Toward A Socially Optimal Wireless Spectrum Management

Zhen Li
Department of Economics and Management
Albion College, USA.
Email: zli@albion.edu

Qi Liao and Aaron Striegel
Department of Computer Science and Engineering
University of Notre Dame, USA.
Email: {qliao,striegel}@nd.edu

Abstract—It is widely recognized that the wireless spectrum is a scarce and limited resource and that the present practice of static spectrum allocation and exclusive licensing is inefficient. The proposed approaches generally either focus on maximization of spectrum utilization or profit maximization of individuals (such as the government or users). In this paper, we consider an efficient, or socially optimal, spectrum sharing that consists of three objectives: full (quantitative) utilization, effective (qualitative) utilization, and zero interference. Through a comparative study of these models using suggested objective criteria, we show a hybrid model consisting of a dynamic spectrum market and dynamic spectrum access supported by cognitive radio technologies that can achieve the social optimum. The dynamic spectrum market enabled by a benevolent social coordinator has fundamental differences from existing dynamic market models in that primary licensed user is not involved in the process of allocating underused spectrum. Moreover, the motivation of social coordinator is to reach socially optimal allocation of spectrum resources rather than to maximize profit or revenue of individuals.

I. INTRODUCTION

Radio spectrum is limited and scarce in the provision of wireless telecommunications services. Historically, the spectrum allocation has been strictly regulated because of externalities in spectrum usage such as electromagnetic interference. The practice is centralized, static, and wholesale type such as the auction, in which governments assign exclusive rights to transmit signals over specific spectrum, and this practice is considered inefficient as it leads to under-utilization of spectrum resources [1].

The under-utilization of spectrum has stimulated the engineering, economics, and regulatory communities in searching for better spectrum management policies and techniques. Three major models have been developed to complement or to replace the current auction model of the Federal Communications Commission (FCC): the dynamic spectrum market model [2]–[6], the cognitive radio (CR) model [7], [8], and the spectrum commons model [9], [10]. The dynamic spectrum market model requires the government assigning property rights to license holders who can resell unused spectrum while the cognitive radio model allows licensed-exempt use by secondary users of frequency owned by a licensed primary user. The spectrum commons model employs open sharing among peer users with an equal right of access as the basis for managing a spectral region. However, it is critical to note these models focus either on the full utilization of spectrum or the profit maximization of governments or primary users, but not the efficient use of spectrum resources from an economic perspective. The dynamic spectrum market model will improve spectrum utilization but is hard to reach full utilization of spectrum due to the deviation of private incentives from social incentives when primary users actively participate in dynamic spectrum markets [11]. The cognitive radio model may reach full utilization but cannot guarantee the effective use of spectrum if secondary users with heterogenous valuation of spectrum usage have equal access to licensed spectrum.

Our approach focuses on a socially optimal spectrum resource management system, which satisfies a good balance of three important objectives: full (quantitative) utilization, effective (qualitative) utilization, and zero interference. Through comparing the four aforementioned models, we propose a hybrid model combining the dynamic spectrum market and the dynamic spectrum access supported by the cognitive radio technology that outperforms the existing models in reaching a good balance of the three objectives of efficiency. The hybrid model allows the secondary users to have access to unused licensed spectrum resources at positive access cost (depending on changing congestion conditions) payable to a social coordinator, who can be the government regulator such as the FCC. Primary users are excluded from the reallocation process of unused spectrum to maximize the possible supply of residual bandwidth. It is important to note that the goal of the benevolent social coordinator is to maximize social welfare of spectrum usage, not to maximize the profits of any party. The optimal spectrum access rate provided in real time by the social coordinator is the minimum cutoff price that induces only the secondary users with higher valuation of spectrum usage to actually use the resources.

While a technical solution is not the focus of this study, the contribution of the paper is both pointing the direction of a socially optimal utilization of wireless spectrum from an economics perspective and defining three objective criteria to reach the social optimum. We illustrate how a hybrid model of dynamic spectrum market and dynamic spectrum access enabled by cognitive radio technologies can actually achieve the social optimum by taking good balance of full and effective

1An allocation of scarce resources is considered efficient if the social welfare of using the resources is maximized. In this paper, “efficiency” and “social optimum” are used interchangeably.
utilization of the limited resources.

II. A COMPARATIVE STUDY OF MAJOR MODELS OF SPECTRUM MANAGEMENT

In this section, we outline the three objectives of efficient spectrum allocation to conduct a survey of the four major models of spectrum resource management: the FCC model, the dynamic market model, the cognitive radio model, and the commons model. The three comparison objectives used in the following analysis are:

- **Full (quantitative) utilization**: Utilization maximization, i.e., the demand for spectrum resources is satisfied to the maximum.
- **Effective (qualitative) utilization**: The spectrum resources are only allocated to those users who value and benefit the most from the spectrum usage.
- **Zero interference**: No overuse so that users do not interfere with each other.

A. License Auctioning: the FCC Model

In the United States, the FCC has been using spectrum licenses to allot spectrums to applicants. A licensed regime provides the certainty needed to ensure broad investment in the band as can be provided by exclusive licensed use. In a well-designed auction, everyone has an equal opportunity to win and the spectrum is sold to bidders who value it the most, hence likely to use it most effectively. Nevertheless, full utilization will be satisfied only if the bandwidth demand by the primary user is greater than or equal to the bandwidth supply. As it has been widely shown, licensed spectrums are often unused or under-utilized, resulting in white spaces and significant waste of spectrum resources [1].

B. The Dynamic Spectrum Market Model

The under-utilization of spectrum has stimulated a large bed of literature exploring the issue of dynamic spectrum sharing and management [2]–[6]. The dynamic spectrum market model we refer to is a combination of spectrum property rights (with exclusive-use) and hierarchical spectrum markets, i.e., the spectrum bands license holders have the rights to resell part of their unused spectrum to secondary users for profit. A hierarchical access structure can be established to coordinate primary and secondary users, thus limiting the interference perceived by primary users. Since such sharing is not mandated by the regulation policy, economy and market will play an important role in driving toward the most profitable (and hence effective) use of spectrum resources.

While well designed, dynamic spectrum markets will create incentives for license owners to share spectrum, such markets are unlikely to eliminate under-utilization from the root because transaction costs of spectrum buyers (secondary users) and sellers (primary users) can be significant and private incentives of license holders may deviate from social incentives [11]. To fully utilize spectrum, flexible short-term secondary licenses are needed on infinitely small slots in terms of the amount of spectrum, the time windows and the area coverage. It is cumbersome for license holders to fully identify the reusability of the spectrum in a very fine granularity. The delay in negotiating and finalizing contracts in auction market can also be problematic when both buyers and sellers are self-interested. A dynamic spectrum market will only arise if the transaction cost of license holders is less than the value of the spectrum to secondary users net of the transaction cost of secondary users. In addition, license holders may not always want to supply idle spectrum to the market as their business profit-maximization decision-making is not necessarily consistent with revenue maximization from selling/renting excess spectrum, due to competition concerns for example.

C. The Cognitive Radio Model

Cognitive radios and opportunistic spectrum access [7], [8] seek technical solutions to the under-utilization problem. While cognitive radio users are capable of accessing both the licensed and the unlicensed spectrums [12], the cognitive radio model we refer to is a licensed system plus non-interfering open access by unlicensed users, which is in line with the common understanding of what cognitive radio techniques shall enable. Cognitive radios create increased efficiency by dynamically allocating spectrum. It differs from the dynamic spectrum market model in that the access is open to any non-interfering usage rather than a limited number of secondary users to which license holders sell for profit, thus can perform better in fully utilizing available spectrum resources.

If cognitive radio technology is neutral, such a regime cannot guarantee the most effective use of spectrum resources in a heterogeneous user environment since secondary users have an equal access to the unused spectrum. In cases when users with less valuation for spectrum usage were selected, efficiency would not be achieved as the resources were not used in the most productive way. In QoS-aware cognitive radio networks, the objective is generally to balance the QoS of admitted secondary users and the tolerable interference from the secondary users to primary users by using some admission control algorithm, which may not necessarily lead to optimal usage of spectrum. For example, secondary users can be selected to maximize revenue of private parties such as base station operators [13], but it cannot guarantee the right set of secondary users are selected because the secondary users may not be truth-telling in requesting for accessing the spectrum. Bidding without truthfulness (or strategy-proofness) is extremely vulnerable to market manipulation and produces very poor outcomes, as already shown by economic theories and concrete examples [14].

D. The Spectrum Commons Model

The spectrum commons model gives users license-exempt access to spectrum, which is open to all and free from either government or private control [9]. The commons model challenges the exclusive use of spectrum by claiming that new spectrum sharing (such as cognitive radio) technologies allow a virtually unlimited number of users to use the same spectrum without causing each other interference. The commons model
TABLE I
COMPARISON OF SPECTRUM RESOURCE MANAGEMENT MODELS

<table>
<thead>
<tr>
<th>Model</th>
<th>Full (Quantitative) Utilization</th>
<th>Effective (Qualitative) Utilization</th>
<th>Zero Interference</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCC Model</td>
<td>highly inferior</td>
<td>primary user only</td>
<td>Yes</td>
</tr>
<tr>
<td>Dynamic Spectrum Market Model</td>
<td>inferior</td>
<td>primary user + ranked secondary users</td>
<td>Yes</td>
</tr>
<tr>
<td>Cognitive Radio Model</td>
<td>Yes</td>
<td>primary user + unranked/ranked secondary users</td>
<td>Yes</td>
</tr>
<tr>
<td>Spectrum Commons Model</td>
<td>Yes</td>
<td>unranked users</td>
<td>Yes (if used with CR)</td>
</tr>
</tbody>
</table>

is not an alternative to command-and-control regulation, but in fact shares many of the same inefficiencies of that system as a commons must be controlled either by private actors or by the government [10]. In addition to the resource over-usage problem characterized by the “tragedy of the commons”, this extreme commons model can be inefficient by itself. The commons model cannot guarantee the effective use of limited spectrum resources: when all potential users of the same spectrum have an equal access, the spectrum may be actually used by users who value the spectrum less.

E. The Comparison of the Four Models

As discussed above, none of the four models is optimal characterized by the three primary objectives. The FCC’s auctioning of exclusive licenses will avoid interference, assure a high quality of service, and foster investment in the band, but not every channel in every band is fully utilized. Market-based dynamic access has the potential to increase spectrum utilization, but it can be costly and suffers from misaligned incentives as primary users may be unable or unwilling to resell unused spectrum. Cognitive radio technology enables license-exempt use of frequency owned by a licensed party but it cannot guarantee the most effective use of spectrum resources. Managing spectrum as a commons can satisfy the full utilization and zero interference objectives (if used with cognitive radio technology) but it cannot guarantee the effective utilization of the spectrum. Table I summarizes the above four models regarding satisfying the three objectives of efficient allocation of spectrum resources. It is not difficult to see the tradeoff is between full utilization and effective utilization. The dynamic spectrum market model is superior to the FCC model regarding full utilization, but it is less competitive to the cognitive radio model on this regard; the cognitive radio model is superior to the spectrum commons model regarding effective utilization, but it is less competitive than the dynamic spectrum market model on this regard. Intuitively, an improved model can be a hybrid of the dynamic spectrum market model (for effective utilization) and the cognitive radio model (for full utilization).

III. MODELING ANALYSIS: SOCIAL OPTIMUM OF SPECTRUM ALLOCATION

While many interests have arisen in dynamic spectrum sharing with market forces, the objectives primarily are either to maximize the profit or revenue of license holders [3], to maximize the primary user’s utility [6], to maximize the profit of all secondary users [5], or to maximize auctioneer’s revenue [4]. None of the objectives is necessarily consistent with social optimum by maximizing the net benefits of spectrum allocation in society as a whole. Instead, we formulate an optimization problem of a social coordinator (e.g., government regulators such as the FCC) to maximize the social benefits of spectrum usage based on the three primary criteria for efficient spectrum allocation discussed in the previous section.

A. System Overview

We consider a license auctioning system plus open access with a varying access price based on instantaneous changes in demand and supply of residual spectrum (Figure 1), in which a primary user (user 1) and multiple secondary users (user 2 to n) who want to share the licensed spectrum auctioned to the primary user. The secondary users send the spectrum request (S_i) and the maximum access rate (P_i) they are willing to pay to the social coordinator. The social coordinator determines the optimal spectrum access rate per unit bandwidth (P^*) based on the availability of the unused spectrum by the primary user and the total requests for spectrum sharing by the secondary users. After allocation, the secondary users transmit in the allocated spectrum using adaptive modulation whose transmission rate can be dynamically adjusted based on the channel quality to enhance the transmission performance. It is assumed that all secondary users can reach social coordinators, which can be co-located with the base stations, and transmission error is within a reasonable threshold for smooth bid collection and data exchange.

The hybrid framework consists of two parts: dynamic spectrum market and dynamic spectrum access. While the dynamic spectrum access component allows open-access enabled by cognitive radio technologies, it is important to observe that in order to select the most effective secondary users, a cost of access has to be implemented so that market forces can work to reveal secondary users’ private valuation of spectrum usage. This cost is determined by the dynamic spectrum market. Notably, there are four major differences with this dynamic spectrum market component compared with other dynamic market models:

- **Benevolent Motivation**: The motivation is totally different: the social coordinator is benevolent whose motivation is to reach socially optimal allocation of spectrum resources rather than to maximize profit or revenue of individuals.

^2The aggregate benefits of spectrum usage by the society including all primary and second users
**No Primary User:** Primary users are not involved in the process of allocating unused or underused spectrum to secondary users since the primary user is guaranteed the prioritized power to get access to the licensed spectrum. In existing dynamic spectrum market models, the auction outcomes resulting from profit maximization for the primary users are not necessarily maximizing social welfare. For example, a profit-driven primary user may not offer available spectrum for auction if doing so is not profitable due to transaction costs or competition concerns. Economic incentives by primary users can deviate from social incentives, leading to unexploited opportunities when the equilibrium price is set above the zero-vacancy zero-interference optimal rate.

**Finer Granularity of Spectrum Supply:** Since any unused spectrum sensed by CR will be automatically on the market, not decided by the primary user\(^3\), the maximum possible supply of licensed spectrum to the secondary market can be reached that helps reduce price and leads to more utilization. The granularity of spectrum supply can be infinitely small, which is impossible to decide in advance as in traditional secondary spectrum auction market.

**Minimum Cost:** Rather than using the actual bid price, the final cost for accessing spectrum would be the minimum cutoff price that is sufficient to select all the secondary users with higher valuations to maximize spectrum utilization. The non-discriminating unit price is identical for all admitted secondary users.

\(^3\)The supply of residual spectrum still depends on the primary user’s actual use of the spectrum.

**B. Formulation of Optimization Problem**

We formulate the optimization problem for the social coordinator and illustrate how such optimal allocation of spectrum can be achieved via the optimal equilibrium price that can ensure both truthfulness of secondary users and efficient usage of spectrum bandwidth. Let the instantaneous occupancy of the spectrum of the primary user (user 1) be \( S_1 \leq S \) where \( S \) is the total amount of spectrum bandwidth. \( R = S - S_1 \) is thus the amount of available bandwidth that can be shared by the secondary users, called the supply of residual bandwidth that varies over time depending on the actual usage of the spectrum by the primary user. The strategy of each of the secondary users is to decide the amount of bandwidth \((S_i)\) to request and which access price \((P_i)\) to bid. Therefore, \( S = \sum_{i=2}^{n} S_i \) is the total demand for the residual bandwidth by the secondary users. The social coordinator sets an access rate per unit of bandwidth according to the unit-price function \( p(S, R, P) \) that is positive and non-decreasing for \( S > 0 \):

\[
p(S, R, P) = \begin{cases} 
\frac{\partial p}{\partial S} & \text{if } \frac{\partial p}{\partial S} \geq 0 \\
\frac{\partial p}{\partial R} & \text{if } \frac{\partial p}{\partial R} \leq 0 
\end{cases}
\]

Therefore, \( \sum_{i=2}^{n} S_i \) is the total demand for the residual bandwidth by the secondary users.

Once the access rate is determined, the social coordinator sets an access rate per unit of bandwidth according to the unit-price function \( p(S, R, P) \) that is positive and non-decreasing for \( S > 0 \):

\[
p(S, R, P) = \frac{\partial p}{\partial S} \frac{\partial p}{\partial R} S \]

The valuation of the spectrum by the secondary user \( i \) depends on his/her utility surplus for using the spectrum, denoted as \( \pi_i \). The revenue is \( r_i \cdot S_i \) where \( r_i \) is the user \( i \)'s per-unit-spectrum revenue which is positively related to the user’s spectrum usage efficiency. The cost of spectrum allocation for the user \( i \) is \( p(S, R, P) \cdot S_i \). The profit of the user \( i \) is therefore

\[
\pi_i = (r_i - p(S, R, P)) \cdot S_i
\]
The demand for bandwidth by the secondary users varies with the secondary users’ needs as a result of the secondary users’ profit-maximization decision-making. The supply of the residual bandwidth varies with the primary user’s instantaneous use of the spectrum. The social coordinator’s goal is to choose the optimal access rate so that the residual bandwidth can be optimally shared by the secondary users.

The social coordinator’s optimization problem is formulated in the following objective function (Equation 3) to set an access rate \( p(S, R, P) \) to reach full and effective utilization of the residual bandwidth:

\[
\min: |R - \sum_{i=2}^{n} b_i| \quad \text{and} \quad \max: \sum_{i=2}^{n} \pi_i \quad (3)
\]

where \( b_i \) is the actual spectrum usage by user \( i \), and

\[
b_i = \begin{cases} 
S_i, & \text{if } p(S, R, P) \leq r_i, \\
0, & \text{if } p(S, R, P) > r_i.
\end{cases} \quad (4)
\]

The direct outcome of the optimization problem is the optimal equilibrium price \( p^*(S, R, P) \), which is by nature a minimum, cutoff price that guarantees the secondary users with higher valuation of the spectrum are actually using the spectrum. The admitted secondary users are not price-discriminated in the sense that they are all charged the same unit price for actually accessing the spectrum. The final price is not the actual bid prices from secondary users, instead, the lowest cutoff price is chosen from the set of bid prices \( P \) that guarantees demand equals supply \( (S = R) \). If the secondary users are ranked as \( r_2 \geq r_3 \geq \ldots \geq r_x \geq \ldots r_{n-1} \geq r_n \), the optimal access rate would be \( p^*(S, R, P) = r_x \) such that \( \sum_{i=2}^{x} b_i = R \). To avoid interference, the demand by the cutoff user \( x \) may only be partially satisfied.\(^4\) In the case of \( S \leq R \), \( p^*(S, R, P) = r_n \), the lowest bid by the secondary users.

For minimizing the delay and data exchange, the dynamic spectrum requests are likely to adopt a one-round simultaneous bid model in order to achieve in real time. Users will not be afraid to reveal their true valuation for the spectrum usage since the final price will not be their bid price but the actual minimum cutoff price. On the other hand, users are also lack of incentives to overbid due to the upper bound revenue constraint specified in Equation 2. The truthfulness of users is therefore guaranteed by this non-discriminating minimum cutoff price charged on all admitted secondary users.

C. Economic Evaluation of Social Optimum of Spectrum Allocation

The social optimum of wireless spectrum utilization can be illustrated and justified in Figure 2. The left vertical axis is the private valuation of per-unit spectrum by potential users. The \( n \) users are ranked by their true valuations of per-unit spectrum usage from the highest to the lowest. When potential users are ranked according to their individual valuations of some spectrum, we can derive the private value curve (analogous to the demand curve) as the descending bars illustrated in Figure 2. The right vertical axis is the cumulative percentage of the total valuation as shown by the cumulative valuation curve provided by the Pareto chart. The horizontal axis represents the spectrum size. For illustration purposes all bars are of an equal width meaning all users demand for an equal share of spectrum resources.

The social coordinator in the proposed optimal model would allocate the spectrum usage to the first \( x \) users when the sum of the demand for the spectrum by these \( x \) users is equal to the fixed supply of the spectrum, i.e., \( \sum_{i=1}^{x} S_i = S \). Note the demand of user \( x \) can only be partially satisfied if the remaining spectrum size after serving the first \( (x-1) \) users is less than \( S_x \), which is the demand for the spectrum by user \( x \).

Such an allocation of spectrum resources is efficient as the social welfare of the spectrum usage can be maximized. Thus, the highest level of effectiveness the spectrum allocation can reach is \( E_{optimum} \), representing the maximum percentage of social valuation that can be realized. If free markets could be well developed to coordinate the demand and supply of spectrum resources, market forces would drive the market price to the efficient level \( p^* \) such that only the first \( x \) users became the actual users of the spectrum, and the efficient usage of the spectrum would be realized. Therefore, the optimal regime balances all the three important but often conflicting objectives of spectrum management: full utilization, effective utilization and no interference, making the hybrid regime superior as it takes care of both spectrum utilization maximization and spectrum effectiveness maximization.

\(^4\)To avoid partial service, the price can be easily adjusted to equal \( r_{(x-1)} \).
As a hypothetical case study to outline one possible scenario of the hybrid model, suppose a wireless carrier $V$ wins Block $B$ of $X$ MHz FCC auction. User $V$ is thus the primary user who has the prioritized right of accessing the block. As a precondition of the auction, the FCC rules are to allow dynamic access to the unused block without $V$’s permission. The FCC, acting as benevolent social coordinator, can use fine-designed standardized (time-region-block) spectrum packages to make an instantaneous market for trading residual Block $R$, analogous to real-time stock exchanges. The FCC acts as the seller of unused spectrum, enabling the maximum possible supply of spectrum bandwidth that leads to full utilization.

Specifically, considering at time $t$, all secondary users in the region who need to access spectrum send bandwidth requests together with their truthful bid prices for a standard package to FCC trading market co-located with the nearest wireless base station. The FCC chooses the optimal unit price and sends it to the secondary users. For example, if there are two packages available and the top two highest bids are $P_i$ and $P_j$ ($P_i < P_j$) sent by user $i$ and $j$, then the price $P_i$ would be the cutoff price, paid by all the secondary users with a bid no less than $P_i$ (i.e., user $j$ who bid $P_j$ only needs to pay $P_i$). In case that the total number of packages requested by the secondary users is no higher than the number of packages available, the optimal price is set at the lowest bid price to ensure a positive cost. The actual trade price is the minimum cutoff price rather than the bid price. By this way, it can not only induce users to be honest in revealing their true evaluation of spectrum, but also reduces the set of information the secondary users have to deal with.

The packages can have a term of early termination: a secondary user’s use of the spectrum can be automatically terminated to avoid interference if the primary owner requests access. Suppose at time $(t+1)$ the cognitive radio of the user $i$ senses signals sent by the primary user $V$, the user $i$ stops using the spectrum and sends a new bandwidth request to the trade market which in turn assigns a different spectrum at a new price in real time. To improve effectiveness, it is the lower-bidding secondary users whose services are terminated first. For example, if the user $k$ bids less than the user $i$, the user $k$ will be terminated and the bandwidth be reallocated to the user $i$. The payment by the secondary users for using the residual spectrum can be made at the end of each time period, based on the actual use of spectrum to accommodate the possibility of early termination.

Lastly, we note that this paper does not target at solving technical challenges but rather focusing on the social optimum and economic perspective. We focus on the interesting idea of bringing together both the dynamic spectrum market and dynamic spectrum access supported by cognitive radio technologies and prove the concept that an efficient, socially optimal spectrum management is possible based on the three important evaluation criteria, and consider the technical support of cognitive radios to be beyond the scope of the paper.

Wireless spectrum resources are limited, making how to efficiently utilize these limited resources an important topic. Cognitive radio technologies make an innovative step toward more utilization of spectrum. However, current models such as FCC auctions, cognitive radios, dynamic markets, or spectrum as commons alone cannot achieve an optimal solution. It is challenging to design a spectrum allocation regime that could satisfy all the three objectives, i.e., full utilization, effective utilization and zero interference at the same time. We argue the importance of social optimum from an economic perspective and make an initial attempt proving that a hybrid of dynamic spectrum market and dynamic spectrum access supported by cognitive radios can satisfy a good balance of the three objectives and can achieve the social optimum of wireless spectrum allocation. Dynamic sharing of spectrum is still in its infancy. Many complex issues in technical, economic, and regulatory aspects need to be addressed before its potential can be assessed and realized. It is our hope that this study will add to the important field of investigation that helps us understand how to effectively utilize wireless spectrum from a socially optimal perspective.

REFERENCES


