Meet an Fantastic Sibyl: A Powerful Model in Cognitive Radio Networks

Wei Yang(✉), Xiaojun Jing, and Hai Huang

School of Information and Communication Engineering,
Beijing University of Posts and Telecommunications, Beijing 100876, China
{yangweibupt, jxiaojun, huanghai}@bupt.edu.cn

Abstract. Dynamic spectrum access is challenging, since an individual secondary user usually just has limited sensing abilities. One key insight is that primary user emergence forecasting among secondary users can help to make the most of the inherent association structure in both time and space, it also enables users to obtain more informed spectrum opportunities. Therefore, primary user presence forecasting is vital to cognitive radio networks (CRNs). With this insight, an auto regressive enhanced primary user emergence reasoning (AR-PUER) model for the occurrence of primary user prediction is derived in this paper. The proposed method combines linear prediction and primary user emergence reasoning. Historical samples are selected to train the AR-PUER model in order to capture the current distinction pattern of primary user. The training samples of the primary user emergence reasoning (PUER) model are combined with the recent samples of auto regressive (AR) model tracking recent parallel. Our scheme does not require the knowledge of the signal or of the noise power. Furthermore, the proposed model in this paper is blind in the detection that it does not require information about the channel. To verify the performance of the proposed model, we apply it to the data during the past two months, and then compare it with other method. The simulation results demonstrate that the AR-PUER model is effective and generates the most accurate forecasting of primary user occasion in several cases. Besides, it also performs much better than the commonly used energy detector, which usually suffers from the noise uncertainty problem.

Keywords: Dynamic spectrum access · Linear prediction
AR-PUER · CRNs

1 Introduction

With the express development of wireless communication system, it posed a tough requirement for the limited radio spectrum resource. On the other hand, in connection with the spectrum report conducted by the Federal Communications Commission (FCC), the majority of the radio spectrum is not in use in reality [1]. Therefore, the perception of CRNs has been proposed as a hopeful technology to deal with the spectrum scarcity as well as the spectrum underutilization problem [2]. The merely consideration is that the secondary users have to vacate the channel within a certain amount of time whenever the primary user becomes active. Thus, the cognitive radio

P. Marques et al. (Eds.): CROWNCOM 2017, LNICST 228, pp. 15–25, 2018.
https://doi.org/10.1007/978-3-319-76207-4_2

network faces the tricky challenge of detecting the presence of the primary user, particularly in a low signal-to-noise ratio region, since the signal of the primary user might be severely alleviated due to multipath and shadowing before reaching the secondary user. To ensure that there will be no harmful intervention to the primary user, the secondary users need to sporadically detect the presence of the primary user. There are several factors that avoid the spectrum sensing from operating in a reliable manner. One factor is that the strength of the primary users' signals could be very weak when they reach the secondary users. If the secondary user makes a defective decision in detection and establishes transmitting when the primary users are active, its own signal will meddle with the primary users' signals.

The most popular spectrum sensing methods that have been proposed are energy detection, cyclostationary feature detection and matched-filtering detection. Up to now, not much work has been done on blind sensing models. Algorithms based on cyclostationarity have been developed in [3]. However, the performance of these algorithms at a low SNR has not been investigated; it also requires some prior information about the primary users, whereas the proposed model does not. The proposed model performs much better than the commonly used energy detector. Moreover, unlike the energy detector [4] and prior cyclostationary methods, the novel AR-PUER model here is blind and does not require information about the multipath channel distortions that the primary users has undergone on its way to reaching the secondary user.

In this paper, we proposed an auto regressive enhanced primary user emergence reasoning (AR-PUER) model, which does not require the knowledge of the signal or of the noise power. Moreover, the proposed detection algorithm in this paper is blind in the sense that it does not require information about the multipath channel deformations the primary user has undergone on its way to reaching the secondary user.

The rest of this paper is organized as follows. In Sect. 2, the auto regressive enhanced primary user emergence reasoning (AR-PUER) model is presented for the proposed scheme, which includes the detection method, primary user behavior characteristic and AR-PUER treatment process. In Sect. 3, the performance of proposed approach is analyzed, and numerical results are presented. Finally, Sect. 4 briefly concludes this paper.

2 System Model

2.1 Time Correlation Estimate

In order to study the spectrum usage of primary users, the concept of time division is introduced. Time division belongs to time interval that based on a statistical analysis of the primary users' spectrum over a period of time. For some primary users, its time division may be more than one.

With the purpose of make the calculation results closer to the actual use of primary users, the following content will revise the error combined with statistical methods to make the consequence more accurate. To evaluate the disparity between the primary users' spectrum access time and the average time, the concept of time similarity is proposed, which can be defined as

$$\tau(\alpha, \beta) = 1 - \delta(\alpha, \beta) \tag{1}$$

where $\delta(\alpha, \beta) = \frac{|\alpha - \beta|}{\varepsilon}$, $\xi = |\alpha - \beta|$.

In (1), α is the initiation spectrum access time which independent with the statistical time of primary users, β is the average time that primary user access to the spectrum, while ε indicate the time is twenty-four hours. In the meantime, we classify ξ as time freeness that has two kinds of values:

(1) Primary users' own actual time ξ' is known as the code value of time freeness; in the same way, $\tau(\xi')$ is the code value of time similarity.
(2) According to $\xi = |\alpha - \beta|$, we can calculate the observation value of time freeness ξ, while the observation value of time similarity is $\tau(\xi)$ or $\tau(\alpha, \beta)$.

The code value of time similarity can constitute two intervals $(\tau(\xi'), 1)$ and $(0, \tau(\xi'))$, when the observation value $\tau(\alpha, \beta)$ is valid in $(\tau(\xi'), 1)$, accordingly we call $(\tau(\xi'), 1)$ as the valid interval; as a result, $(0, \tau(\xi'))$ is the invalid interval.

Considering the situation that a primary user maybe repeated access to the spectrum, therefore, the primary user exist multiple time similarity, that is to say, there is more than one time similarity located in the valid interval. Consequently, the time similarity of primary user A at $t_1, t_2, \cdots, t_k, \cdots$ can indicated as $\tau_1\Big|_{t_1}^{A}, \tau_2\Big|_{t_2}^{A}, \ldots, \tau_k\Big|_{t_k}^{A}, \ldots$ respectively. Analogously, we see that $\tau_1(\alpha, t_1)\Big|_{t_1}^{A}, \tau_2(\alpha, t_2)\Big|_{t_2}^{A}, \ldots, \tau_k(\alpha, t_k)\Big|_{t_k}^{A}, \ldots$. In the valid interval, the maximum time similarity can characterized as $\tau_{\max}(\alpha, t_i)\Big|_{t_i}^{A}$, while the minimum time similarity can indicated as $\tau_{\min}(\alpha, t_j)\Big|_{t_j}^{A}$. Hence, they can illustrated as $\tau_{\max}\Big|_{t_i}^{A}$ and $\tau_{\min}\Big|_{t_j}^{A}$ for short.

The above time similarity in which interval is valid is mainly determined by the time freeness ξ', because the standard deviation is the best way to evaluate time fluctuations, we can define its standard deviation as the time freeness ξ', which implies

$$\xi' = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{2}$$

When ξ' is little, the time similarity τ will be relatively large, and the effective interval will be narrowed accordingly, which reveal the primary users' spectrum access behavior has a strong regularity. On the contrary, when ξ' is large, the time similarity τ will be little at once, and the effective interval will be broaden. This phenomenon expose that the primary users' spectrum access behavior has a weak regularity, which reveal a greater volatility.

2.2 Primary User Behavior Characteristic

At first, we define a fixed number of days (m) as criteria. In order to minimize the interference of human factors, the maximum allowed range of time periods should not exceed the code value of time freeness. The primary users' spectrum access duration begin with D, and end up with $D + m$, which located in the area from line R to $R + n$, with the maximum trigger value during the time period $R + n$.

The access behavior employs natural number coding, that is, for the access behavior which length is n, each of them take an integer value from 1 to m (the maximum value). The access behavior is constructed as follows

R	n	D	m

In the process of operation, we need to determine the evaluation of each primary user. The evaluation can be calculated on the basis of assessment function

$$f_k = \sum_{e \in P} \omega(e) \tag{3}$$

where f_k is the evaluation of primary user k, e is the unit grid, P is the enclosed area with primary users' access behavior, and $\omega(e)$ is the value of unit.

In the progress for the primary users' access behavior characteristic, the procedures are as follows:

(1) Setup the initial number of days as $P(0)$ after initialization;
(2) Evaluate the primary users in accordance with the former assessment function (3), and calculate the evaluation value of primary users in $P(t)$;
(3) Carry on interleaved computation, we can obtain $P(t+1)$ from $P(t)$ after the mean value calculation;
(4) Calculate the end conditions, if $t \leq T$, then $t \to t+1$, and next go to step (2); if $t > T$, then output the primary user with maximum time similarity $\tau_{\max}\Big|_{t_i}^{A}$ as the optimal solution, and stop the calculation.

2.3 AR-PUER Model

The sensing model in this paper is based on linear prediction and case based reasoning [5, 6], which involves predicting a future value of a stationary discrete-time stochastic process, given a set of past samples of the process.

Considering the time series of received signal vectors as $\{W_1, W_{t-1}, W_{t-2}, \cdots, W_{t-p}\}$. In the auto regressive $AR(p)$ model, samples $\{W_1, W_{t-1}, W_{t-2}, \cdots, W_{t-p}\}$ are used to predict the condition of primary user a_t. In this paper, forward linear prediction will be used. The forward prediction at time t is denoted by a_t and is given by

$$W_t - \varphi_1 W_{t-1} - \varphi_2 W_{t-2} - \cdots - \varphi_p W_{t-p} = a_t, \;\; t = 0, \pm 1, \pm 2, \cdots \tag{4}$$

The retardation factor B is an operator, which have the arithmetic operation $B^k W_t = W_{t-k}, \;\; k \geq 1$. The stipulation of primary user a_t is then defined as

$$(1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p) \cdot W_t = a_t \tag{5}$$

When $\Phi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p$, the mentioned formula above can illustrated as

$$\Phi(B) \cdot W_t = a_t, \;\; t = 0, \pm 1, \pm 2, \cdots \tag{6}$$

The blind AR-PUER sensing model in this paper combines linear prediction with the primary user emergence reasoning of the received data. Auto regressive model is a widely used technique in numerical analysis and linear algebra [7]. Based on this approach, some signal statistics will be defined, which reveal distinctive features only in the presence of primary users. This discrimination will improve the probability of detection.

In the function $W_t = \Phi^{-1}(B) \cdot a_t$, in which $G(B) = \Phi^{-1}(B) = \sum_{k=0}^{\infty} G_k B^k$, $|B| < 1$, we see that $W_t = \sum_{k=0}^{\infty} G_k a_{t-k}$ where $G_k, \;\; k = 0, 1, 2 \cdots$ is the Green's function of the $AR(p)$ model. Analogously, in $a_t = \Phi(B) \cdot W_t$ where $I(B) = \Phi(B) = G^{-1}(B) = I_0 - \sum_{k=1}^{\infty} I_k B^k$, $|B| < 1$ $(I_0 = 1)$, we can obtain $a_t = W_t - \sum_{k=1}^{\infty} I_k W_{t-k}$. It can be shown from the function that $I_k, \;\; k = 0, 1, 2, \cdots$ is the inverse function for $AR(p)$.

From the stationary sequence $\{W_t\}$, which $W_k, \quad W_{k-1}, \quad W_{k-2}, \cdots$ was known, we can have $\hat{W}_j = W_j (j \leq k), \;\; \hat{a}_{k+l} = 0 (l \geq 1)$. In doing so, some signal statistics forecast formulas based on the auto regressive model will be obtained below.

$$\hat{W}_k(l) = \varphi_1 \hat{W}_k(l-1) + \varphi_2 \hat{W}_k(l-2) + \ldots + \varphi_p \hat{W}_k(l-p)\,, l > q \tag{7}$$

$$\hat{W}_k(l) = \sum_{j=1}^{\infty} I_j^{(l)} . W_{k+1-j}\,, l \geq 1 \tag{8}$$

where $\begin{cases} I_j^{(1)} = I_j & j \geq 1 \\ I_j^{(1)} = I_{j+l-1} + \sum_{m=1}^{l-1} I_m I_J^{(l-m)} & j \geq 1, \;\; l \geq 2 \end{cases}$

Since the primary user emergence reasoning model shares the same training samples and procedures with the auto regressive model, it is possible to combine the training samples selected from both models, and improve the forecasting accuracy with near vectors in both temporal and feature space. Thus we propose an auto regressive

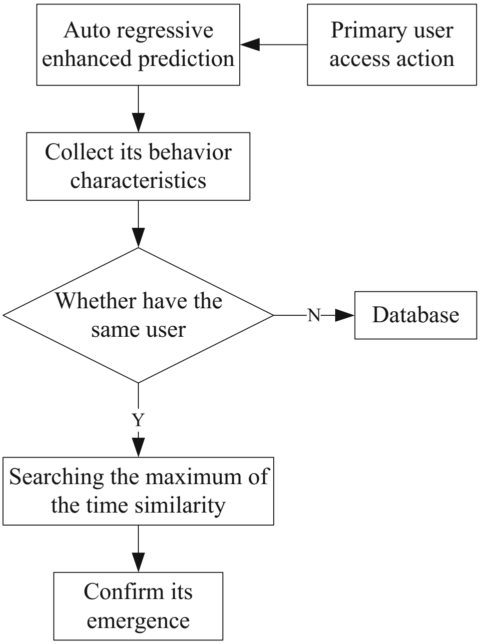

Fig. 1. Auto regressive enhanced primary user emergence reasoning model

enhanced primary user emergence reasoning (AR-PUER) model to enhance the precision of primary user presence without requiring the information about the channel. The spectrum access behavior reasoning model of primary users is shown in Fig. 1.

For the AR-PUER model training with historical data, the procedures of the combined AR and PUER model are as follows:

First of all, searching all the users in the stored access behavior, which have the similar spectrum access time with the real primary user, and then utilize the auto regressive model to enhance the prediction of primary users. Moreover, collect its access behavior characteristics and analysis to determine the primary users which in accordance with the daily access habits according to the time similarity. Once again, chase down its nearest neighbor by means of the above mentioned primary user, and then calculate the time span of them. Furthermore, update the searched user in the light of its access behavior and correction rule. Eventually, make a review and revision, thus evaluate the necessity whether hold it back.

As to the situation that primary users may have a delay during the spectrum access procedure because of some special circumstance, which will result in $\tau(\alpha, \beta)$ is invalid. Accordingly, we introduce the conception of behavior similarity. Frequently, the primary user's construction set is consist of feature set and relation set, which can indicated as $U : U = \{Feature, \mathrm{Re}lation\}$. Besides, the configuration set U can also include some other user's properties, for example, with the primary user's different access behavior, it may contain the weight coefficient W, that is to say, $U : U = \{Feature, \mathrm{Re}lation, Weight\}$.

Supposing that $User1$ and $User2$ match all the characteristics, then we have $U_1 = U_2$. If they only have some characteristic values in common, they are partial similarity, and we can utilize Sim to express the similarity between them. It can also be defined as the ratio between the matched characteristics with all the features, that is to say $Sim \in [0, 1]$. The larger the relation value, the higher the similarity between them. When Sim = 1, they are the same user; while different users when Sim = 0. The assemblage of any two users can be defined as $V_A = \{a_1, a_2, \cdots, a_n\}$, $V_B = \{b_1, b_2, \cdots, b_n\}$, and then the similarity between them can be expressed as

$$Sim(A, B) = \frac{1}{n}\sum_{i=1}^{n} \sin(a_i, b_i) \tag{9}$$

If the user's access behavior is not stable, we can set the corresponding weight. Consequently, the expression of behavioral similarity, which include the weight coefficient, can represented as

$$Sim(A, B) \frac{\sum_{i=1}^{n} \sin(a_i, b_i) w_i}{\sum_{i=1}^{n} w_i} \tag{10}$$

where $\sin(a_i, b_i) = \begin{cases} 1, a_i = b_i \\ 0, a_i \neq b_i \end{cases}$, and α_i stands for the accessed primary user, while β_i express the stored primary user access action. At the moment of $a_i = b_i$, the forecast result is true, moreover, $\sin(a_i, b_i) = 1$; conversely, $\sin(a_i, b_i) = 0$. Simultaneously, we can define the weight coefficient w_i flexibility on the basis of the actual usage condition.

3 Simulations and Analysis

Simulation experiments based on matlab platform are made to check the detection performance in this section. Since the proposed model is a combination of AR model and PUER model, each of these models are also separately utilized. The training window size is mainly set to catch more similar pattern in the historical data. Since the combined AR-PUER model is used for primary user presence forecasting, the training window can be more time-correlated by setting two months of primary user data for training.

The performance of spectrum sensing is characterized by the probability of detection and the probability of false alarm. In the case, the secondary users may cause a severe interference to the primary user. On the contrary, the probability of false alarm means the probability that the secondary user judges primary user to be detected even though the primary user does not occur. Besides, the consumption of radio resource becomes worse because the secondary user changes the currently used channel into another available channel. It has been recognized that there is a tradeoff between the probability of detection and the probability of false alarm according to the sensing threshold [8, 9].

Taking into account the condition that spectrum access of primary users at the weekend and weekday may be different; however, the overall usage of the primary users is similar. Therefore, the experiments only consider the working days, and exploit different numbers to represent various primary users. Subsequently, the simulation is carried out on the trial platform to analyze whether the AR-PUER model is effective. Through the experiment, we can get the statistics data of primary users' presence time and the model successful forecasting number of times. As shown in Fig. 2, where the whole represents the total spectrum access number of primary users, while the success means that the success predictive number for secondary user through the AR-PUER model.

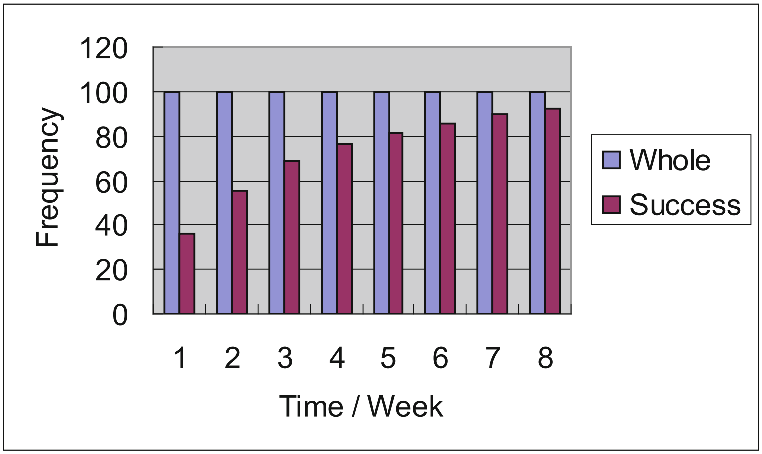

Fig. 2. Service chart of AR-PUER model

For the secondary user, the dynamic spectrum access successful implementation depends largely on the model of AR-PUER detection algorithm and the intelligent degree of the model mainly reflects in how many times it can successfully predict the primary users' emergence. That is to say, the more service time provided by the AR-PUER model is just the number of services that the secondary user would like to get, the higher successful forecasting rate it is. Additionally, the AR-PUER model's intelligence level is also improved and vice versa. The detection performance comparison with energy detector scheme in AWGN channel as shown in Fig. 3.

Primary users in the daily use of the spectrum often occurs in the following situations, such as one day occupy multiple spectrum bands, the same primary user during different time periods of the spectrum is not the same, some spectrum resources even have not been used. Even though there are circumstances exist as above-mentioned, AR-PUER model is still more effective in providing the spectrum access prediction services for secondary users. After analysis and calculation, the success rate of AR-PUER model during the past three months are improved dramatically at the initial time, after that it tends to smooth stably, this can be seen from the data in Fig. 4.

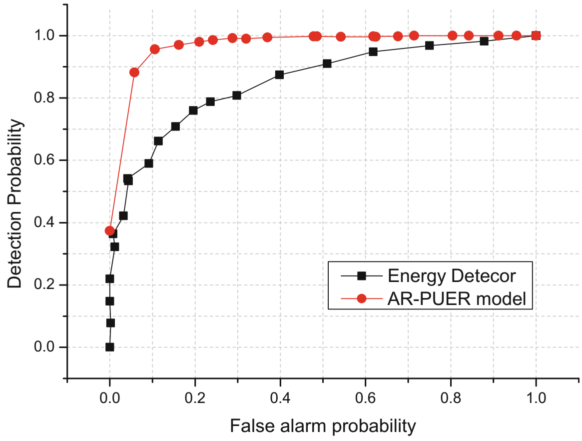

Fig. 3. Comparison with energy detector scheme in AWGN channel

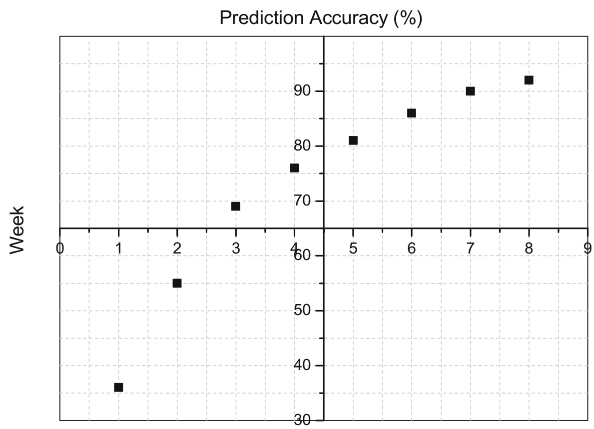

Fig. 4. Prediction trend chart of AR-PUER model

Through the data of first four weeks, it can be seen that the successful rate of the AR-PUER model to forecast the presence of primary users began to rise as time goes by, this reflect that the learning ability of the algorithm increases gradually with time. The following four weeks data show that through a long time of learning, the AR-PUER model can provide stable and high quality service for secondary users, which also demonstrate that the AR-PUER model is effective in the implementation of dynamic spectrum access in secondary users. This process can also be illustrated as a 3D color pie chart in Fig. 5.

The numerical experiment illustrated that the forecasting performance was improved by combining the training samples of the AR and PUER model, and the improvement was repeatable for different primary user data. Accordingly, the proposed

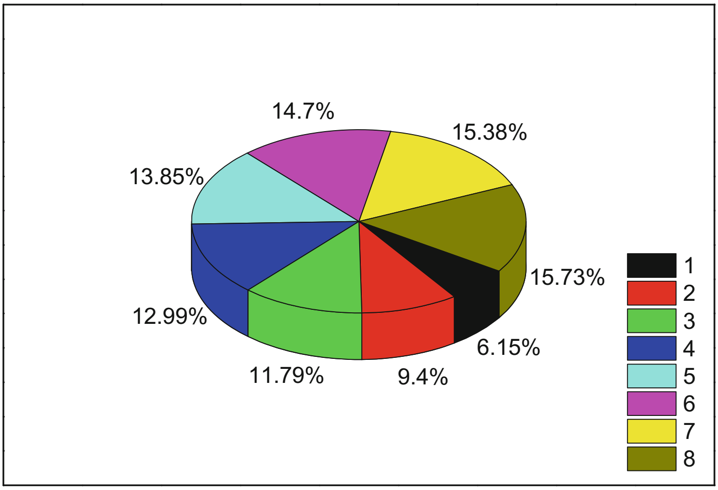

Fig. 5. 3D color pie chart of the prediction outcome

AR-PUER model could improve the probability of detection in the extant radio environment. Summarizing the results of the numerical experiments, the proposed model achieves the best forecasting performance.

4 Conclusion

This paper has introduced a blind auto regressive enhanced primary user emergence reasoning (AR-PUER) model for the cognitive radio. The method is based on linear prediction and case based reasoning. By using a combination of auto regressive prediction and primary user emergence reasoning, time similarity and time freeness are computed in our method. This is useful since it is desirable for a sensing model to operate without requiring the knowledge of the noise statistics. The time similarity is, thus, an indicator of the presence of the primary users in the signal that is received by the secondary user.

The analysis results show that the performance of the proposed model is better than the conventional manners for the detection of primary users. Besides, the simulation outcomes demonstrate that the AR-PUER model is effective and generates the most accurate forecasting of primary users occurrence in several cases.

Acknowledgement. This work was supported in part by the National Natural Science Foundation of China under Grant Nos. 61143008, 61471066, National High Technology Research and Development Program of China under Grant No. 2011AA01A204, and the Fundamental Research Funds for the Central Universities.

References

1. Kolodzy, P.: Spectrum policy task force. Fed. Commun. Comm. Technol. Rep. Docket **40**(4), 147–158 (2002)
2. Jajszczyk, A.: Cognitive Wireless Communication Networks (Hossian, E., Bhargava, V.) [Book Review]. IEEE Communications Magazine, p. 18, November 2008
3. Gardner, W.A.: Signal interception: performance advantages of cyclicfeature detectors. IEEE Trans. Commun. **40**(1), 149–159 (1992)
4. Sonnenschien, A., Fishman, P.M.: Radiometric detection of spread-spectrum signals in noise of uncertain power. IEEE Trans. Aerosp. Electron. Syst. **28**(3), 654–660 (1992)
5. Juell, P., Paulson, P.: Using reinforcement learning for similarity assessment in case-based systems. IEEE Intell. Syst. **18**(4), 60–67 (2003)
6. Jeng, B.C., Liang, T.P.: Fuzzy indexing and retrieval in case-based systems. Expert Syst. Appl. **88**(1), 135–142 (1995)
7. Golub, G.H., Van Loan, C.F.: Matrix Computations. Johns Hopkins University Press, Baltimore (1983)
8. Digham, F.F., Alouini, M.-S., Simon, M.K.: On the energy detection of unknown signals over fading channels. IEEE Transactions on Commun. **55**(1), 21–24 (2007)
9. Liang, Y.-C., Zeng, Y., Peh, E., Hoang, A.T.: Sensing-throughput tradeoff for cognitive radio networks. In: IEEE International Conference on Communications, pp. 5330–5335, June 2007