

# Utility-based Dynamic Spectrum Aggregation algorithm in Cognitive Radio Networks

Haeyoung Lee, Seiamak Vahid and Klaus Moessner

Centre for Communication Systems Research (CCSR), University of Surrey, United Kingdom

{Haeyoung.Lee, S.Vahid, K.Moessner}@surrey.ac.uk

**Abstract**— In this paper, we propose a utility-based spectrum aggregation algorithm to enhance the performance of a cognitive radio network considering multiple objectives: (i) maximization of overall throughput, (ii) reduction of channel switching, (iii) reducing the number of sub-channels comprising the aggregate channel, aimed at opportunistic spectrum use by secondary users (SUs). These three objectives are integrated into a weighted sum utility function. The weight associated with each objective can be set differently (typically done manually) depending on the metric to be optimized. In this article however, we propose and evaluate an automatic mechanism for setting weights. The proposed algorithm including the learning module allows for automated (no manual intervention) adaptable setting of objective-function weights depending on the environment changes and its performance is also shown via the simulation results.

**Index Terms**—Spectrum aggregation, Opportunistic spectrum use, Dynamic spectrum allocation, Q-learning

## I. INTRODUCTION

Underutilization of some licensed spectrum has motivated research into the opportunistic spectrum use by cognitive radios (CRs) [1][2]. Through spectrum sensing for example, CRs are able to identify and utilize discrete available spectrum opportunities. Since the spectrum opportunities could change over time [1], CRs need to adapt operation to spectrum availability. Thus, the problem of selection and allocation of spectrum in cognitive radio networks poses new challenges that do not arise in traditional wireless technologies.

When appearance of primary users (PUs) is detected, communication of secondary users (SUs) has to be interrupted and data to transmit must be buffered. The communication can be resumed when the connection is successfully moved to the new channel. Thus, channel switching could lead to large queuing delays and potential packets losses [3]. Moreover, PUs can experience short-term interference to transmissions before being detected by neighboring SUs [4]. Considering the disruption caused by channel switching, prediction/estimation of future spectrum availability based on monitoring current/historical PU activity has been investigated in literature [3][4][5][6].

The spectrum opportunities identified by CR users could be too narrow to support required bandwidths [7]. In this case, multiple spectrum holes can be utilized simultaneously for cognitive radios through spectrum aggregation (SA). With advances in OFDM-based techniques, for example, discontinuous orthogonal frequency division multiplexing (DOFDM), will support spectrum aggregation in a single radio

frequency (RF) unit [8]. Although spectrum aggregation does indeed contribute to better spectrum utilization, it can lead to system overheads and complexity. If the level of spectrum fragmentation is serious, transmitters might be required to aggregate many small sub-channels and then excessive filtering and/or guard bands to protect adjacent users would be required [9]. Increasing the number of sub-channels will result in increasing channel search times [9]. Thus, the number of sub-channels composing the aggregate channel should be small [7].

Spectrum allocation with aggregation capability has also been investigated recently. The aggregation algorithms proposed aim to increase the spectrum utilization level given the spectrum aggregation range supported by the hardware [6][10][11]. In [12], in order to measure the channel fragmentation impacts, a number of algorithms are proposed and evaluated. Whereas aforementioned works have assumed that the channel quality is the same for different users, the work in [7] exploits difference in channel quality in designing the spectrum aggregation scheme. The proposed aggregation algorithm selects the spectrum with the highest spectral efficiency to maximize the total throughput [7]. Making use of the different channel characteristics is particularly beneficial to allocation of spectrum/channel composed of small number of sub-channels and narrower bandwidths to users. As the bandwidth increases, the power consumption increases and operation in wider spectrum could increase the possibility of channel switching. When considering the spectrum allocation to cognitive radio users having the spectrum aggregation capability, many factors should therefore be considered simultaneously.

To the best of our knowledge however, aforementioned factors have not been considered together (under single formulation) within the scope of spectrum aggregation algorithms for opportunistic spectrum access in the literature. In this paper, we propose a new multi-objective spectrum aggregation algorithm that aims to maximize the total throughput, minimize channel switching, and reduce the number of sub-channels in the composite/aggregate channel. Three objectives are integrated into a weighted sum [13] utility function. The weights associated with each objective can be different depending on the performance metric of interest. In this paper, we also propose the framework for spectrum aggregation/allocation with the learning technique. The system continuously monitors the performance of each individual objective and decides to trigger the learning module for the case of the degraded performance. As a result of learning, the

optimal weight values are chosen and applied to the spectrum aggregation method. Thus, the weight values are adaptively set/learned based on the radio environment encountered so that performance metrics can be satisfied given the thresholds.

The rest of this paper is organized as follows. In section II, the system model is presented. The proposed algorithm is explained in section III. The framework of spectrum allocation for the proposed algorithm is also proposed in Section III. After analyzing the results of the proposed algorithm by the simulation in section IV, conclusions are presented in section V.

## II. SYSTEM MODEL AND ASSUMPTIONS

In a cognitive radio network, assume that there are  $N$  secondary users with their request transmission rate known as  $R_i^{req}$  bps for  $i=\{1,2,\dots,N\}$ . We also define the following notations:

- Channel: is considered the basic resource unit - Based on the PUs' activity pattern, the spectrum is channelized logically into basic units of equal bandwidth.
- Sub-Channel: is a grouping of contiguous idle channels and combination of a number of sub-channels constitutes an aggregate channel.
- Aggregate channel: It is a collection/pool of sub-channels that is allocated to secondary users.

The spectrum of total bandwidth  $BW_{total}$  MHz is channelized into  $C$  channels of bandwidth  $BW_c$ . The discrete channel occupancy pattern by PU's spectrum usage is modeled as an ON/OFF random process [4][14] as depicted in Figure 1. The occupancy pattern of channel  $i$  is represented as the sequences of activity/inactivity period lengths,  $ON_i$  and  $OFF_i$ , respectively. The primary activity patterns are assumed to be independent across channels and its statistical information is stored in a database such as Radio Environment Map (REM) [15] for adequate observation time to achieve a good level of convergence. Then, with the knowledge about the PU traffic retained in the database, the average duration of the  $OFF_i$  period,  $E(OFF_i)$  and so-far observed OFF duration of idle channel  $i$ ,  $T_{Cnt\_Idle}^i$  could be tracked and remaining free time,  $T_{Re\_Idle}^i$ , can be estimated in Eq. (1) [14].

$$T_{Re\_Idle}^i = E(OFF_i) - T_{Cnt\_Idle}^i \quad (1)$$

There are  $M$  idle channels and the estimated remaining idle time for idle channels is used to select the aggregate spectrum. For simplicity, it is assumed that consecutive ON/OFF periods are not correlated in this paper. For the case that there is correlation between consecutive ON/OFF periods,  $E(OFF_i/ON_i)$  can be used instead of  $E(OFF_i)$  in Eq. (1) [14].

While the availability of spectrum opportunities is varying, spectrum opportunities are assumed to be identified without errors. We further assume that spectrum occupancy status does not change during spectrum aggregation/allocation intervals.

In order to utilize differences in channel quality, it is assumed that the channel characteristics for all nodes on all channels are known. By adopting Adaptive Modulation and Coding (AMC) with OFDM technique for instance, the supported transmission rate,  $R_{i,j}$ , for any secondary user  $i$  on

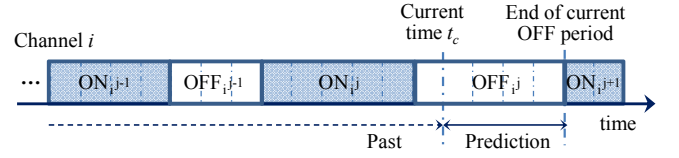


Figure 1. On-Off Channel Occupancy model

channel  $j$  could be determined. If channel  $j$  is occupied by PUs, set  $R_{i,j}=0$  for all  $i$  [7]. The information on spectrum availability and channel quality and channel occupancy pattern are used as input to the spectrum aggregation algorithm. In order to satisfy the secondary user's throughput requirement, multiple channels can thus be aggregated.

## III. MULTI-OBJECTIVE UTILITY-BASED SPECTRUM AGGREGATION ALGORITHM

The proposed utility-based spectrum aggregation algorithm is presented in this section. The approach focuses on finding the most appropriate aggregate channel consisting of multiple sub-channels for each CR node whilst taking into account the following three objectives: to maximize total throughput (i.e. spectrum utilization), to reduce channel switching, and to reduce the number of sub-channels comprising the aggregate channel. The approach of a weighted sum utility of the objectives is used to solve the multi-objective problem. The utility value of each objective for each channel and each user is calculated and the total utility value is calculated as a weighted sum approach with the weight vector  $\{w_1, w_2, w_3\}$  as shown in Eq. (2). The aggregation algorithm finds the aggregate channel satisfying the user's required throughput. The problem is formulated as shown in Eq. (2).

$$F(U) = \sum_{k=1}^3 w_k \cdot \sum_{i=1}^N \sum_{j=1}^M a_{i,j} \cdot u_{k,i,j} \quad (2)$$

where  $a_{i,j}$  is the binary channel assignment indicator. It denotes whether user  $i$  occupies channel  $j$ . If  $a_{i,j}=1$ , channel  $j$  is assigned to user  $i$ , otherwise, is not.  $u_{k,i,j}$  is the normalized utility value of  $k^{th}$  objective when  $j^{th}$  channel is allocated to  $i^{th}$  user, and  $w_k$  is the weight of the  $k^{th}$  objective.

### 1) Throughput utility function for $u_{1,i,j}$

In order to allocate the channel to users with good channel quality, the utility value is calculated in Eq. (7) where  $R_{i,j}$  is the transmission rate of channel  $j$  used by a node  $i$  and  $R_{max}$  is the maximum value of  $R_{i,j}$  for all users and all channels. The modified hyperbolic tangent function of Eq. (10) explained in [16] is exploited. The threshold  $\eta$  and spread parameter  $\sigma$  are chosen such that i) the utility is 0.95 when the metric  $x$  achieves the target  $x_0$  and ii) the utility is 0.05 when the metric is one decade away. This function is monotonic and bounded by zero and one. Since higher value is calculated for the pair of users and channels of good quality, this objective will contribute to improving the overall throughput.

### 2) Channel switching utility function for $u_{2,i,j}$

The identified primary behavior is exploited in the selection of the channel with the least-likelihood of the appearance of PUs, thus reducing channel switching. For each idle channel,

the remaining idle time is estimated as denoted in Eq. (1). The utility values for all idle channels are also calculated with the modified hyperbolic tangent function as shown in Eq. (8). Since the utility value of this objective depends on the statistical channel usage pattern, the utility value of the certain channel is the same for all users, that is,  $u_{2,i-1,j} = u_{2,i,j} (1 \leq i \leq N)$ .

### 3) Channel fragmentation utility function for $u_{3,i,j}$

In order to reduce the number of sub-channels, the contiguous channel allocation is preferred to non-contiguous channel allocation. When the algorithm considers allocating a channel  $j$  for the user  $i$ , it decides the type of spectrum aggregation. When it considers allocating idle channel  $j$  with the adjacent channel  $j-1$  or  $j+1$  together, it can be considered as contiguous spectrum aggregation. In this case, the utility value becomes 0.95, the maximum utility value same as the other objectives. If the aggregation type is non-contiguous aggregation, 0.05 will be allocated as the utility value as shown in Eq. (9).

The spectrum allocation with aggregation based on the multi-objective utility function is formulated as following.

For given  $N, M, BW_c, R_{i,j}, T_{Re\_Idle\_i},$  and  $R_i^{req}$ ,

$$A^* : \arg \max_A F(U) \quad (3)$$

$$F(U) = \sum_{i=1}^N \sum_{j=1}^M a_{i,j} \cdot (w_1 u_{1,i,j} + w_2 u_{2,i,j} + w_3 u_{3,i,j}) \quad (4)$$

subject to

$$a_{i,j} = \{0, 1\} \text{ and } \sum_{j=1}^M a_{i,j} \leq 1, \forall i, j \quad (5)$$

$$\sum_{k=1}^3 w_k = 1, 0 < w_k < 1 \quad (6)$$

$$u_{1,i,j} = f(R_{i,j}, R_{\max}; \eta, \sigma), \forall i, j \quad (7)$$

$$u_{2,i,j} = f(T_{Re\_Idle\_j}, T_{Re\_Idle\_max}; \eta, \sigma), \forall i, j \quad (8)$$

$$u_{3,i,j} = \{0.95, 0.05\}, \forall i, j \quad (9)$$

$$f(x, x_0; \eta, \sigma) = (1/2) \{ \tanh[\log(x/x_0) - \eta] \sigma + 1 \} \quad (10)$$

$$BW_c \times (C_{\max}^i - C_{\min}^i + 1) \leq BW_{Agg}, \forall i$$

$$C_{\max}^i = \max(j) \text{ for } a_{i,j} = 1, C_{\min}^i = \min(j) \text{ for } a_{i,j} = 1, \forall i \quad (11)$$

To solve the problem in Eq. (3), the proposed algorithm selects sequentially each channel that will experience the highest utility for the aggregate channel for each user until the set of selected channels satisfies users' requested throughput.

### 4) Automated setting of weights

Assume that the system has set/pre-defined thresholds for each performance metric (i.e. total system throughput  $TH_C$ , the number of channel switching  $SW_C$ , and the number of sub-channels  $SUB_C$ ) based on system level Key Performance Indicators (KPIs) to be met. The problem to solve is then the selection of optimal set of weights of multi-objective utility in

Eq. (4) so that the performance of each objective remains close as possible to the pre-defined thresholds. Figure 2 shows the framework of spectrum aggregation including the module for automatic selection of weights via Q-learning, an off-policy Reinforcement Learning (RL) algorithm [17]. The system continuously monitors the spectrum aggregation & allocation results (monitoring phase), and triggers the learning module when the performance is not met/thresholds crossed (triggering phase). Through the learning process (learning phase), the set of optimal weight vector is found (selection phase) and used by the spectrum aggregation & allocation algorithm. Thus the cycle of monitoring, triggering, learning and selection repeats. The learning model to decide the weight values comprises the possible states of the environment  $S$ , the set of possible actions  $A$ , the reward function  $R$ , and the policy  $\pi: S \rightarrow A$ . Q-learning learns the optimal policy through simple Q-value iterations to take note of recent policy success/failure as feedback and the weighted average of past values observed. The details on Q-learning can be referred to [18]. In the considered scenario, the elements of learning model are defined as follows:

1) *States*: The state is the monitored average performance, total throughput  $TH_t$ , channel switching  $SW_t$ , and the number of sub-channels  $SUB_t$ .  $s_t = (TH_t, SW_t, SUB_t)$  (12)

2) *Actions*: the weights vector set

$$a_t = (w_1, w_2, w_3) \text{ where } 0 < w_i < 1 \text{ and } w_1 + w_2 + w_3 = 1 \quad (13)$$

3) *Rewards*: The reward of applying a certain set of weights is the satisfaction degree compared to the predefined criteria. It is calculated by the following Eq. (14).

$$r_t = 1 + \max \left\{ 0, \frac{TH_t}{TH_C} - 1 \right\} + \max \left\{ 0, 1 - \frac{SW_t}{SW_C} \right\} + \max \left\{ 0, 1 - \frac{SUB_t}{SUB_C} \right\} \quad (14)$$

When the achieved performance (i.e.  $TH_t, SW_t$ , and  $SUB_t$ ) meets the criteria/thresholds (i.e.  $TH_C, SW_C$ , and  $SUB_C$ ), the reward is one as the maximum. Otherwise, it will have the value reflected to the satisfaction degree.

## IV. SIMULATION AND PERFORMANCE ANALYSIS

This section presents evaluation results of the multi-objective utility-based spectrum selection with aggregation technique. In our simulation, we assume a total available spectrum of 30 MHz. The PU activity is modeled through an ON/OFF process for the unit of channel of 200 kHz bandwidth.

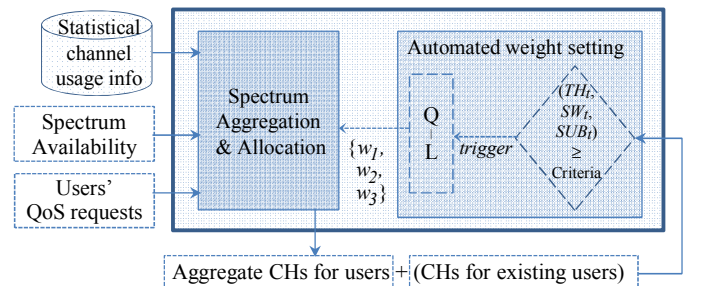


Figure 2. The framework of spectrum aggregation & allocation (for SUs) using Reinforcement learning

Requested throughput (average throughput 5 Mb/s assumed) of the secondary users are drawn randomly from a uniform distribution, and the service time of resource requests is set to a mean of 5 seconds with exponential distribution. The configuration parameters used are shown in Table 1.

In the simulations, we start off by assigning the same priority to three different objectives, i.e. the weight vector of multi-objective utility  $\{w_1, w_2, w_3\}$  is set as  $\{1/3, 1/3, 1/3\}$ . However, the weight of each objective can be set differently depending on metric/objective to be optimized. Performance evaluations in the following sections compare performance of the random aggregation method (labeled as “Random”) with different settings of the weight vector.

Initial results in Figure 3, 4, and 5 represent performance prior to integration of the learning technique. Figure 3 presents the normalized total throughput. It compares the case with equal-weight setting  $\{1/3, 1/3, 1/3\}$  against the optimal setting for  $\{1, 0, 0\}$  that allocates the channel with the best quality and thus is optimal from the perspective of throughput. It is observed that the proposed equal-weight setting outperforms Random algorithm and can reach more than 90% performance of the optimal setting. Channel switching performance is evaluated in Figure 4. The algorithm which only considers the remaining idle time through the setting  $\{0, 1, 0\}$  becomes the optimal algorithm for channel switching. The equal-weight setting results in more channel switching but presents better performance than the Random aggregation in Figure 4. Lastly, Figure 5 evaluates the number of sub-channels. The algorithm which only considers the sub-channel number through the setting  $\{0, 0, 1\}$  is the optimal from the perspective of this criterion. However, the equal-weight setting achieves a performance close to this optimum.

As the number of users increases, the total throughput also increases. However, when the number of users is over 15, it becomes saturated due to a lack of available spectrum. As the number of resource requests increases, the probability that channels having short remaining idle time are utilized increases and this leads to increase in the channel switching. However, due to increasing competition between users, the amount of spectrum resource allocated to users is reduced. Thus, smaller number of sub-channels is aggregated into an aggregate channel.

The proposed method for the automated setting of weights

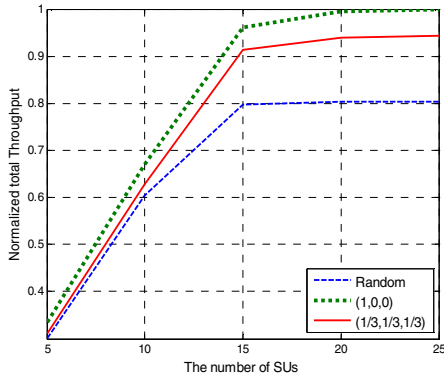


Figure 3. Normalized System Throughput

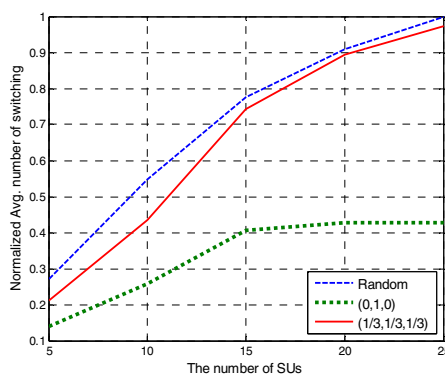


Figure 4. Normalized number of Channel Switching

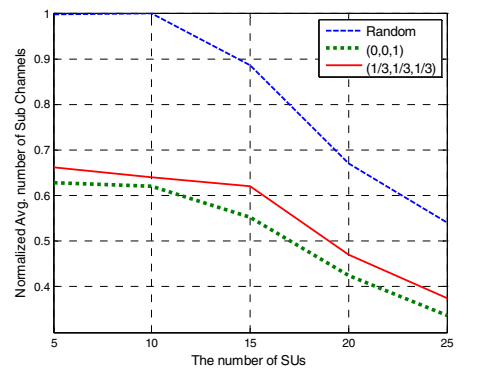


Figure 5. Normalized number of Sub-channels

Table 1. The configuration of the simulation

Parameters		Values
Spectrum	Total Spectrum	30 MHz
	BW of a CH, $BW_c$	i. 200 kHz, ii. 250kHz
	PU's ON/OFF pattern of CHs	Set i. $\{(1/13), (2/12), \dots, (1/13)\}$ , $E[\text{Off}] = 7$ Set ii. $\{(1/5), (2/4), \dots, (5/1)\}$ , $E[\text{Off}] = 3$ (Uniform dist.)
	Modulation, $Mod$	8PSK, 16QAM, 32QAM [7] Set i. Uniform dist, Set ii. distribution prob. = $\{0.3, 0.6, 0.1\}$
SUs	Service time	5 seconds (Exponential distribution)
	Requested Th.	5 Mb/s (Uniform distribution)
T-Q	learning rate, $\alpha$	$\alpha(s_t, a_t) = 1/\beta(s_t, a_t)$ [18]
	Exploration prob. $\epsilon$	Initially 1 and decrease as time increase [19]

relies on pre-set thresholds per objective. Although at the outset system KPIs can be used to determine and set the required performance thresholds, in our simulations as depicted in Figure 6, the average performance of the first observation interval (period 0 –  $T_A$ ) plus a 10% margin is used to set the actual thresholds for the channel switching and the number of sub-channels. Since the number of users requesting resources is 15, the threshold for throughput performance is then set as the sum of requested throughputs i.e. 75 Mbps. The equal-weight setting is used to set the initial weight values. During the period  $T_A$  to  $T_B$ , an increasing number of channel switchings is detected at time  $T_B$ . (actually triggered by adjusting PU's ON/OFF pattern from set i→ii as depicted in Table 1). Then the learning algorithm is run during learning phase i.e. period  $T_B$ - $T_C$ , and the optimal weight vector is obtained. Then, the selected weight vector is used by the spectrum aggregation algorithms at  $T_C$  at which point the monitoring phase also begins. While the performance of “No-Learning” from  $T_C$  to  $T_D$  shows no improvement compared to the performance from  $T_A$  to  $T_B$ , it is observed that the proposed approach using RL learning does improve the channel switching performance. Although the corresponding period throughput performance does not show any major changes, the number of sub-channels within period  $T_C$ - $T_D$  is degraded and shows an increase to a level close to the set threshold. This is because the proposed spectrum aggregation



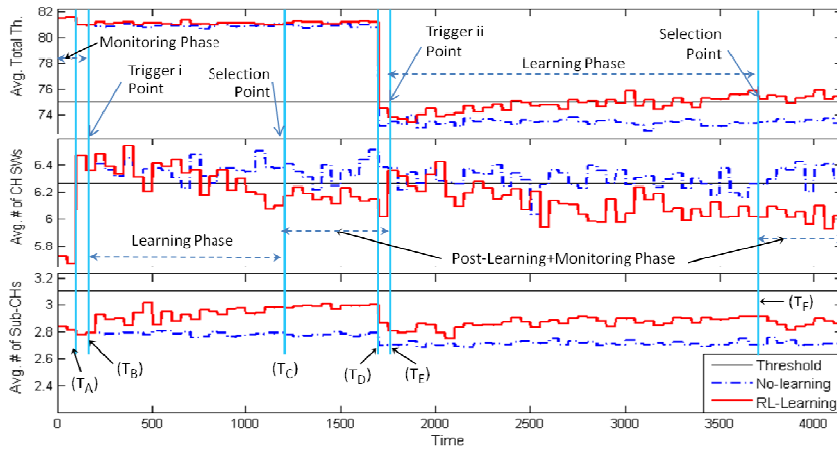


Figure 6. Performance comparison of RL-Learning and No-Learning approach

algorithm prefers to select the sub-channels with lower remaining idle time although more sub-channels may be required. From time  $T_D$  to  $T_E$ , i.e. during the next monitoring phase, throughput performance is degraded (actually triggered through adjusting channel quality, by going from *Mod* set  $i \rightarrow ii$  in Table 1). The learning process is triggered at time  $T_E$  and following the new learning phase from  $T_E$  to  $T_F$ . Then the new optimal weights are used by the proposed spectrum aggregation algorithm. In this case, the obtained improvement in throughput is accompanied by a reduction in number of channel switchings as result of leaning process.

## V. SUMMARY AND CONCLUSION

In this paper, a utility-based multi-objective spectrum aggregation algorithm is proposed. Three objectives (max. throughput, min. channel switching and min. of the number sub-channels) are integrated into a weighted sum utility function. Since the optimal weights setting can be different depending on the performance metric of interest, a RL learning approach is used to determine the optimal weights. The proposed approach is shown to be adaptable to changes amongst different objectives of interest. The proposed framework including a learning module allows for the management of complex interactions and trade-offs between different metrics without manual intervention.

## ACKNOWLEDGMENT

This work is performed in the framework of the European-Union funded project OneFIT ([www.ict-onefit.eu](http://www.ict-onefit.eu)) and ACROPOLIS ([www.ict-acropolis.eu](http://www.ict-acropolis.eu)). The projects are supported by the European Community's Seventh Framework Program (FP7). The views expressed in this document do not necessarily represent the views of the complete consortium.

## REFERENCES

- [1] I. F. Akyildiz, W.-Y. Lee, M.C. Vuran, and S. Mohanty, "NeXt generation /dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer Networks Journal (Elsevier)*, Sept. 2006
- [2] S. Haykin, "Cognitive Radio: Brain-Empowered Wireless Communications," *IEEE Journal on Selected Areas in Communicatoin*s, vol. 23, no. 3, pp. 201-220, Feb. 2005
- [3] A. Lertsinsruttavee, N. Malouch, and S. Fdida, "Spectrum Handoff Strategies for Multiple Channels Cognitive Radio Network", *ACM CoNEXT Conf.*, Philadelphia, 2010, USA
- [4] L. Yang, L. Cao, and H. Zheng, "Proactive channel access in dynamic spectrum networks," *Phys. Commun.*, vol. 1, no. 2, pp. 103-111, 2008.
- [5] P. A. K. Acharya, S. Singh, and H. Zheng, "Reliable Open Spectrum Communications through Proactive Spectrum Access", *TAPAS 2006*.
- [6] F. Huang, W. Wang, H. Luo, G. Yu, and Z. Zhang, "Prediction-based Spectrum Aggregation with Hardware Limitation in Cognitive Radio Networks" *IEEE 71st Vehicular Technology Conference*, 2010
- [7] J. Lin, L. Shen, N. Bao, B. Su, Z. Deng, and D. Wang, "Channel Characteristic Aware Spectrum Aggregation Algorithm in Cognitive Radio Networks," *IEEE 36th Conference on Local Computer Networks (LCN)*, pp. 634 - 639 , 4-7 Oct. 2011
- [8] J. D. Poston and W. D. Horne, "Discontiguous OFDM considerations for dynamic spectrum access in idle TV channels," *IEEE Int. Symp. New Frontiers Dynamic Spectrum Access Networks (DySPAN)*, Nov.2005
- [9] A. Shukla, B. Williamson, J. Burns, E. Burbidge, A. Taylor, and D. Robinson, "A Study for the Provision of Aggregation of Frequency to Provide Wider Bandwidth Services," *QnetIQ*, August 2006
- [10] D. Chen, Q. Zhang, and W. Jia, "Aggregation Aware Spectrum Assignment in Cognitive Ad-hoc Networks" *CrownCom 2008*
- [11] J. Li, Z. Tan, C. Tao, and S. Xu, "A new spectrum aggregation algorithm for IMT-Advanced based on cognitive science", *Wireless Communications and Signal Processing (WCSP)*, 21-23 Oct. 2010
- [12] E. Coffman, P. Robert, F. Simatos, S. Tarumi, and G. Zussman, "Channel fragmentation in dynamic spectrum access systems: A theoretical study" *ACM SIGMETRICS Perf. Eval. Review*, June 2010
- [13] R. T. Marler and J. S. Arora, "Surey of multi-objective optimization methods for engineering," *Structural and Multidisciplinary Optimization*, vol. 26, no. 6, pp. 369-395, 2004.
- [14] F. Bouali, O. Sallent, J. P. Romero, and R. Agusti, "Strengthening radio environment maps with primary-user statistical patterns for enhancing cognitive radio operation," *CROWNCOM 2011*
- [15] Y. Zhao, L. Morales, J. Gaedert, K. K. Bae, J.-S. Um, and J. H. Reed, "Applying radio environment maps to cognitive wireless regional area networks," *DySPAN 2007*, april 2007, pp. 115-118
- [16] Y. Zhao, S. Mao, J. Neel, and J. Reed, "Performance evaluation of cognitive radios: Metrics, utility functions, and methodology," *Proc. IEEE*, vol. 97, no. 4, pp. 642-659, Apr. 2009.
- [17] J. Nie and S. Haykin, "A Dynamic Channel Assignment Policy Through Q-Learning", *IEEE Transaction on Neural Networks*, vol. 10, Nov. 1999
- [18] Richard S. Sutton and Andrew G. B. *Reinforcement learning : an introduction*, MIT Press, Cambridge, MA, 1998
- [19] Y. Yao and Z. Feng, "Centralized Channel and Power Allocation for Cognitive Radio Networks: A Q-learning solution", *Future Network and Mobile Summit*, 2010