power	21 dBm
BS antenna gain	10 dB
User antenna height	1.5 m
User antenna gain	1 dB
Antenna efficiency	100%
Bandwidth	3 MHz
Frequency	900 MHz
Noise power	−102.7 dBm

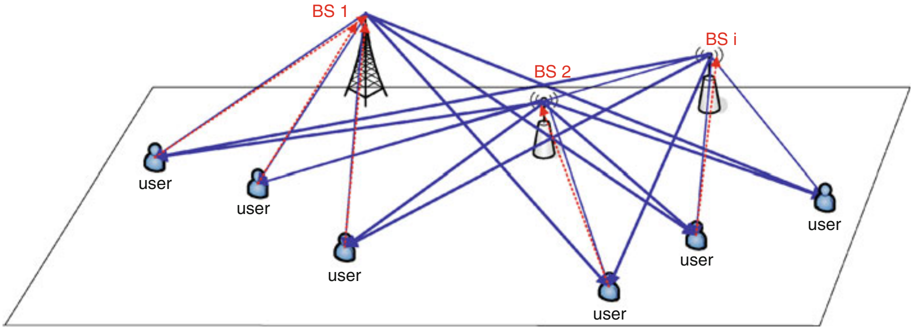


Fig. 1 The multicast scenario

the distributed schemes may be more vulnerable to shadowing and the ‘hidden node problem’ [13], they are more scaleable and should deliver a system with lower node complexity. A scheme, which we refer to here as the random picking distributed scheme, applies the reinforcement learning to select the most appropriate available channel.

Applying reinforcement learning to the distributed channel assignment schemes is aimed at improving the performance of more conventional schemes by using previously obtained knowledge to aid future decisions. This is shown to further improve the assignment stability and general performance of the cognitive radio system [14–16]. The following content briefly describes the basic rules of the distributed reinforcement learning algorithm used here: Initially, all channels start with an equal weighting, but after each activation, the weights are updated on a channel according to the conditions shown in Table 2—successful or unsuccessful channels will result in different positive or negative changes of the weights. That means the weights are increased or reduced on each activation. On the next activation, the BSs will choose the channel with the highest weight as the highest priority channel to attempt to be assigned using any channel assignment schemes as required. So reinforcement learning will help the channels to find the most appropriate way to be assigned and maximally reduce the interference and significantly avoid the collisions. The weighting values can be represented by Eq. 5 based on [17,18].

$$W_i = FW_{i-1} + W_f \tag{5}$$

Table 2 Threshold and weighting factor

Type of BSs	Thresholds	Weights (W_f)
New		
Acceptance	T > 4.3 dB	+2
No acceptance	T <= 4.3 dB	0
Existing		
Reassignment	2.3dB =< T <= 3 dB	−1
Dropping	T < 2.3 dB	−2
Reassignment		
New acceptance	T > 4.3 dB	+1

Where W_i is the weight of current iteration after updating the information from the previous weight W_{i-1}. W_f is the weighting factor that is the weight value associated with the operation of current system and channel assignment, as shown in Table 2. F is the parameter used to adjust the proportion of the previous weight that will be considered later, and is set as 1 in this paper.

Table 2 is based on different base station assignment conditions. We define the weighting factors of the user population for different thresholds. Here 2.3 dB is used as the threshold of GMSK modulation [12] for it to be demodulated with an adequate bit error rate. For most cases we select a common 4.3 dB threshold based on the original acceptance threshold plus a 2 dB margin. here are three different types of assignment conditions. If a base station is newly accepted on a channel, the weight associated with that channel is increased by 2 (a reward), or 0 if it fails (no punishment) as shown in Table 2. If an existing base station is forced off its existing assignment we assign the channel −1 as a punishment for requiring reassignment. We use −2 if dropping occurs after failing reassignment on 5 different channels successively. For the reassignment, we define +1 as the weight increase associated with a new channel following a successful reassignment, because it is not as important as a new base stations' initial assignment, but it still needs to be rewarded for successful assignment. If a reassignment is forced due to interference then the weight associated with the channel is reduced by 1. These values at the moment have been chosen arbitrarily, and further work will be needed to choose optimal ones. In addition, the weight factor could be designed adaptively, not merely limited to integers. We will pay more attention for building a whole weight system, instead of choosing them arbitrarily.

The flowchart below shows how different types of base stations operate the distributed channel assignment scheme with reassignment, blocking and dropping phases, and the ways they update the weights. The area inside the dashed line is the step that is responsible for determining whether the channel should be assigned, reassigned, blocked or dropped. We define these combined steps here to simplify the later flow chart shown in Fig. 2.

The simulation steps of the random picking distributed scheme with reinforcement learning are briefly described below:

Step 1: Initial step
Number all the base stations and channels from 1, respectively. Keep the best 3 channel sets from the last iteration for each BS (the best 3 channel set is obtained by selecting the highest 3 weights for each testing BS corresponding to each channel)
Step 2: Channel assignment
Initial activation: All the channel weights in each base station are the same. This scheme selects 3 random channels for each base station as the best 3 channels to start.

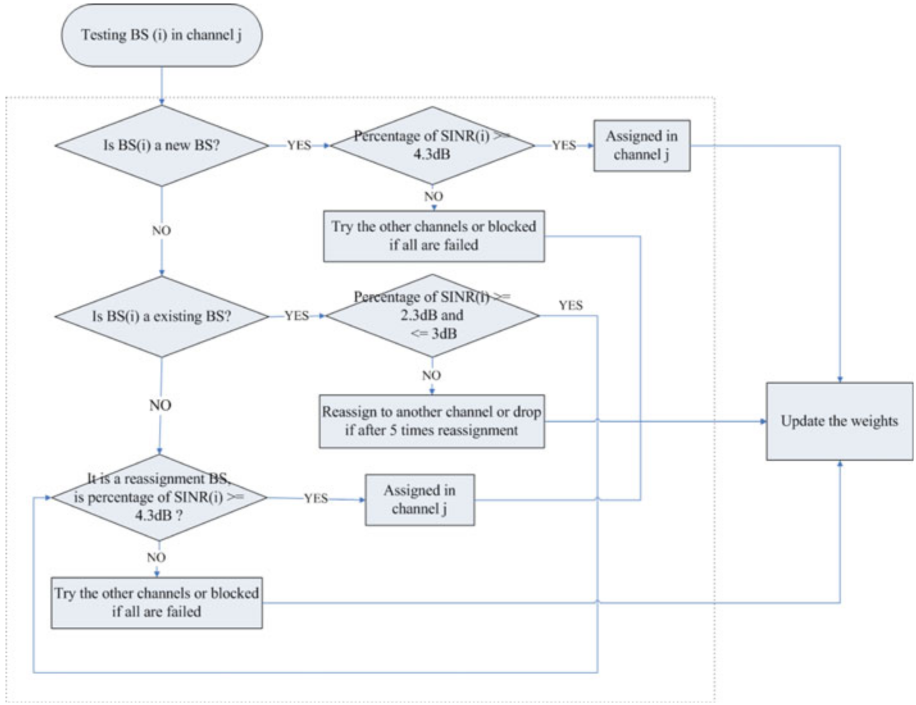


Fig. 2 Flowchart for distributed channel assignment

Subsequent activation: The best 3 channel set is used for each BS. A random channel from the best channel set is chosen, with the remaining channels from the best channel set chosen randomly if the channel(s) fail. Once the best channel set has been exhausted, the remaining channels from the pool are selected randomly. The initial assignment and reassignment will be stopped after trying five channels.
Step 3: Weight update

The weights are changed and recorded after each activation using the values in Table 2. This means the best 3 channel sets may also change after each activation.

4 Model and Results

We will discuss the effect on performance of different user populations applied in the random picking distributed channel assignment scheme, after a different number of weight update iterations. Base stations are assigned in random order, and 1,000 sets of user locations are used in order to provide an adequate number of trials for obtaining correspondingly statistically significant results. In the past, we used to select the model with 5 channels and 30 base stations [9], but due to loading quality of service dropped below that which is considered acceptable, so here a model with 10 channels and 50 base stations is used in this paper to better illustrate the channel assignment.

From the results in [17], we know that the number of users per base station play an important role in improving the performance, particularly in reducing the need for reassignment, but also dropping. In this situation, more hidden node situations can be prevented if the

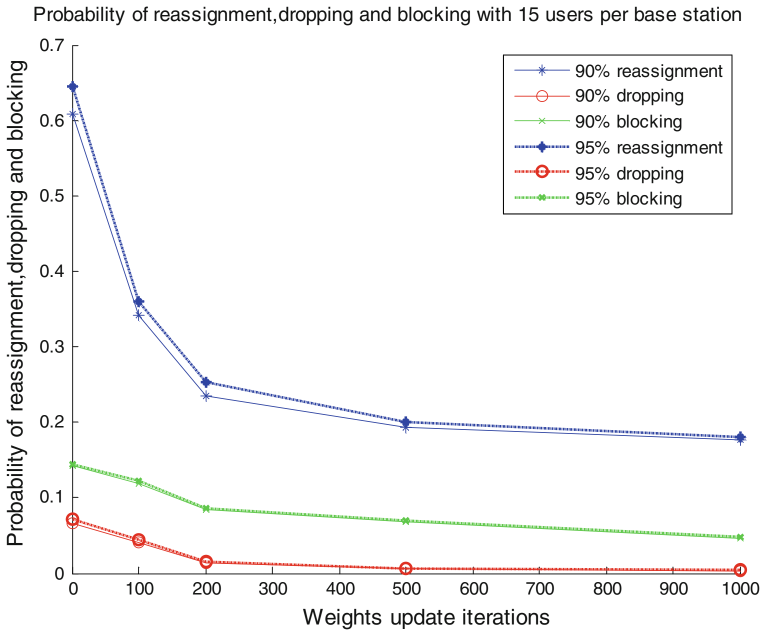


Fig. 3 Average number of reassignments, probability of dropping and blocking with 15 users per base station

number of users per base station is large enough. We use the following equation to find the suitable number of users per base station:

$$N_o = K_d \frac{S}{N_u N_b} \tag{6}$$

Where K_d is the density of users, which effectively adjusts the size of potential coverage area of a base station based on the number of users per base station, S is the total area of coverage area and N_u is the total number of users and N_b is the number base stations. So the suitable number of users is approximately 14 in this scenario here, corresponding to a 95% level when we use 15 users per base station to obtain the channel assignment results [17]. It is important that there are sufficient users to help alleviate the ‘hidden node problem’. Either a very high density or low density of users will result in the channel assignment performance being degraded. This number of users per base station is also linked to the power adjustment which will be discussed later.

In Fig. 3, the performance with respect to the number of weight update iterations at each base station is shown. That means the system is tested from no reinforcement learning applied in the distributed schemes (iteration 0), to weights obtained after 1,000 iterations. We include the average number of reassignments, the probability of dropping and blocking in one figure in order to see the proportion they contribute, and also for obtaining the whole trend. The simulation here is not a true Monte-Carlo simulation, because the weight derivation is a cumulative process that is increased or reduced after each activation, with all the results obtained from the previous results. The learning process is never stopped. It is found that with dropping and blocking, the reinforcement learning plays an important role for improving the performance, particularly in reducing the need for reassignment. The number of reassignments is significantly reduced after a large number of iterations have been used to set the channel weights. When the weight update iterations increase, not only is the reassignment probability greatly reduced, but it also improves the rate at which base stations find suitable channels for assignment. This means that the base stations are much more likely to find a suitable channel, because the high weighted channels will help the users avoid the interference. Blocking and dropping performance are also improved after 1,000 iterations; the probability of dropping decreases to a residual value, which shows that reinforcement learning when used with a multicast user population helps the system avoid dropping. The convergence rate of reinforcement learning is also an interesting direction for future work, along with applying different machine learning knowledge and algorithms.

The behavior of the assignment process in distributed channel assignment schemes can be divided into three periods (investigation period, accumulation period and mature period) based on a selection of a percentage of users across the entire coverage area, first seen in [9]. This user distributed occupancy situation is also seen in this case. The difference here is the decreasing rate of reassignment here is slightly lower, which is a result of using a limited user population, rather than obtaining performance from test users regularly spaced over the coverage area. These groups of users are not able to cover all the area, which may result in less information exchange in the system than before. As the density of users increases a greater number will be used for detection at each base station. The number per base station can be limited by defining a nominal service area, which is less than the actual coverage area (due to minimum SINR), but coverage area overlap will still exist.

Figure 4 shows the probability of users with the ability to connect to a different number of base stations. The total number of users here is 15 per base station. In the scenario tested here, nearly 90% of the users can potentially connect to 2 base stations on different channels, indicating that there is a high level of overlap. Also, the users in the central area are more frequently selected, which may cause the weights of the base stations to be influenced faster than others. The level of overlap is affected partly because of the random location of base stations and the density of users. This kind of situation needs to be avoided in the real communications system. Power adjustment will help reduce the overlap and will be explained in the next section.

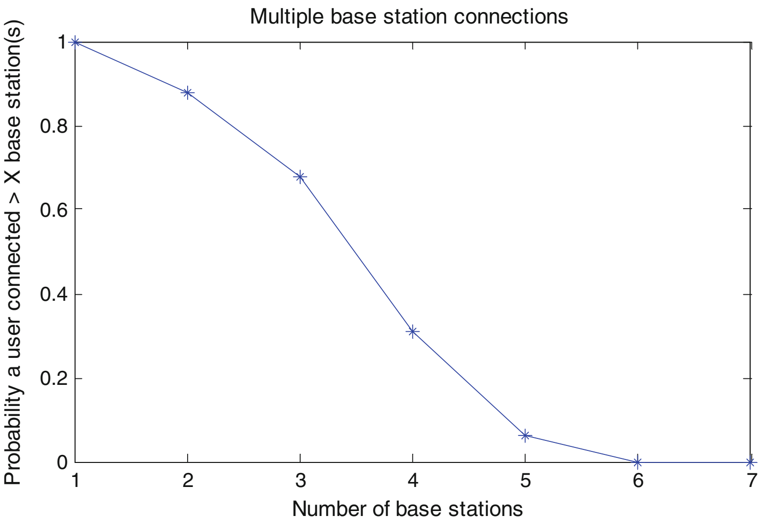


Fig. 4 Multiple base station connection for individual users

Table 3 Percentage thresholds and weighting factors

Type of BSs	Thresholds	Weights (W_f)
New		
Acceptance	95%T > 4.3 dB	+2
No acceptance	95%T <= 4.3 dB	0
Existing		
Reassignment	95%T <= 4.3 dB	−1
Dropping	85%T < 2.3 dB	−2
Reassignment		
New acceptance	95%T > 4.3 dB	+1

Now we investigate the effects of adjusting the channel assignment thresholds. We update the original thresholds of Table 2 according to Table 3, which is a new attempt to take into account the user population, and the reinforcement learning. Previously the same percentage of minimum user population served was adopted for blocking, dropping and reassignment, so here we focus more maintaining specific percentage of the user population served in each category and less on the SINR thresholds. The 4.3 dB threshold is still based on the original acceptance threshold plus a 2 dB margin, but now this is coupled with a minimum population percentage threshold. This margin is included to cope with small fluctuations in interference, e.g. as a result of new arrivals on the same channel in locations far from the coverage area, which would result in connections being reassigned or dropped. In practice, interference will also fluctuate due to other effects, such as multipath, change in transmission parameters etc. However, if the margin is too large then connections are blocked unnecessarily, hence the margin of 2 dB is a compromise figure. For dropping we set an 85% threshold at 2.3 dB instead of 95%, as a minimum level. This slightly relaxed condition still keeps the vast majority of users active in the systems, and seems a better option rather than forcibly dropping otherwise active users in the system. An alternative strategy that is beyond the scope of this paper is to keep all systems active providing some users are benefiting from service, and instead recording the connection as being disturbed (for a period).

Comparing Figs. 3 and 5, we find that using the modified percentage threshold yields better performance and is a more flexible way of controlling behavior. There are two ways to identify the thresholds, one is based on a fixed percentage of the user group and another is changing performance by selecting an appropriate percentage of the user population rather than adjust the SINR threshold value. The reassignment threshold in Table 3 is stricter than the range of SINR values in Table 2, so the base station will be reallocated more frequently and this helps dropping and blocking rate decrease. Dropping is also reduced due to the more relaxed threshold.

5 Power Adjustment

Reducing overlap is an important issue to consider further because it can result in high unwanted levels of interference. The following figure shows the spatial layout of adjacent base stations on different non-interfering channels. In order to solve the overlap problem in Fig. 4, we will adjust the power over the coverage area, assuming a fixed number of users are connected to each BS, in order to minimize the overlap and further reduce the interference.

In this base station service area with power adjustment, as an example, we assume 'BS' is the base station whose transmit power will be adjusted, BS1, BS2 and BS3 are three adjacent

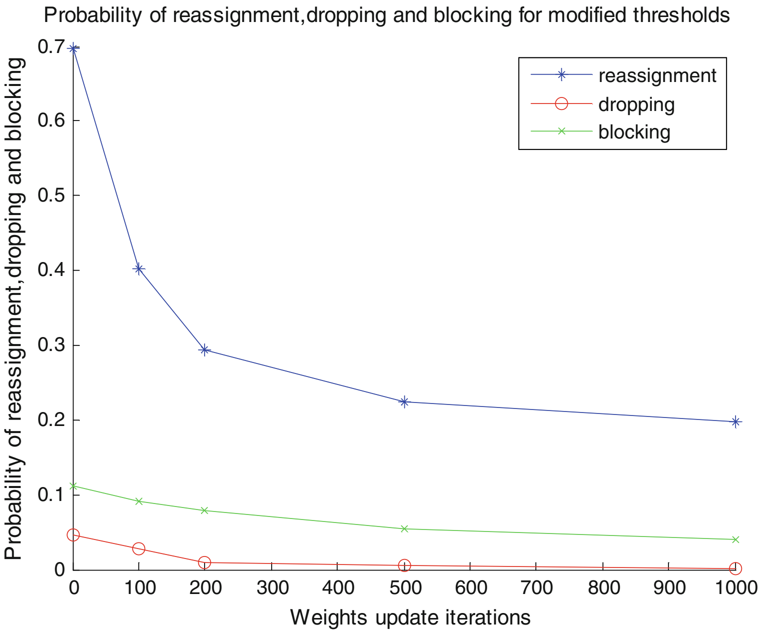


Fig. 5 Probability of reassignment, dropping and blocking with modified thresholds

BSs. The dashed circles for each base station represent different service areas, which shows that they could work for different modulation schemes as required. In the fixed transmitter power situation, the higher the SINR threshold required, the smaller service area that will be covered. Initially, the minimum percentile threshold is 2.3 dB, but this will be varied later. Due to the different locations of users for each iteration, the service area for each BS changes, meaning that the transmit power also needs to be changed each iteration. To find a suitable transmit power for each BS, we need to define a minimum receive power for the users at the boundary between different BSs, which means inside this boundary, the users will belong to this specific BS, meaning that the overlap can be reduced. The position of BS, BS1, BS2 and BS3 are shown in Fig. 6. If the service area of two base stations overlap, the users connected to each BS will decrease. Therefore, the equations for calculating the transmit power are:

$$P_{TXL'} = \max\left(S_{\min dB} + L_1, S_{\min dB} + L_2, \ldots\right) \tag{7}$$

$$P_{TXL} = \min\left(P_{TXL'}, P_{\max}\right) \tag{8}$$

$S_{\min}$ can be calculated from the equation below:

$$S_T = 10 \log_{10} \frac{S_{\min}}{I + N} \tag{9}$$

$$S_{\min} = 10^{\frac{S_T}{10}} (I + N) \tag{10}$$

Where $S_{\min}$ is the minimum received power at the edge of coverage area. L_i is the path loss to the ideal edge of cell boundary to prevent no overlap, which is the same as the pass loss in the Okumura-Hata model we used in Sect. 2. S_T is the SINR threshold for different modulation schemes. We assume the INR is 10 dB as a typical value to estimate the interference level, the minimum SINR threshold is used 2.3 dB. After this process, the transmit power for 'BS' is

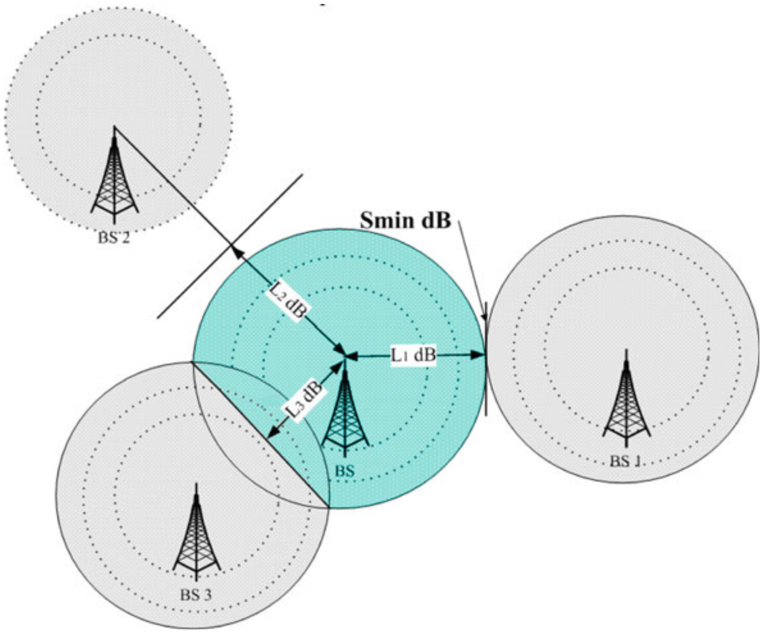
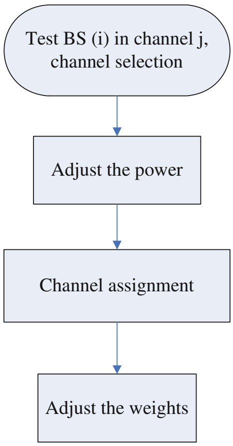


Fig. 6 Base station nominal service area with power adjustment-All base stations are in different channels

Fig. 7 Flowchart for channel selection and assignment with power adjustment in system



reset, and this is repeated at each BS, in order to obtain new transmit powers; all are limited by the $P_{\max}$ constraint.

Compared to the previous flowchart, we add one more step for power adjustment between the channel selection and assignment parts. The improvement here does not directly contribute to the weight itself, but greatly improves the user distributed occupancy detection and also is independent of the channel assignment scheme. The new flowchart is shown above (Fig. 7).

Figure 8 shows the cumulative probability of users with the ability to connect to a different number of base stations. The total number of users here again is 15 per base station and

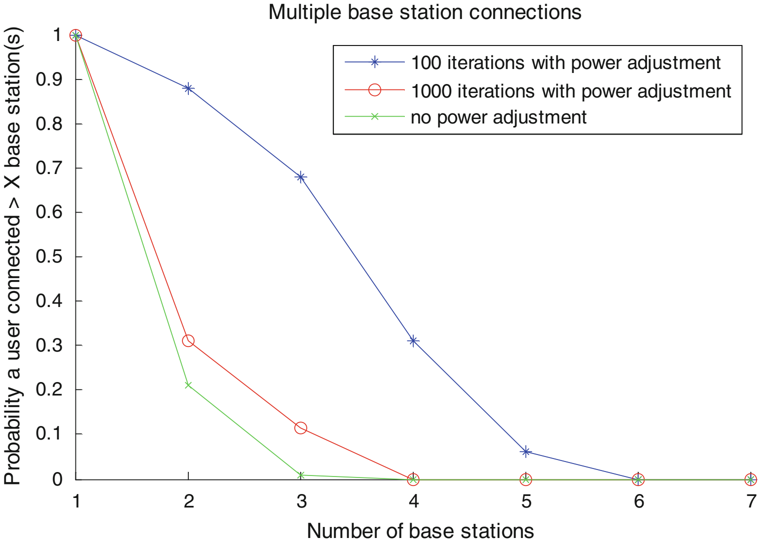


Fig. 8 Multiple base station connection for individual users

the minimum percentage threshold is 2.3 dB. Compared with the no power adjustment case, about 30% for 100 iterations and 20% for 1,000 iterations of the users are shared by 2 base stations in different channels, i.e. the power adjustment reduces the level of overlap, and the overlap is further reduced by the reinforcement learning. Due to the random location of users it is only possible to reduce the overlap, not completely eliminate it. As BS and BS3 shown in Fig. 6, if the service area boundaries intersect, there is still a small area of overlap, and users in this area can connect to two BSs. This is especially true for the users in the central area.

Figure 9 shows the percentage power reduction in transmit power for different reassignment thresholds. 2.3 dB is used for GMSK as a minimum threshold for reassignment, other thresholds are increased by each 2 dB margin as we explained before in weights part. When the threshold is 2.3 dB, the new transmit power is lower on average by -10.8 dB for 100 iterations, and -11 dB for 1,000 iterations compared with the original level. For both of 100 and 1,000 iterations, it is found that after power adjustment, the transmit power level has been reduced significantly but the users still satisfy the acceptance threshold. This significantly reduces the overall energy required in the communications system. Learning activation here is not that obvious for power adjustment from 100 to 1,000 iterations, especially when the iteration close to 1,000. Compared to the no power adjustment case, the learning process with power adjustment obtains enough information earlier, which means learning is more mature for this number of iterations, as service areas change little after this time.

From Fig. 10, it is found that compared to the no power adjustment scheme, the scheme with power adjustment starts with less reassignment. After a large number of iterations, the decreasing rate of reassignment is still higher than the scheme without power adjustment. When the power adjustment is based on the SINR boundary threshold of 2.3 dB, the transmit power level is much lower than the original transmit power level. It shows that the reinforcement learning also works well for this scheme, as the performance is greatly improved. Due to the new transmit power being much lower than the original case, base stations can be located in relatively random positions without explicitly considering the channel assignment. In the

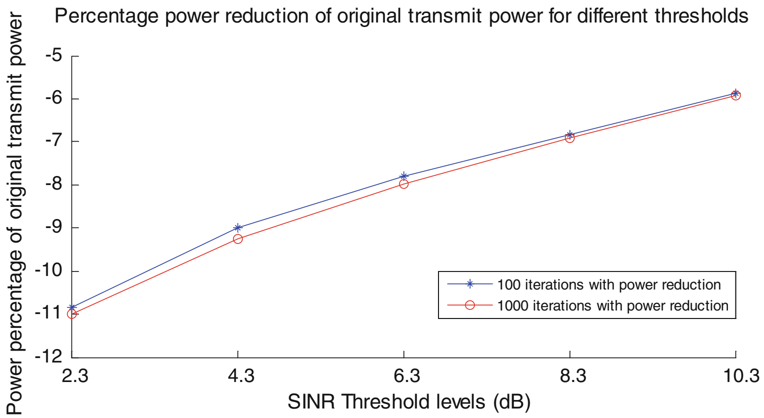


Fig. 9 Power reduction (dB) of original transmit power with different SINR thresholds after 100 iterations

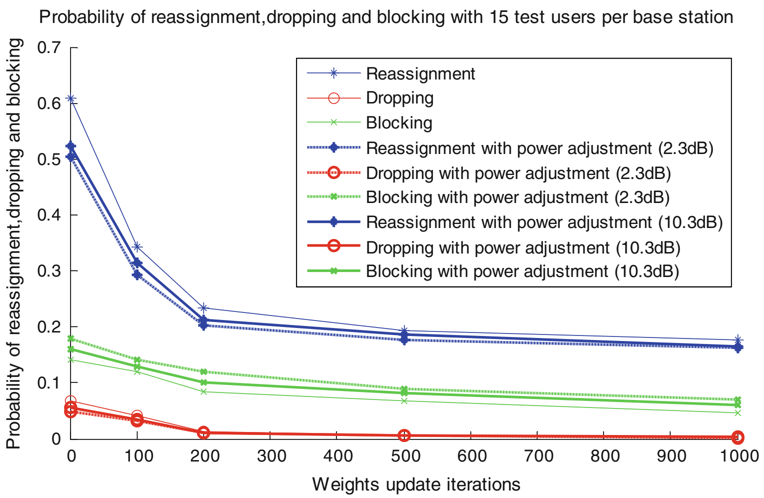


Fig. 10 Performance of power adjustment for reassignments, dropping and blocking with 15 users per base station

same random location situation, the scheme with power adjustment causes a lower interference than the scheme without power adjustment. This also explains why reassignment using a 10.3 dB threshold is worse than a 2.3 dB threshold. The three periods defined before for reinforcement learning process are still seen, only that the decreasing rate of reassignment here is smaller, because the initial channel assignment for the scheme with power adjustment is much better than the scheme without power adjustment. The improvement space is not as much as the scheme without power adjustment.

The blocking performance with power adjustment is worse than with the no adjustment cases. In the case of reduced power level situations, more base stations are being packed onto the same channel for initial assignment, which will cause the blocking to increase. The decrease in dropping will also cause a further increase in the blocking as space is not freed up on the assigned channels. Dropping improves significantly because the reduced power means that overall there is less interference from other base stations on the same channel.

6 Conclusion

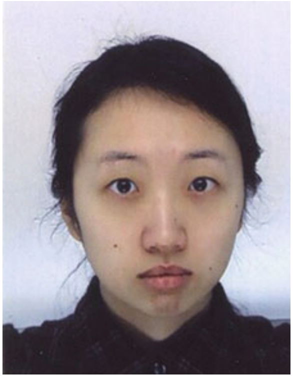
This paper has presented a random picking distributed channel assignment scheme applied to a cognitive radio system exploiting reinforcement learning with a user population receiving multicast downlink transmissions, with performance improved by power adjustment. It is found that distributed channel assignment schemes with reinforcement learning can efficiently improve the performance of channel assignment by limiting the reassignment, blocking and dropping rates. Moreover, adding power adjustment into the system helps base stations reduce their overlapping coverage areas and further reduces the interference from other BSs. The results show how the number of reassignments in the various schemes is significantly reduced after a large number of weight update iterations are used. By using the multicast architecture it is possible to exploit channels that utilize occupancy detection from multiple users, which helps solve the hidden node problem, resulting in a reduced number of reassignments and improving the dropping probability. However, this is at the expense of higher blocking because it is more difficult to find suitable free channels because they are more likely to be occupied by the multiple users. Different minimum quality of service threshold percentages can be used to control and improve performance, in place of the more traditional SINR threshold levels. It is found that significantly reducing the levels of overlap between adjacent base stations improves the performance of reassignment, dropping and blocking, while also reducing interference and saving transmitter power.

References

1. Mitola, J., III, & Maguire, G. Q., Jr. (1999). Cognitive radio: Making software radios more personal. *IEEE Personal Communications, 6*(4), 13–18 (see also IEEE Wireless Communications).
2. Alsarhan, & Agarwal, A. (2009, August). Cluster based spectrum management using cognitive radios in wireless mesh network. In *COGCOM 2009*. San Francisco, USA.
3. Fette, B. A. (2006). *Cognitive Radio Technology*. Jordan Hill, Oxford: Elsevier Science and Technology Books.
4. Katzela, I., & Naghshineh, M. (1996). Channel assignment schemes for cellular mobile telecommunication systems: A comprehensive survey. *IEEE Personal Communications, 3*, 10–31.
5. Berlemann, L., Mangold, S., Hiertz, G., & Walke, B. (2006, April). Policy defined spectrum sharing and medium access for cognitive radios. *Journal of communications, 1*(1), 1–12.
6. Cheng, M. M.-L., & Chuang, J. C.-I. (1996). Performance evaluation of distributed measurement-based dynamic channel assignment in local wireless communications. *IEEE Journal on Selected Areas in Communications, 14*, 698–710.
7. Peha, J. M. (2005). Approaches to spectrum sharing. *IEEE Communications Magazine, 43*, 10–12.
8. Yang, M., & Grace, D. (2008). Cognitive radio based spectrum assignment for heterogeneous multicast terrestrial communication systems with different transmission rate requirements, Cognitive Radio and Software Defined Radios: Technologies and Techniques, IET Seminar on 18th Sep. 2008.
9. Yang, M., & Grace, D. (2009, June). Cognitive radio with reinforcement learning applied to multicast terrestrial communication systems. In *CROWNCOM*. Hannover, German.
10. Zhao, J., Zheng, H., & Yang, G. H. (2005). Distributed coordination in dynamic spectrum allocation networks. In *IEEE DySPAN* (pp. 259–268).
11. Weiss, T. A., Hillenbrand, J., Krohn, A., & Jondral, F. K. (2003). Efficient signaling of spectral resources in spectrum pooling systems. In *10th Symposium on Communications Vehicular Technology (SCVT)*.
12. Saunders, S. R. (reprinted in August, 2004). Antennas and Propagation for wireless communication systems. University of Surrey, Guilford, UK, John Wiley & Son Ltd.
13. Harrold, T. J., Faris, P. C., & Beach, M. A. (2008). Distributed Spectrum Detection Algorithms for Cognitive, Cognitive Radio and Software Defined Radios: Technologies and Techniques, IET Seminar on 18th Sep. 2008.
14. Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research, 4*, 237–285.

15. Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge: The MIT Press.
16. Mitchell, T. M. (1997). *Machine learning*. New York: McGraw-Hill.
17. Yang, M., & Grace, D. (2009, August). Cognitive radio with reinforcement learning applied to multicast downlink transmission and distributed occupancy detection. In *COGCOM*. San Francisco, USA.
18. Jiang, T., Grace, D., & Liu, Y. (2008). Cognitive radio spectrum sharing schemes with reduced spectrum sensing requirements, IET Seminar on 18th Sep. 2008.

Author Biographies



Mengfei Yang received her B.E. (Hons) degree in Electronics Engineering from the Department of Electronics, University of Central Lancashire, UK, in 2005. She is currently a Ph.D. student in the Communications Research Group, University of York, under the supervision of Dr. David Grace. Her research interests include dynamic channel assignment, interaction and coexistence of heterogeneous multicast terrestrial communication systems, applied with artificial intelligence technique, especially on reinforcement based learning. She jointed the Worldwide Universities Network Cognitive communications Consortium and presented the work in the second meeting in Hannover, Jun, 2009.

David Grace is Head of Communications Research Group and Senior Research Fellow within the Department of Electronics at the University of York. Current research interests include cognitive radio, particularly applying distributed artificial intelligence to spectrum assignment, and cognitive networking, interference management and coexistence. Currently, he is a co-investigator of the FP7 BuNGee project dealing with broadband next generation access, and recently he was the principal investigator of a UK MOD project on 'Cognitive Routing for Tactical Ad Hoc Networks'. From 2003–2007 he was the technical lead for the 14-partner FP6 CAPANINA project. He is currently chair of the Worldwide Universities Network Cognitive Communications Consortium, which has 50+ member organisations worldwide. He is an author of 140+ papers, and a member of COST IC0902 and the IEEE Committee on Cognitive Networks. From 2005–2009 he was COST 297 WG1 chair which dealt with radio communications for high altitude platforms. In 2000, he jointly founded SkyLARC Technologies Ltd, and was one of its directors.