AN ABSTRACT OF THE THESIS OF

Tamara AlShammari for the degree of Master of Science in Electrical and Computer Engineering presented on July 17, 2014.

Title: Malicious-Proof and Fair Credit-based Resource Allocation Techniques for DSA Systems

Abstract approved: ____

Bechir Hamdaoui

In this paper, we propose a credit-based resource allocation technique for dynamic spectrum access (DSA) systems that is robust against malicious and selfish behaviors and ensures good overall system fairness performance while also allowing spectrum users to achieve high amounts of service. We also propose a new objective function that, when combined with the proposed credit-based technique, leads to further improvements of the system fairness performance. Our proposed techniques overcome user misbehavior by masking the impact of the users' pursued private objectives on the overall system performance. They also improve fairness among users by allocating service to users adaptively by accounting for how much service each user has received in the past. Our simulation results show that our proposed techniques maintain high system performance by allowing users to achieve high amounts of service and by ensuring fair allocation of spectrum resources among users even in the presence of misbehaved users. Using simulations, we also show that these high performances are also achievable under various different network scenarios.

Malicious-Proof and Fair Credit-based Resource Allocation Techniques for DSA Systems

by

Tamara AlShammari

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented July 17, 2014
Commencement June 2015

Master of Science thesis of Tamara AlShammari presented on July 2014.

APPROVED:

Major Professor, representing Electrical and Computer Engineering

---

Director of the School of Electrical Engineering and Computer Science

---

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Tamara AlShammari, Author

ACKNOWLEDGEMENTS

"And finally let me tell you, if you passed all this successfully never forget who helped, supported, or motivated you because while climbing that ladder to success, those were the ones who held it steady"

At first, I would like to thank God "Allah" for all the beautiful things that he gave to me, for one more dream came true, and for being there all the time wherever I go. Secondly, I would like to express my gratitude to my supervisor Bechir Hamdaoui for the useful comments, remarks and engagement through the learning process of this master thesis. Many thanks go to my beautiful gorgeous mother for raising the passion for learning and knowledge in me, to my uncle who supported me all the way, and to my little sister for all the time she spent listening to my research challenges. Furthermore, I would like to special thank Yousef Qassim for his inspirations, keeping me harmonious, his believe in my abilities, and helping me putting pieces together. Finally, I would like to thank my friends Asma'a and Nadia for being there. I will be grateful forever for your love and care.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1: Introduction

Dynamic spectrum access (DSA) is a new paradigm that allows spectrum users to seek and use spectrum bands opportunistically, in that users can switch across and access different bands dynamically and can share the same spectrum at the same time. In DSA systems, there are two types of users: licensed or primary users (PUs) and unlicensed or secondary users (SUs). DSA systems allow secondary users to sense the licensed spectra, and to occupy and use any spectrum band when it is not used by its primary users. However, these secondary users must be transparent to the primary users in that they have to leave the spectrum band as soon as they sense the presence of any primary users.

DSA has great potentials for overcoming spectrum shortage problems [4, 5]. As a result, several researches have been conducted to address various aspects of DSA systems, such as designing efficient sensing techniques [7, 8], balancing between system utilization and its reliability [2], and overcoming user malicious behaviors during sensing [6, 18]. One of the key challenges arising from the complex and diverse nature of nowadays emerging wireless systems is the design and development of efficient spectrum access techniques that can be implemented in a distributed manner and scale well with the number of users. Learning-based techniques have been considered as a potential solution candidate for such a challenge due to their inherent distributed nature [9–12, 16, 17]. These techniques essentially propose distributed DSA access methods that can perform without the need for any control/central unit, thereby enabling secondary users to distribute themselves among available bands/channels without any guidance or direction from any third party or entity. Learning-based techniques allow secondary users to do so by using their knowledge to be acquired through interaction with the environment (their present and past experiences) to decide what to do best in the future.

Throughout literature, various researches have shown that poorly designed objective functions can lead to poor system performances [15]. As a result of this, some research efforts have been put to develop efficient objective functions for DSA systems [3, 10, 12]. These previously proposed objective functions such as those proposed for elastic [12] and inelastic [3, 10] traffic models are shown to have good performances in terms of optimality, scalability, and distributivity. However, they do have some shortcomings which we aim to address in this work. At first, they are not robust against malicious behavior in the sense that if some users choose (intentionally or unintentionally) to pursue selfish and greedy objectives, the overall system performance degrades, and such a degradation can be very severe depending on the level of maliciousness/selfishness. Secondly, they are unfair in that users that employ these proposed techniques may not receive equal amounts of service.

In this paper, we propose a credit-based spectrum resource allocation technique for DSA systems that, unlike the previous techniques, is robust against malicious and selfish behaviors and improves fairness among users. In addition, we propose a new objective function that, when combined with the credit-based technique, fairness among users is improved even further while still maintaining high system performance. The robustness of our proposed techniques against users' misbehavior (selfishness and maliciousness) lies in its ability to mask the impact that the users' pursued private objectives have on the overall system performance. Fairness improvements, on the other hand, are achieved by allocating service to users adaptively while accounting for the amount of service each user has received in the past. Using simulations, we show that the proposed techniques achieve high system performance by allowing users to receive high service levels and by ensuring fair allocation of spectrum resources among users even when the system contains misbehaved users. We also show that these high performances are also achievable under various different network scenarios by considering dynamic primary and secondary user activities.

The rest of this thesis goes as follows. In Chapter 2, we describe our system model. In Chapter 3, we state our motivation and objective by illustrating the shortcomings of existing techniques. Chapters 4 and 5 present our proposed credit-based technique and our proposed fairness objective function, respectively. In Chapter 6, we evaluate and show the performance of our proposed techniques under various network settings. Chapter 7 highlights and discusses some implementation and practical aspects of the proposed techniques. Finally, we conclude the thesis in Chapter 8.

Chapter 2: System Model

In this work, we assume a fully-connected topology, i.e. all users interfere with each other. Also, we consider a wireless system with m non-overlapping spectrum bands (or channels). We assume that each band j offers an amount of service denoted by V_j; the service that the band offers could, for example, be throughput, reliability, data rate, etc. We also assume that there is an access point (or a monitoring agent) deployed in the system whose responsibility is to keep track of what and when users join the spectrum bands.

We consider the elastic traffic model in which a user's received reward corresponds to the amount of service it receives from using the spectrum when this received reward exceeds a certain threshold, Q. On the other hand, when the received amount of service is less than the threshold, the user's reward drops very quickly and becomes unacceptable. We assume that users do not leave their spectrum band (and try to find another band) unless their received level of service goes below their required level. In addition, we adopt the adaptive service model, proposed in [11], where the users' required level of service changes depending on what they have received so far. Mathematically, the reward, $r_i(t)$, of user i at time t can be written as [11]:

$$r_i(t) = \begin{cases} S_i(t) & \text{if} \quad S_i(t) \geq Q(t) \\ Q(t)e^{-\beta\frac{Q(t)-S_i(t)}{S_i(t)}} & \text{otherwise} \end{cases} \tag{2.1}$$

where $S_i(t)$ is user i's received level of service at time t, $Q(t)$ is the required level of service at time t, and β is the decaying factor. At last, we assume a time-slotted resource access and sharing scheme, where users are assumed to arrive at the beginning and leave at the end of time steps.

Chapter 3: Fairness and Misbehavior

In learning-based DSA techniques, after a user determines its objective, it tries to maximize it using a learning algorithm. Two intuitive objective functions can be considered. The first one is the intrinsic reward function, r_i, given in Eq. (2.1), where here a user aims to maximize its own received reward; this function reflects the users' expected selfish behaviors when going after maximizing their own received rewards. The other one is the global/total reward function, G, which aims to maximize the total rewards received by all users. At time t, $G(t)$ can formally be written as

$$G(t) = \sum_{i=1}^{n(t)} r_i(t) \tag{3.1}$$

where $n(t)$ is the total number of users accessing the system at time t. The main drawback of using these two functions is that they lead to poor system performance. This is because in the intrinsic function case, users' objectives are not aligned with one another, and in the global function case, users' objectives are not sensitive enough to their own actions to lead to high rewards. Detailed and good explanations of such performance behaviors can be found in [12].

To address this performance issue, the difference objective function, D_i, has instead been used in DSA networks for supporting both elastic [12] and inelastic [3] traffic models, and is shown to achieve near-optimal performances by outperforming r_i and G substantially. The reason behind the high performance that D_i achieves lies in the fact that when the number of users in the system exceeds the channels' capacities, D_i leads to a near-optimal distribution of the users among the different available spectrum bands. As shown in [12], the optimal distribution occurs when $(m-1)$ channels/bands each has exactly a number of users equaling the channel's capacity, and the m^{th} band has all the other remaining users[1].

[1]This is when all channels are assumed to offer the same service; i.e., $V_j = V$ for all j. Refer to [11] when V_js are not the same.

For illustration purposes, we consider in this section a DSA system with 10 bands and 500 users. Also, for simplicity and without loss of generality, we assume that all bands offer the same amount of service; i.e., $V_j = V = 20$ for all j. In our figures, we normalize the global received reward with respect to an approximation of the maximal global achievable reward, given in [12]. We plot in Fig. 3.1 the normalized achievable global/total reward under D_i, r_i and G. As stated above, observe that r_i and G result in very poor performance, whereas D_i results in high performance. In addition to achieving high rewards, D_i is shown to scale well with the number of users, and can be implemented in a fully distributed manner in fully connected networks, as reported in [12].

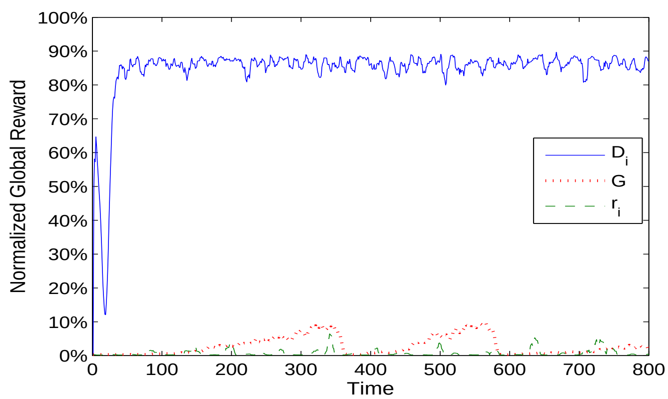

Figure 3.1: Normalized global reward under various time steps.

Despite of its performance advantages, D_i has some shortcomings. First, it is unfair. This is because, under D_i, some users may end up staying in the most crowded channel more than others, thereby receiving smaller amounts of service. To illustrate, we show in Fig. 3.2 the standard deviation of users' received rewards under the D_i function for different number of users. As it can be seen from the figure, the standard deviations can be relatively high, implying that users may receive unequal amounts of service when D_i is used.

Second, the D_i function is not robust against misbehaved users[2]. The issue is that even though D_i can increase the achievable performances, it can only do so when all users pursue it as their objectives. In other words, when some users choose (intentionally

[2]Specifically, we look at misbehaving in terms of being greedy and looking for receiving more service even if the user does not need it.

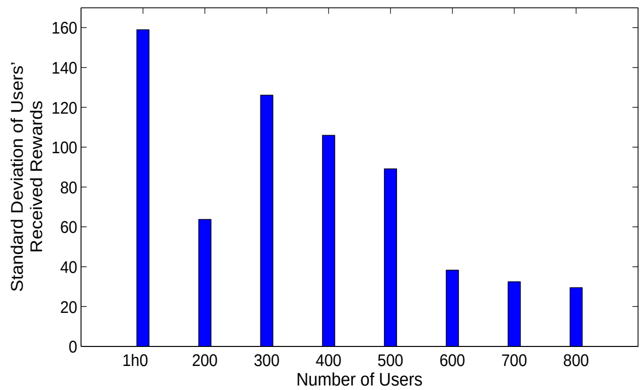

Figure 3.2: Standard deviation of users' received rewards for different number of users under when D_i is used.

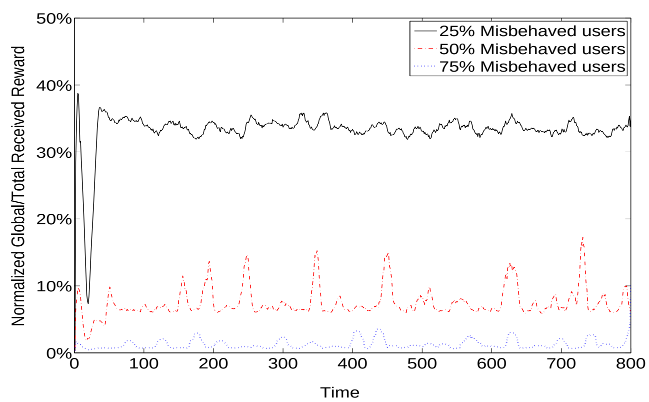

Figure 3.3: Normalized global reward achievable under D_i for various percentages of misbehaved users.

or unintentionally) not to pursue this function, the function can no longer lead to good performances. To illustrate this, we show in Fig. 3.3 the performance of D_i in the presence of misbehaved users. In the figure, "x% Misbehaved users" refers to the case when x% of the users choose to pursue their greedy objective function r_i, while the other $(100 - x)$% users implement the D_i function. The figure shows that as the percentage (i.e., number) of misbehaved users increases, the overall system received reward decreases, and hence, so does the per-user average received reward. This overall performance degradation gets

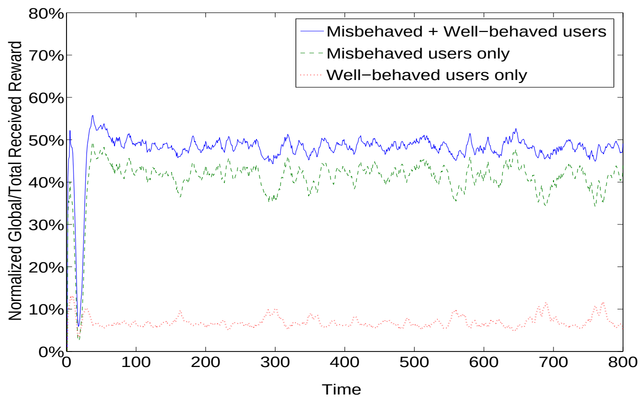

Figure 3.4: Normalized global/total received reward when 20% of the users misbehave.

even worse when the percentage of misbehaved users becomes higher and higher.

What's even worse, not only do these misbehaved users lead to poor overall system performance, but also receive most of the available service, thereby leaving those well behaved users with no to little service. This is illustrated in Fig. **??**. Note that the misbehaved users (which represent only 20% of all users) receive about 40% of the optimal/total amount of achievable service, whereas the well behaved ones (which represent 80% of all users) receive all together only about 10% of the total possible amount of service. Also, it is worth mentioning that because of the presence of these misbehaved users, the overall global/total system reward goes down from about 85% when all users behave well (as shown in Fig. 3.1) to about 50% only.

It is therefore important to devise efficient strategies and techniques that ensure fair allocation of resources among users while also maximizing the achievable system performance even in the presence of misbehaved users. Our proposed strategy for doing so consists of developing new resource allocation techniques that are immune from the users' objective function choices, in that even when users choose to deploy and pursue greedy goals and objectives, their collective selfish behavior does not lead to poor and unfair system performance. In addition, with our developed techniques, the amount of service a user receives depends not only on which channel the user selects, but also on how much service it has received so far, thereby ensuring fairness among users. To sum up, the proposed techniques possess two characteristics:

- **Robustness against user selfishness.** It achieves high performance regardless of what objective functions users choose. In other words, it reduces the effect that objective functions have on the achievable performance.

- **Fairness among users.** It improves fairness among users by allocating service to users adaptively based on how much service each user has received in the past.

In essence, this work makes two contributions: *i*) it proposes a credit-based resource allocation approach that allocates spectrum resources on an adaptive manner by accounting for what users receive in the past; and *ii*) it proposes a new objective function that aims to maximize the users' contribution to the overall system fairness. By combining the proposed objective function with the proposed credit-based allocation approach, we then propose an integrated resource management framework that allocates spectrum resources among users fairly while ensuring robust against malicious and selfish behavior.

Chapter 4: Credit-based Resource Allocation

We propose a credit-based resource allocation technique that overcomes selfish behaviors and ensures fairness among users. In this technique, each user is assigned a credit value with an initial value of one. This credit value determines the proportion of service that each user should receive; the greater the credit value, the higher the service to be received.

Under the proposed credit-based resource allocation technique, users' credit values get updated depending on the amount of service they receive when compared to the system fair-share. We define the system fair-share as the amount of service that each user should receive in order to ensure fair allocation of available service among all users. The maximum amount of system service is achieved when the users are distributed among the channels as follows: each channel j contains exactly $b_j = V_j/Q$ users except the channel with the minimum capacity which should contain the remaining number of users [11]. When $V_j = V$ for all channel j, this optimal distribution leads to a maximal value of global received reward that can be approximated to [11]: $\hat{G}_{max} = (m-1)V$. When the spectrum resources are allocated fairly among all users, each user i should then receive at time t a system fair-share, $R_i(t)$, that is equal to:

$$R_i(t) = \sum_{t'=t_i}^{t} \frac{(m-1)V}{n(t')} \tag{4.1}$$

where t_i is the time step at which user i joins the DSA system and $n(t)$ is again the total number of users present in the system at time t.

Now, we want to generalize our service model to allow service differentiation among users. That is, different users may have different service priorities, and those that have higher priority should receive higher amounts of service. We assume that there are c classes, $\{1, 2, \ldots, c\}$, and we let p_i denote the priority of class i. When considering service differentiation, the system fair-share is to be defined so that users with higher priorities should be able to achieve amounts of service that are greater than those achievable by users with lower priorities. Formally, the system fair-share, $R_i(t)$, of user i with priority

p_i can be written as

$$R_i(t) = (m-1)V \sum_{t'=t_i}^{t} \frac{p_i}{\sum_{k \in \mathcal{A}(t')} p_k} \tag{4.2}$$

where $\mathcal{A}(t)$ denotes the set of all users using the system at time t. Note that Eq. (4.1) is a special case of Eq. (4.2) in which all users have the same priority.

At the end of each time step, if the user receives less than the system fair-share, its credit value increases by one if it does not exceed a certain threshold, otherwise it is set to that threshold value. On the other hand, if the user receives more than the system fair-share, its credit value gets decreased by one when it is greater than a certain threshold, otherwise it is set to that threshold value. Consequently, users' credit values could be positive, zero, or negative. Whenever a user credit value reaches zero or a negative value, the user is no longer able to receive service until its credit value becomes positive again.

Indeed, when the user receives less than the system fair-share, our approach requires that its credit value does not exceed a certain threshold so as to prevent it from becoming very large. Otherwise, once the user receives its fair-share of the spectrum, it will take a relatively long time for its received service to be reduced so as to not exceed its fair-share. Likewise, in the case of receiving more than the system fair-share, a user credit value must not go below a certain threshold because otherwise its credit value would keep decreasing, and can reach a small value. When this happens, if this user, after sometime, wants to ramp up its share again, it will take it a long time before it can actually start receiving service.

Mathematically, a user i's credit value at time t, $Cr_i(t)$, is calculated as:

$$Cr_i(t) = \begin{cases} \max\{Cr_i(t-1)-1, Cr_i^{th}(t)\} & \text{if } R_i(t) < \sum_{t'=t_i}^{t} S_i(t') \\ \min\{Cr_i(t-1)+1, Cr_i^{th}(t)\} & \text{otherwise} \end{cases} \tag{4.3}$$

where the credit threshold bound is defined as:

$$Cr_i^{th}(t) = (R_i(t) - \sum_{t'=t_i}^{t} S_i(t'))/Q(t)$$

The numerator in the above equation represents either the missing service in case the user did not receive its whole fair-share or the extra service in case the user received more than its fair-share. In this equation, the amount of missing/extra service is represented as a multiple/fraction of user's required service at time t. This means that the user's credit value threshold represents how many Qs the user has to receive/miss in order to receive its fair-share.

For more clarification, let us consider the following example where we assume that at time t, user A is missing 4 units of service. If we further assume that user A's $Q(t)$ is equal to 2, then its credit threshold is 2. This means that user A needs $2Q$'s to compensate the missing service. Let us also assume, for illustration, that at the same time t, its credit value reaches that threshold, i.e. $Cr(t) = 2$. Since its credit value is positive, this user is able to receive service. At time $t+1$, this user receives an amount of service that is equal to Q plus the fair-share amount for this time step. Thus, its threshold at time $t+1$ gets updated, and becomes equal to 1. This implies that user A needs now just one Q to cover the missing service. Since its previous credit value is greater than its current threshold, its current credit becomes equal to the threshold, i.e. $Cr(t+1) = 1$. In the same manner, if this user at time $t+2$ receives $2Q$ plus the fair-share for one time step, its threshold value gets updated and becomes equal to -1. This means that this user is no longer missing any service, and has actually received one extra Q. The same credit updating process continues with reversing the condition for the credit value. That is, the user credit value must not go below the new threshold.

Using user i's credit at time $t-1$, the amount of service that user i receives from accessing band j at time t is:

$$S_i(t) = \begin{cases} \frac{Cr_i(t-1)}{\sum_{k \in B_j(t): Cr_k(t-1)>0} Cr_k(t-1)} V_j & \text{if } Cr_i(t-1) > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $B_j(t)$ is the set of all users belonging to band j at time t.

It is worth mentioning that we here assume that our system is associated with an access point whose task is to monitor and keep track of users (their ID, their activities, their check-in and check-out times). One cannot rely on users to report their credit values as they might cheat in order to receive more service. Thus, we consider that both users and the access point are calculating and updating users' credit values. This way when a user lies about its credit value, the access point will know about the mismatch between the value it calculated and the value the user reported. As a result of this user behavior, the access point will block this user from accessing and using the DSA system.

Chapter 5: Fairness-Aware Objective Function

In order to increase the system fairness, we propose a new fairness-aware objective function that, when combined with our proposed credit-based technique, ensures *near-fair* system performance. In addition to improving fairness, the proposed objective function, coupled with the credit-based allocation, achieves not only performances that are as high as those achievable under D_i, but also ensures robustness against misbehaved users.

Our proposed objective function, referred to as F_i, essentially aims to maximize user i's contribution to the overall system fairness. Formally, it can be written as

$$F_i(t) = \begin{cases} R_i(t) - \sum_{t'=t_i}^{t} S_i(t') & \text{if } \sum_{t'=t_i}^{t} S_i(t') \geq R_i(t) \\ \sum_{t'=t_i}^{t} S_i(t') & \text{otherwise} \end{cases} \tag{5.1}$$

The intuition behind our objective function is that when a user receives more than its system fair-share, then it will try to minimize the gap between its fair-share and its received service. That is, users will try to positively contribute in a way that ensures an overall system fairness. On the other hand, if a user receives less than its fair-share, then it will try to maximize the amount of service it gets to also contribute to the system fairness. Here, we assume that users are willing to cooperate and to have a good contribution on the system fairness. However, if this is not the case, and some users use other objective functions, then that fortunately will not hurt the system fairness that much, thanks to our credit-based resource allocation approach, which alone is capable of maintaining a relatively good system fairness. These performance enhancements will be shown via simulation as reported in the next section.

It is worthy to mention that the total number of interfering users that are using the system at this time step is needed to calculate user's fair function. The methodology of knowing this information is left for future work.

Chapter 6: Performance Evaluation

In this section, we evaluate the performance of the proposed resource allocation technique in terms of the achievable global reward as well as the coefficient of variations (CoV)[1] of users' normalized[2] received rewards. We apply this technique for each of the four objective functions: r_i, G, D_i, and F_i. For simulation purposes, we use the ϵ-greedy Q-learning algorithm. In this algorithm, at the end of each time step, the user selects the channel whose Q-value is the highest with probability 1-ϵ, and selects a random channel with probability ϵ. Whenever a user is tuned to a channel, it measures the service it receives, and then uses it to update the Q-value entry corresponding to the channel being currently used. Readers are advised to refer to [12, 13] for more details.

In our simulations, we consider three different scenarios/cases: static scenario, dynamic scenario without PUs, and dynamic scenario with PUs. In the static scenario, all SUs arrive to and leave the system at the same time, whereas in the dynamic scenarios, SUs may arrive and/or leave at different times. In the following subsections, we study and present our evaluation results under each of these three scenarios. Throughout the simulation section, we set $c = 10$, and assign users' priorities randomly; that is, each user i is assigned a priority level p_i, selected randomly from $1, 2, \ldots, 10$.

6.1 Static DSA system

In this section, we consider a static DSA system, i.e. SUs arrive and leave at the same time. We assume that the system does not have any PUs. We set: $n(t)$=n=500 for all t, V_j=V=20 for all j unless stated otherwise.

[1](CoV) is defined as the ratio of the standard deviation to the mean.

[2]All normalized received rewards presented in this section are with respect to the maximal possible achievable reward approximated in [12].

6.1.1 Optimality

We show in Fig. 6.1 the normalized global reward of r_i and G with and without the proposed credit-based resource allocation technique. Fig. 6.1 shows that the proposed credit-based technique allows users to achieve high rewards/service even when they choose to chase their greedy objective, r_i, or the global objective, G. Fig. 6.2, on the other hand, shows the normalized global reward also under the proposed credit-based resource allocation technique but for the F_i and D_i functions. While the function D_i already performs well in terms of the amount of achievable rewards, adding the credit-based feature does improve its achievable performance even more. The figure also shows that, when coupled with the credit-based resource allocation approach, our proposed fairness-aware function F_i outperforms the function D_i in terms of the amount of achievable service. As we stated it earlier, the real benefit of our proposed techniques lie, however, in their robustness against maliciousness and in their ability to ensure fairness, as will be illustrated in the next section and the other following ones.

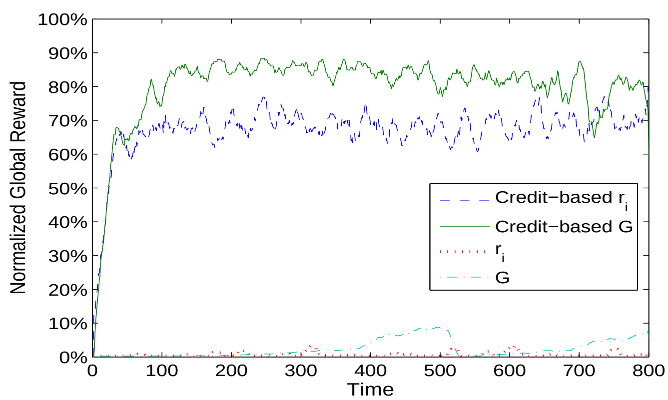

Figure 6.1: Normalized global reward of r_i and G without and with credit-based resource allocation (static case).

6.1.2 Robustness against objective function choice

We now show that the proposed credit-based technique reduces the impact of the objective function choice on the system performance. Fig. 6.3 shows the system performance

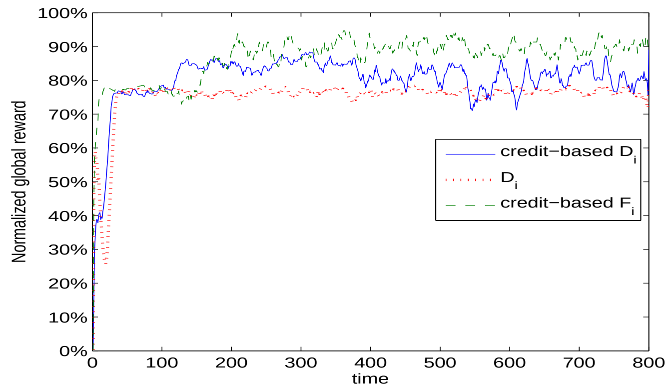

Figure 6.2: Normalized global reward of D_i and F_i without and with credit-based resource allocation (static case).

achievable when users choose to pursue different objectives under the credit-based resource allocation technique. Consider the case when 25% of the users misbehave (i.e., use the r_i function instead of the D_i function) while the other users use D_i (the plot in the figure corresponding to "25% r_i, 75% D_i"). Observe that when the proposed credit-based technique is not used, the normalized overall system performance is about 40% only, whereas when the technique is used, the system performance reaches about 80%. In this specific scenario, the adoption of our technique doubles the overall achievable performance. As you can see from the figure, this performance improvement can be even greater when the percentage of misbehaved users is higher. This is illustrated in the figure via the case corresponding to when 75% of the users misbehave while the rest of the users use the function D_i (i.e., plot: "75% r_i, 25% D_i"). Note how low the achievable system performance is in this case.

In conclusion, our results discussed in the above paragraph show that the incorporation of the credit-based approach does improve the achievable performance significantly in terms of robustness against misbehaved users even under D_i.

Let us now reflect on the performance behaviors when considering the proposed fairness-aware function. In the same figure, Fig. 6.3, we also show the performance when 25% of users misbehave by using the r_i function while the rest of the users use the proposed F_i function. Observe that even when some users pursue their greedy objectives, the system can still achieve high performance. This achieved performance is also

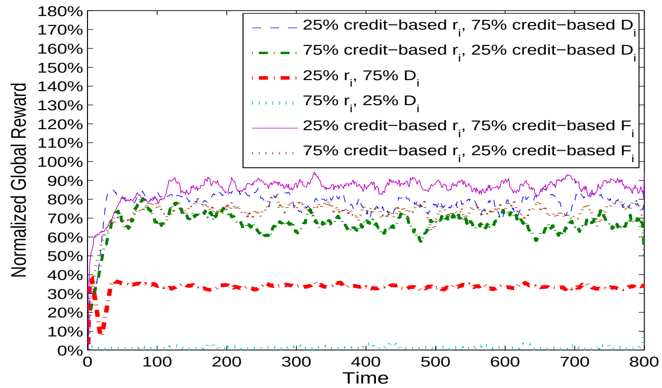

Figure 6.3: Normalized global reward without and with credit-based resource allocation in the presence of misbehaved users (static case).

higher than what D_i achieves even when combined with our credit-based approach. We also show the performance even when 75% of users pursue greedy goals (i.e., r_i) while the rest use the proposed F_i function (i.e., plot: "75% *credit-based* r_i, 25% *credit-based* F_i"). Observe that the achievable performance is still high, and regardless of how large the misbehaving percentage is, the credit-based technique can still prove its robustness against users' objective function choices. Also, this performance, achieved under F_i, is higher than that achievable under D_i even when combined with the credit-based technique. Therefore, we conclude that the proposed credit-based, fair objective function outperforms the D_i function in terms of its robustness vis-a-vis of malicious and selfish behavior.

6.1.3 Fairness

Fig. 6.4 shows the CoV of users' normalized received rewards under D_i without using the credit-based technique, as well as under r_i, G, D_i, and F_i when using the credit-based technique for different numbers of users ($n(t)$=n=400, 700, and 1000). Since the priorities in the case of the D_i function when not using the credit-based allocation have no effect, for the sake of comparison, we assume in this paragraph that all users have the same priority. There are two observations we want to make out of Fig 6.4. First, observe that the proposed credit-based F_i function reduces CoV drastically and outperforms all other cases significantly, thereby improving fairness among users substantially. Second, observe also that our proposed credit-based technique does reduce CoVs even for other function choices. Note, for e.g., the gap difference in CoVs for the function D_i without and with the credit-based technique. (Similar trends have also been observed for the r_i and G functions, but they are not plotted in the figure, so as to focus on the D_i and F_i functions).

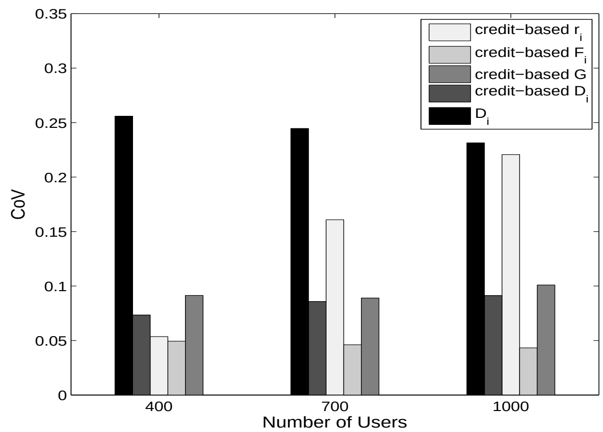

Figure 6.4: Coefficient of variation of users' received rewards under different objective functions (static case).

6.2 Dynamic DSA system without PUs

We now assume that SUs arrive and leave at different times. Also, we assume that SUs arrive according to a Poisson process with arrival rate λ, and they stay in the system for an exponentially distributed duration of mean μ. On average, at each time step, the number of users (κ) is equal then to ($\lambda\mu$). In this section, we investigate dynamic DSA systems without considering the existence of PUs.

6.2.1 Optimality

Figs. 6.5 and 6.6 show the normalized global reward of r_i, G, F_i, and D_i with and without the proposed credit-based resource allocation technique when (κ=1000, λ/μ=0.1) and

(κ=700, λ/μ=0.14), respectively. When λ/μ=0.1, The figures show that, even in the dynamic scenarios, the proposed credit-based technique is able to achieve high amounts of service regardless of the objective function that users choose to maximize; i.e., even when users choose the intrinsic, r_i, or the global objectives, G, as their objectives. For example, the global achievable reward under the function D_i goes from about 50% when the credit-based approach is not used to about 85% when the approach is used. The second point we observe is that the credit-based approach, when used with the F_i function, achieves higher amounts of service than what other functions achieve.

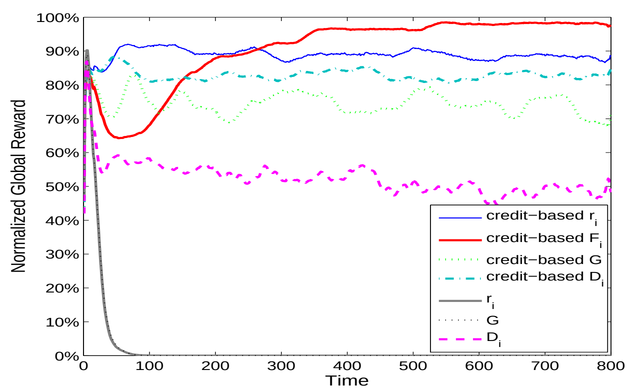

Figure 6.5: Normalized global reward of all objective functions with and without credit-based resource allocation: λ/μ=0.1, κ=1000 (dynamic without PUs case).

6.2.2 Robustness against objective function choice

We now show that the proposed credit-based technique reduces the impact of objective function choice on the system performance even in the dynamic scenarios. Figs. 6.7 and 6.8 show the system performance achievable when users choose to pursue different objective functions under the credit-based resource allocation technique for (κ=1000, λ/μ=0.1) and (κ=700, λ/μ=0.14), respectively. Consider the case when 50% of the users use either F_i or D_i, and the other 50% use the r_i function under the credit-based allocation scheme. Note that regardless of what users choose to pursue as their objectives, the credit-based technique ensures that the system as a whole reaches and maintains high performance. Our proposed F_i function does, however, achieve better performance

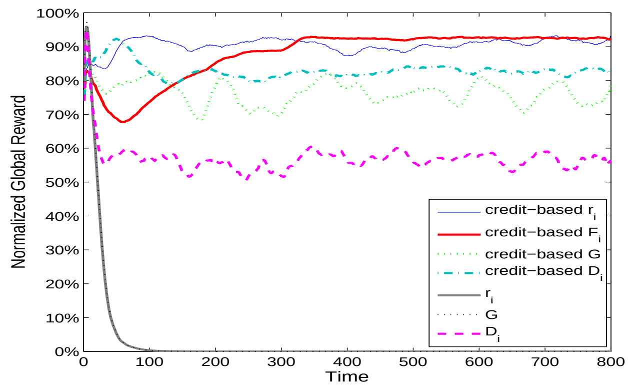

Figure 6.6: Normalized global reward of all objective functions with and without credit-based resource allocation: λ/μ=0.14, κ=700 (dynamic without PUs case).

than the other functions, including D_i. Although we show our results only for the case when 50% of the users choose greedy objectives, similar performance behaviors have been observed for other percentages as well. Finally, note that when the credit-based technique is not used, having misbehaved users can degrade the system performance substantially. This can be seen in the figures when considering the D_i function without the credit-based approach (plot: "50% r_i, 50% D_i").

In conclusion, our results show that the F_i function coupled with the credit-based technique is more robust against misbehavior and reaches higher system performances than the other functions.

6.2.3 Fairness

Fig. 6.9 shows the CoV of users' normalized received rewards under D_i when not using the credit-based technique, as well as under r_i, G, D_i, and F_i when using the credit-based technique when the number of users, on average, equals 400, 700, and 1000; i.e., κ = 400, 700, and 1000. Like in the static case, we also make two observations vis-a-vis of fairness when considering dynamic behavior of secondary users. First, observe that the proposed credit-based F_i function achieves better fairness levels than the other functions, especially for high average numbers of users. Second, note that the use of the proposed credit-based technique helps in reducing the CoV of the D_i function as well, thus making

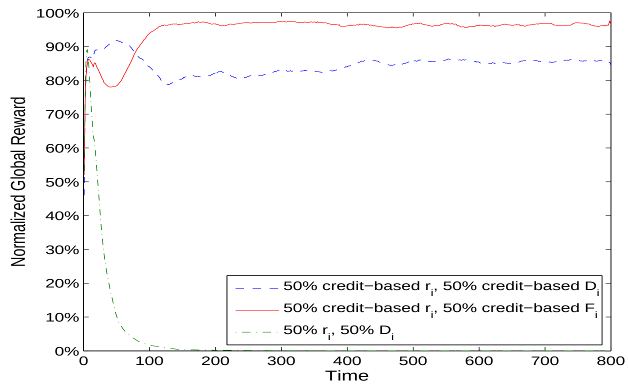

Figure 6.7: Normalized global reward with and without credit-based resource allocation: λ/μ=0.1, κ=1000 (dynamic without PUs case).

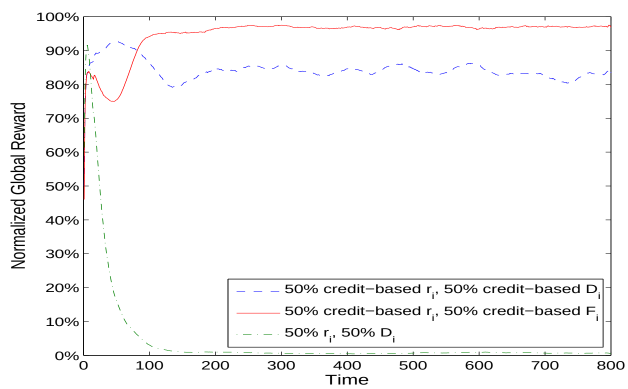

Figure 6.8: Normalized global reward with and without credit-based resource allocation: λ/μ=0.14, κ=700 (dynamic without PUs case).

it even fairer; similar trends have also been observed for the r_i and G functions, but are not plotted in the figure, so as to focus on the D_i and F_i functions.

6.3 Dynamic DSA system with PUs

We now consider a dynamic DSA system with the presence of PUs. We use a ON/OFF renewal process to mimic PUs' activities and presence in the system. For each spec-

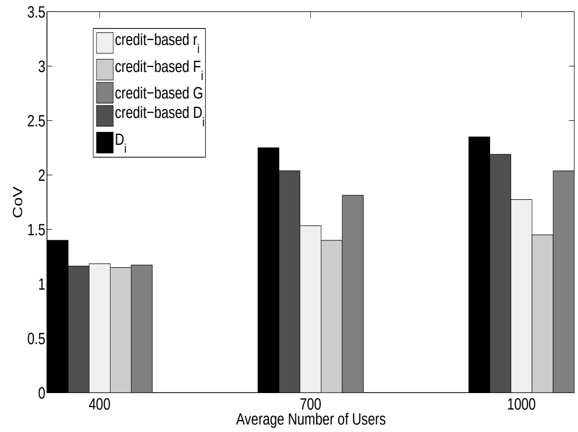

Figure 6.9: Coefficient of variation of users' normalized received rewards under different objective functions (dynamic case without PUs).

trum band j, we assume that ON and OFF durations are exponentially distributed with means μ_{ON} and μ_{OFF}, respectively. We define PUs load (η) to be $\mu_{ON}/(\mu_{OFF}+\mu_{ON})$. Moreover, we assume that PUs arrive to and leave the system at different times.

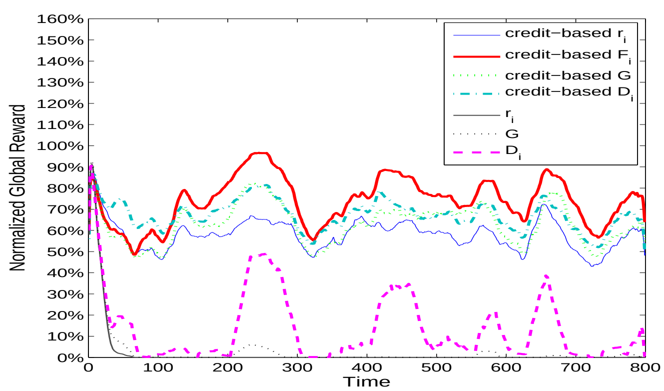

Figure 6.10: Normalized global reward of all objective functions with and without credit-based resource allocation: λ/μ=0.1, κ=1000, η=20% (dynamic with PUs case).

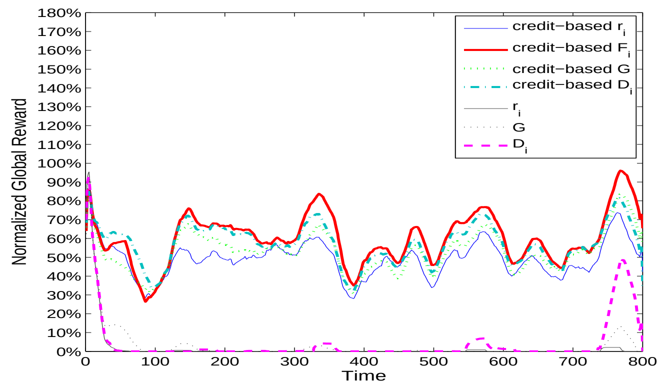

Figure 6.11: Normalized global reward of all objective functions with and without credit-based resource allocation: λ/μ=0.1, κ=1000, η=40% (dynamic with PUs case).

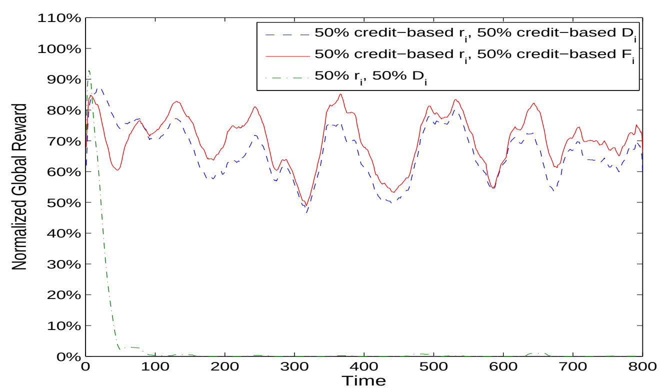

Figure 6.12: Normalized global reward of all objective functions with and without credit-based resource allocation when users chase different objective functions: λ/μ=0.1, κ=1000, η=20% (dynamic with PUs case).

6.3.1 Optimality

We show in Figs. 6.10 and 6.11 the normalized global reward of r_i, G, F_i, and D_i with and without the proposed credit-based resource allocation technique for different PUs loads. It is clear that even when considering primary users' activities, the proposed credit-based technique still improves significantly the achievable system rewards, and this is regardless of the objective function being used. This is true in both cases: η=20% and η=40%. We

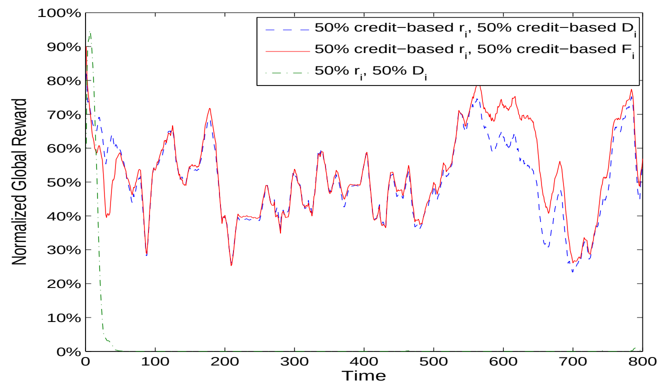

Figure 6.13: Normalized global reward of all objective functions with and without credit-based resource allocation when users chase different objective functions: λ/μ=0.1, κ=1000, η=40% (dynamic with PUs case).

also observe that our proposed function does lead to better performance when compared with that achievable under the D_i function even when D_i uses the credit-based technique.

We want to mention that the fluctuations in the achievable performance are due to the random presence and return of PUs to their channels, which makes the system unavailable for SUs from time to time. During the presence of PUs, the performance in terms of achievable rewards degrades since SUs must leave the system upon the return of any PUs.

6.3.2 Robustness against objective function choice

In this subsection, we show that the proposed credit-based technique minimizes the effect of the objective function choice on the system performance even in the presence of PUs. Figs. 6.12 and 6.13 show the system performance achievable with and without the credit-based technique when 50% of users pursue the F_i or D_i functions while the other half pursues r_i. Again, two observations could be drawn: At first, observe that the achievable performance when the credit-based technique is used is insensitive to the users' objective function choices. That is, when there are selfish users in the system that choose to go after their intrinsic objectives, employing our proposed credit-based technique with the objective functions prevents the overall system performance from degrading. For example, observe that when 50% of the users use either D_i or F_i function while the rest use the r_i function (e.g.: plot: "50% *credit-based* r_i, 50% *credit-based* F_i"), the use of the credit-based approach moves the overall achievable rewards from almost zero to about 70% (when η=20%, Fig. 6.12) and 40% (when η=40%, Fig. 6.13). Also, observe that the F_i function is more robust against malicious behavior than the D_i function, but as PUs load (η) increases, the performance difference becomes smaller.

Also as mentioned in the previous subsection, the channel outage due to the return of PUs to their channels results in fluctuating performance behaviors, since SUs cannot access the band when PUs are using it. Such PUs' activities do, however, as expected affect the performance achievable under any of the objective functions.

6.3.3 Fairness

Figs. 6.14 shows the CoV of users' normalized received rewards under D_i without the credit-based technique, as well as under r_i, G, D_i, and F_i with the credit-based technique for different average numbers of secondary users, 400, 700, and 1000, while considering primary users' activities with $\eta = 20\%$. We again observe that even in the presence of primary users, the use of our proposed credit-based technique reduces substantially the CoV value achievable under the D_i function, especially for high average numbers of users. We also observe that the proposed credit-based technique when coupled with the proposed F_i function outperforms all other functions, including the D_i function even when using the credit-based technique.

It is worth mentioning that nearly similar CoVs are achieved under different objective functions, and regardless of average number of users. That is due the fact that the credit system splits the spectrum among existing users in a way that no matter what their objective is, they will not take more than their fair-shares. The same applies for different average number of users.

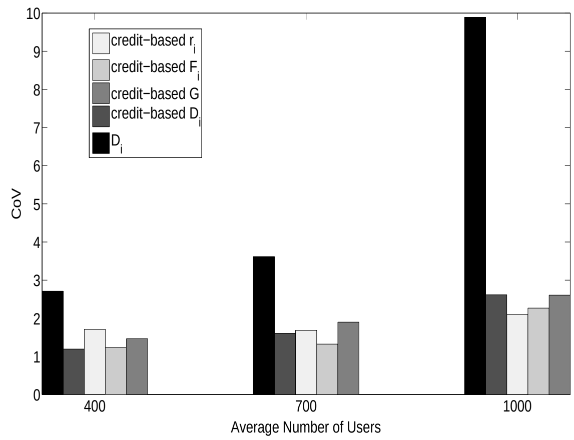

Figure 6.14: Coefficient of variation of users' normalized received rewards under different objective functions: η=20% (dynamic case with PUs).

Chapter 7: Discussion

We have seen that the use of the proposed credit-based technique makes learning techniques robust against objective function choice and improves system fairness while still achieving high system rewards. Without this technique, the D_i function, on the other hand, does achieve good rewards as well, but only when all users use it as their objectives. In other words, if some (or all) users pursue other objectives, the overall system performance can degrade substantially. In addition, ensuring fairness can be very challenging due to the way users end up distributing themselves among the channels. One key advantage of the D_i function, however, lies in its fully distributed capability; it can be implemented and fully realized without needing any cooperation or centralized entity[1]. Ours can still be viewed as a distributed technique in the sense that users can still choose and switch to their bands on their own will and without having any third entity tell them to do so. However, since our approach relies on and accounts for what users have received in the past to be able to decide what should be allocated in the future so that misbehavior can be prevented, it requires the deployment of access points to keep track of and monitor users' activities. This, however, is not unrealistic and can be done with minimum overhead.

[1] This depends to some extent on the network topology and on other factors as well [12].

Chapter 8: Conclusion

This paper proposes a credit-based resource allocation technique that improves fairness among spectrum users and combats malicious and selfish behavior in DSA systems. The proposed technique reduces the effect of user's objective function choice on the system performance. It is robust against misbehavior and achieves good performance independently of what users choose to pursue as their objectives. It also improves fairness by allocating equal amounts of spectrum service to users. We also propose fairness-aware objective function that when coupled with credit-based technique, the overall fairness is improved even further.

Bibliography

[1] T. AlShammari, B. Hamdaoui, M. Guizani, and A. Rayes. Overcoming user selfishness in dsa systems through credit-based resource allocation. *In Proc. of IEEE International Conference on Communications*, June 2014.

[2] L. Cao and H. Zheng. Balancing reliability and utilization in dynamic spectrum access. *IEEE/ACM Transcations on Networking*, 20:651–661, June 2012.

[3] B. Hamdaoui, M. NoroozOliaee, K. Tumer, and A. Rayes. Coordinating secondary-user behaviors for inelastic traffic reward maximization in large-scale DSA networks. *IEEE Transactions on Network and Service Management*, December 2012.

[4] S. Haykin. Cognitive radio: Brain-empowered wireless communications. *IEEE Journal on Selected Areas in Communications*, 23(2):201–219, Feb 2005.

[5] W. Horne, P. Weed, and D. Schaefer. Adaptive spectrum radio: a feasibility platform on the path to dynamic spectrum access. *Proc. of Int'l Symp. on Advanced Radio Technologies*, March 2003.

[6] H. Li and Z. Han. Catch me if you can: An abnormality detection approach for collaborative spectrum sensing in cognitive radio networks. *IEEE Transactions on Wireless Communications*, 9:3554–3565, November 2010.

[7] X. Li, H. S., Z. Han, and Z. Wu. Bayesian compressed sensing based dynamic joint spectrum sensing and primary user localization for dynamic spectrum access. *IEEE Global Telecommunications Conference*, pages 1–5, December 2011.

[8] Y. Lin, K. Liu, and H. Hsieh. On using interference-aware spectrum sensing for dynamic spectrum access in cognitive radio networks. *IEEE Tran. on Mobile Computing*, 12:461–474, March 2013.

[9] K. Liu and Q. Zhao. Distributed learning in cognitive radio networks: multi-armed brandit with distributed mulitple players. *IEEE International Conference on Acousitcs, Speech, and Signal Processing*, 2010.

[10] M. NoroozOliaee, B. Hamdaoui, and M. Guizani. Maximizing secondary-user satisfaction in large-scale dsa systems through distributed team cooperation. *IEEE Transcations on Wireless Communications*, 11(10):3588–3597, October 2012.

[11] M. NoroozOliaee, B. Hamdaoui, and M. Guizani. Adaptive service function for system reward maximization under elastic traffic model. *In Proc. of IEEE Global Communications Conf.*, December 2013.

[12] M. NoroozOliaee, B. Hamdaoui, and K. Tumer. Efficient objective functions for coordinated learning in large-scale distributed OSA systems. *IEEE Transactions on Mobile Computing*, 12(5):931–944, May 2013.

[13] R.S. Sutton and A.G. Barto. *Reinforcement learning: An intoduction*. MIT, 1998.

[14] K. Tumer and A. Agogino. Distributed agent-based air traffic flow management. In *Proceedings of the Sixth International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 330–337, Honolulu, HI, May 2007.

[15] K. Tumer and D. Wolpert, editors. *Collectives and the Design of Complex Systems*. Springer, New York, 2004.

[16] J. Unnikrishnan and V. V. Veeravalli. Algorithms for dynamic spectrum access with learning for cognitive radio. *IEEE Transactions on Signal Processing*, 58(2), August 2010.

[17] P. Venkatraman, B. Hamdaoui, and M. Guizani. Opportunistic bandwidth sharing through reinforcement learning. *IEEE Tranactions on Vehicular Technology*, 59(6):3148–3153, July 2010.

[18] W. Wang, H. Li, S. Y.L., and Z. Han. Attack-proof collaborative spectrum sensing in cognitive radio networks. *IEEE Annual Conference on Information Sciences and Systems*, pages 130–134, March 2009.

[19] T. Weiss and F. Jondral. Spectrum pooling: An innovative strategy for the enhancement of spectrum efficiency. *IEEE Communication Magazine*, 42:S8–14, March 2004.