Sequential Multilateral Search for a Common Goal

Igor Rochlin, David Sarne
Department of Computer Science
Bar-Ilan University
Ramat-Gan, Israel

Gil Zussman
Department of Electrical Engineering
Columbia University
New York, NY

Abstract—**Motivated by applications in Dynamic Spectrum Access Networks, we focus on a system in which a few agents are engaged in a costly individual search where each agent's benefit is determined according to the minimum obtained value. Such a search pattern is applicable to many systems, including shipment and travel planning. This paper formally introduces and analyzes a sequential variant of the general model. According to that variant, only a single agent searches at any given time, and when an agent initiates its search, it has complete information about the minimum value obtained by the other agents so far. We prove that the search strategy of each agent, according to the equilibrium of the resulting Stackelberg game, is reservation-value based, and show how the reservation values can be calculated. We also analyze the agents' optimal search strategies when they are fully cooperative (i.e., when they aim to maximize the expected joint benefit). The equilibrium strategies and the expected benefit of each agent are illustrated using a synthetic homogeneous environment, thereby demonstrating the properties of this new search scheme and the benefits of cooperation.**

Keywords-**Multilateral Search, Cooperation, Dynamic spectrum access networks**

I. Introduction

This paper focuses on search problems stemming from the spectrum sensing process of users in a Cognitive Radio Network (also known as Dynamic Spectrum Access Network). A Cognitive Radio was first defined by Mitola [21] as a radio that can adapt its transmitter parameters to the environment in which it operates. According to the Federal Communications Commission (FCC), a large portion of the assigned spectrum is used only sporadically [8]. Due to their adaptability and capability to utilize the wireless spectrum opportunistically, Cognitive Radios are considered key enablers for efficient use of the spectrum [1], [9].

Under the basic model of Dynamic Spectrum Access Networks [1], Secondary Users (SUs) can use *white spaces* that are not used by the Primary Users (PUs) but must avoid interfering with active PUs.[1] In order to identify available PU channels, the SUs have to sense the spectrum and to evaluate the quality of the different available channels. In particular, a *spectrum sensing* mechanism has to determine how and when the SUs sense the different channels and a *spectrum decision* mechanism has to determine which channel best satisfies the application requirements (different channels may have different qualities) [1], [12]. While spectrum sensing is primarily a physical layer issue, we focus on the functionalities above the physical layer that determine how the sensing should be performed.
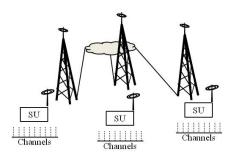

Figure 1. An illustration of the search process of the SUs which are connected via an infrastructure network.

Specifically, we focus on the search process of SUs which is illustrated in Figure 1.[2] In our model, there are a few non-interfering SUs that are served by different base stations or access points. These SUs need to maintain a connection through the infrastructure network. Each of the SUs has to sense the channels in its own environment (i.e., to its own base station) and to select a specific channel based on the channels' qualities (these qualities are not known in advance as they depend on PUs' activity). Sensing a specific channel consumes the SU's resources (e.g., energy), and therefore, an SU will usually not sense all the channels in its environment.

The key point is that since all SUs participate in the connection, the overall performance of the system (e.g., bandwidth allocated to the connection) is a function of the quality of the worst channel. Each SU's decision of whether or not to sense and evaluate an additional channel should thus be based on the tradeoff between the expected incremental improvement that can be obtained in the connection's quality and the cost (in terms of energy spending) associated with the additional sensing operation. The fact that the marginal improvement also depends on the findings of the other searcher(s) substantially complicates the calculation of the SUs' search strategy. Particularly, when the SUs are self-interested, the set of search strategies should be derived based on equilibrium analysis.

Models of agents engaged in costly search processes involving the evaluation of different available options (opportunities), out of which only one needs to be chosen, are quite common in multi-agent systems (MAS) [6], [16]. In these models, the goal of the searcher is not necessarily to find the opportunity associated with the maximum value but rather to maximize the overall benefit, defined as the value of an opportunity eventually picked minus the costs accumulated during the search process. Economic search theory provides a framework

---

[1]PUs and SUs are also referred to as Licensed and Opportunistic Users, respectively.

[2]We note that modeling the effects of the PU arrival process [15] is out of the scope of this work.

for optimizing the performance of a searcher in such costly search settings [19], [24]. The optimal search strategy in such models is commonly reservation-value based. Namely, the searcher evaluates opportunities as long as the best value found is lesser (or greater, depending on the application) than a pre-defined threshold.

Despite the richness of research in search theory, the models used commonly assume a *single* search process. Yet, as illustrated by the Dynamic Spectrum Access example, in reality agents often need to take into consideration the performance of other agents engaged in the search. Other examples of such dependencies can be found in complex transportation problems that involve ground, air and sea transportation, as well as in travel planning. In the first, assume that for each segment of the trip different offers should be received from shipment companies. The selection of a container/vehicle for each segment of the trip dictates the amount and type of cargo that can be transported overall (a single bottleneck). Similarly, when planning a trip and requesting quotes from different vendors (e.g., for flights and accommodation), the correlation between the results of different search efforts substantially influences the overall performance.

In this paper, we formally introduce the model of a multilateral search in which different agents need to search individually and the performance of each agent is affected by the performance of all the others. In particular, we focus on a protocol under which only one agent searches at a time, and each agent's search starts upon receiving the final outcome of all agents that have searched prior to it. Since the agents are searching sequentially, the problem can be considered from a game theory perspective and formulated as a Stackelberg game where each agent is a follower for the agent performing the prior search and a leader for the agent performing the subsequent search. We provide a comprehensive analysis of the problem and prove that the equilibrium strategies for the sequential multilateral search are reservation-value based. Based on the analysis, we obtain the appropriate equations from which the reservation value of each opportunity can be calculated. Complementary analysis for the case of fully cooperative agents and defection from cooperation scenarios is also provided. We use homogeneous environments (i.e., where all opportunities available to an agent share the same search cost and probabilistic properties) to illustrate the effect of different parameters on the equilibrium search strategies and the expected benefit of the different searchers.

II. THE MODEL

For exposition purposes, we first introduce and analyze the two-agents case. Then, in Section V we generalize for the k-agents case. The two-agents model considers agents A_1 and A_2, both interested in forming an ad-hoc partnership from which they can both benefit. Each agent t contributes a value x^t ($t = 1, 2$) to the partnership. The benefit of agent A_t from the partnership, denoted v_t, is the minimum of the values contributed, i.e., $v_t = min\{x^i | i = 1, 2\} = v^*$.

The value x^t, contributed by agent A_t, is the value of one out of n_t opportunities, denoted $O_t = \{o_1^t, .., o_{n_t}^t\}$, available to that agent. While the value of each opportunity o_i^t is a priori

unknown to agent A_t, the agent is acquainted with the probability distribution function, $f_i^t(x)$, from which it is derived. In order to obtain the true value of opportunity $o_i^t \in O_t$, agent A_t needs to consume some of its resources. This is modeled by the cost c_i^t, expressed in terms of opportunity values. Therefore, the agents are required to conduct a search process which takes into consideration the tradeoff between the marginal improvement in the value they contribute to the partnership and the cost incurred along the process.

There are many protocols the agents can follow for executing this multilateral search, differing in the parallelism and levels of cooperation along the process. For example, the agents can search simultaneously with no interaction between them until each terminates its search. Another example is when they take turns searching and share the values they find along the process. Each search model variant is associated with different advantages, disadvantages, and computational complexities. In this paper, we investigate a sequential search protocol by which agent A_1 performs its search first and agent A_2 follows, after obtaining the value contributed by agent A_1. Furthermore, we assume that the value of opportunities cannot be obtained simultaneously, and therefore, each agent performs its search sequentially, obtaining the values one at a time [19], [29] (e.g., spectrum sensing technology precludes the evaluation of several channels simultaneously). The value x^t, contributed by agent A_t, is the maximum among the values obtained along the individual search of that agent.

We distinguish between several settings, varying in the way each agent's expected benefit is defined. In the first setting, both agents are self-interested. Namely, each agent A_t attempts to maximize its own overall expected benefit, denoted EB_t, defined as the minimum of the values contributed, v^* minus the expected accumulated cost incurred by that agent along its search. In the second setting, both agents are cooperative. Thus, each agent's goal is to maximize the aggregate of both agents' expected benefits, i.e., $2v^*$ minus the sum of the costs accumulated along both agents' searches. The cooperative setting is common when the agents involved represent individuals from the same family or organization.

III. ANALYSIS

We first introduce the optimal search strategy for a single searcher facing a sequential search problem, without restricting the value found by the other agent. We then augment that strategy and adapt it to the case of two-agents with the minimum value restriction, distinguishing between different cooperation variants of the model.

A. Optimal Sequential Search with no Restrictions on the Value Found

When relaxing the restriction over the value obtained by the searcher, each agent's search problem can be mapped to the canonical sequential search problem described by Weitzman [29]. Weitzman's model considers a single searcher facing a setting similar to the one used for each of the agents in our model, except that the agent's expected benefit is the highest value it finds minus the expected cost incurred along its search.

The optimal search strategy in Weitzman's model is based on setting a reservation value (a threshold), denoted r_i for

each opportunity o_i. The reservation value to be used should satisfy:

$$c_i = \int_{y=r_i}^{\infty} (y - r_i) f_i(y) dy \qquad (1)$$

Intuitively, r_i is the value where the searcher is precisely indifferent: the expected marginal benefit from obtaining the value of the opportunity exactly equals the cost of obtaining that additional value. The searcher should always choose to obtain the value of the opportunity associated with the maximum reservation value and terminate the search once the maximum value obtained so far is greater than the maximum reservation value of any of the opportunities which has not yet been obtained.

B. Optimal Search Strategy for the Sequential Multilateral Search with Value Restriction

We begin by developing the optimal search strategy for agent A_2, given the value v received from agent A_1. Obviously, if A_2 obtains a value greater than v along its search then it necessarily terminates its search, as its benefit from the partnership cannot be improved further.

Theorem 1: The optimal search strategy for agent A_2, given a value v obtained from A_1, is to set a reservation value $r_i^2 < v$ for each opportunity o_i^2, where r_i^2 is derived from:

$$c_i^2 = \int_{y=r_i^2}^{\infty} (min(y, v) - r_i^2) f_i^2(y) dy \qquad (2)$$

The agent should always choose to obtain the value of the opportunity associated with the maximum reservation value and terminate the search once the maximum value obtained so far is greater than the maximum reservation value of any of the opportunities, which has not yet been obtained.

Proof: The proof augments the inductive proof given in [29] for the case where no restrictions are made on the value found. We first prove that the right-hand side of (2) is a monotonic decreasing function in r_i^2, starting from infinity for $r_i^2 \to -\infty$, thus the r_i^2 that satisfies (2) always exists. For the inductive part, we begin with the case of having a single opportunity. Here, the right-hand side of (2) can be interpreted as the expected additional gain from obtaining the minimum between the value of that opportunity and v, if the searcher is already guaranteed a value $r_i^2 < v$. Obtaining the value of the opportunity in this case is thus beneficial only if the expected additional gain is greater than c_i^2. Therefore, since for any guaranteed value which is less than r_i^2, the right hand side of (2) is necessarily less than c_i^2 and thus r_i^2 is, in fact, a reservation value in this case.

Assume that the reservation-value based strategy is optimal for the case of $n_2' < n_2$ available opportunities when the best value found so far by A_2 is y. We need to prove that for the $n_2' + 1$ available opportunities and best known value y case, the optimal strategy is also reservation-value based. We consider the opportunity o_i^2 associated with the highest reservation value among the $n_2' + 1$ opportunities. In order to obtain the value of that opportunity we distinguish between the cases where $y \geq r_i^2$ and $y < r_i^2$. If obtaining the value of o_i^2 when $y \geq r_i^2$, then we are left with n_2' opportunities whose reservation values are necessarily less than y. Therefore, according to the inductive assumption, the search should

terminate. The decision should be made solely based on the benefit of obtaining the value of o_i^2 (constrained by v) and the cost c_i^2 which, according to (2), is negative and thus the value should not be obtained. When obtaining the value of o_i^2 when $y < r_i^2$, one or more opportunity values should be obtained (there is at least one opportunity where obtaining its value is beneficial according to (2)). We need to prove that obtaining the value of o_i^2, rather than the value of any other opportunity, is optimal. This is achieved by showing that the expected benefit from obtaining the value of o_i^2 and then following the optimal strategy for the remaining n_2' opportunities according to the induction assumption, is greater than obtaining the value of any opportunity o_j^2 ($i \neq j$) and then following the optimal strategy for the remaining n_2' opportunities according to the induction assumption. ∎

From (2), we observe that the reservation value r_i^2 increases as v increases and decreases as c_i^2 increases. Also, r_i^2 is always smaller than v, otherwise the right-hand side of (2) becomes zero. Furthermore, Proposition 1 presents an important property of the correlation between r_i^2 and v.

Proposition 1: The difference between v and r_i^2 increases as r_i^2 increases, i.e., $0 < \frac{dr_i^2}{dv} < 1$.

Proof: For any two values v and v', satisfying $v < v'$, and the difference $v - r_i^2$, we prove that (2) can be satisfied only by using $r_i^{2'}$ that satisfies: $r_i^{2'} < v' - (v - r_i^2)$. ∎

Given a value v_1 and the search strategy described in Theorem 1, A_2 returns a value v_2. The benefit of both agents from v_2 is $v_2' = min(v_2, v_1)$.[3] In order to calculate $E[v_2'|v_1]$, we use $S = (v_1, w, i)$ to denote A_2's state along its search, where i is the index of the opportunity whose value is to be obtained next, v_1 is the value obtained from A_1 and w is the highest value found by A_2 thus far. The expected effective value returned eventually by A_2, if at state (v_1, w, i), is thus given by the recursive equation:

$$E[v_2'|(v_1, w, i)] = \int_{y=-\infty}^{r_i^2} E[v_2'|(v_1, max(w, y), i+1)] f_i^2(y) dy +$$

$$+ \int_{y=r_i^2}^{\infty} min(v_1, max(w, y)) f_i^2(y) dy \qquad (3)$$

where: $E[v_2'|(v_1, w, n_2 + 1)] = w$

Therefore, $E[v_2'|v_1] = E[v_2'|(v_1, -\infty, 1)]$.

Using $E[v_2'|v_1]$, agent A_1 can construct its optimal strategy, as given in Theorem 2.

Theorem 2: The optimal search strategy for agent A_1, is to set a reservation value r_i^1 for each opportunity o_i^1, where r_i^1 is derived from:

$$c_i^1 = \int_{y=r_i^1}^{\infty} (E[v_2'|y] - E[v_2'|r_i^1]) f_i^1(y) dy \qquad (4)$$

Always choose to obtain the value of the opportunity associated with the maximum reservation value and terminate the search once the maximum value obtained so far is greater than the maximum reservation value of any of the remaining opportunities.

Proof: Similar to the proof given for Theorem 1 with the appropriate modifications of expected value calculation. ∎

---

[3]In fact, for the two-agents case, $v_2' = v^*$. However, in (3)-(8) we prefer the use of v_2' as it facilitates the transition to the multi-agent case.

The expected value obtained by each agent from the partnership, $E[v_2']$, can be calculated using a recursive equation similar to (3). We represent the state of A_1 by (w, i), where w is the best value found so far by A_1 and i is the index of the next opportunity whose value it needs to obtain. We use $E[v_2'|(w, i)]$ to denote the expected effective value resulting from A_2's search, if A_1 is currently in state (w, i).

$$E[v_2'|(w, i)] = \int_{y=-\infty}^{r_i^1} E[v_2'|(max(w, y), i + 1)]f_i^1(y)dy +$$
$$+ \int_{y=r_i^1}^{\infty} E[v_2'|max(w, y)]f_i^1(y)dy \qquad (5)$$
where: $E[v_2'|(w, n_1 + 1)] = w$

Therefore $E[v_2'] = E[v_2'|(-\infty, 1)]$. Since the effective value v_2' applies to both agents, the expected benefit of each agent differs only in its accumulated cost component. The expected cost of agent A_t is given by $\sum_{i=1}^{n_t} c_i^t P_t(i)$, where $P_t(i)$ denotes the probability that A_t will eventually obtain, along its search, the value of o_i^t. The value of $P_t(i)$ is given by $\prod_{j=1}^{i-1} F_j^t(r_i^t)$. The expected benefit of any agent A_t, denoted EB_t, is thus given by:

$$EB_t = E[v_2'] - \sum_{i=1}^{n^t} c_i^t \prod_{j=1}^{i-1} F_j^t(r_i^t) \qquad (6)$$

C. Cooperative Behavior and Defection

In the previous section, we assumed that both agents A_1 and A_2 behave in a self-interested manner, i.e., each attempts to maximize its own expected benefit. We now consider the setting where both agents aim to maximize the sum of their expected benefits, i.e., $\sum_{t=1}^{2}(E[v_2'] - \sum_{i=1}^{n_t} c_i^t P_t(i))$. Naturally, the search strategies that maximize the latter are different from those used for the self-interested case. Furthermore, while the expected overall joint benefit increases when both agents are searching cooperatively, there is often an incentive for both agents to deviate from the cooperative strategy in order to improve their individual expected benefit. In the following paragraphs, we present the optimal strategies to be used in a fully cooperative setting and discuss the dynamics that occur when either of the agents (or both) defect from the cooperative strategy.

1) Fully Cooperative Setting: When both agents are fully cooperative and when obtaining the value of an additional opportunity, each agent should consider not only the marginal benefit from that value to itself, but also the benefit from the increase in the joint value v_2' to the other agent. In this case we can prove, similar to the proof used for Theorems 1 and 2, that the optimal strategy of both agents is reservation-value based, though different in its value. The reservation value r_i^2 of A_2 in this case should satisfy:

$$c_i^2 = \int_{y=r_i^2}^{\infty} 2(min(y, v) - r_i^2)f_i^2(y)dy \qquad (7)$$

and r_i^1 of A_1 should satisfy:

$$c_i^1 = \int_{y=r_i^1}^{\infty} 2(E[v_2'|y] - E[v_2'|r_i^1])f_i^1(y)dy \qquad (8)$$

Proposition 2: Both the individual accumulated costs and the expected value $E[v_2']$ at the end of the search, in the fully

cooperative case, are greater than those resulting in the self-interested case. Overall, the expected difference among the two components (i.e., sum of expected values minus accumulated costs) is greater in the cooperative case.

Proof: It is easy to see from (2) and (7) that the reservation value r_i^2 for any value v contributed by A_1 is greater in the cooperative case. Therefore agent A_2, upon receiving a value v, will necessarily search more than in the self-interested case (incurring a greater cost) and its search will result in finding a greater expected value. Consequently, agent A_1 receives greater values for each value with which it terminates its search, and according to (8), its reservation value r_i^1 necessarily increases. The increase in r_i^1 suggests a longer search process, i.e., greater search costs. Since A_2 receives higher values with increased probability, and r_i^2 increases as v increases, A_2 ends up searching more, overall, and terminates its search with a greater expected value in comparison to the self-interested case. Finally, the joint expected benefit in the cooperative case is greater simply because the agents attempt to maximize this sum directly rather than separately maximizing each of its parts. ∎

2) Incentives to Defect: The cooperative strategies are beneficial when the cooperation can be enforced, or when both agents are obligated to the same goal (e.g., working for the same user or users from the same organization). When the cooperation cannot be guaranteed, it will never hold and the agents will use different reservation values. Regardless of the strategy used by A_1, A_2 should use a reservation value $r_i^{2,\text{defect}}$ according to (2), rather than $r_i^{2,\text{cooperative}}$ according to (7), as this strategy maximizes its expected benefit for any value v received from A_1. If A_2 is using its self-interested strategy, then we should distinguish between the case where A_1 believes that A_2 is cooperative and where A_1 believes it isn't cooperative. Assuming that agent A_2 is cooperative, A_1 should use a reservation value $r_i^{1,\text{defect}}$ according to a modification of (2) that makes use of $E'[v_2'|v_1]$, which is the expected effective returned value given v_1, when substituting $r_i^2 = r_i^{2,\text{cooperative}}$ ($1 \le i \le n_2$) in (5). Finally, we consider the case where both agents decide to deviate from their cooperative behavior (while each assumes that the other is using a cooperative strategy). Here, each agent A_t uses $r_i^{t,\text{defect}}$ ($1 \le i \le n_t$).

IV. STACKELBERG EQUILIBRIUM FOR HOMOGENEOUS ENVIRONMENTS

To demonstrate the agents' search strategies and the resulting equilibrium dynamics in the sequential multilateral model, we use a synthetic setting with an infinite number of homogeneous opportunities available to each agent. The use of the homogeneous setting simplifies the analysis. It enables the illustration of the main differences between the sequential multilateral search strategies model and others the agents may use, as well as the differences between the strategies of A_1 and A_2 under different cooperation schemes. In the homogeneous setting used for this section, all opportunities available to agent A_1 share the same search cost and probability distribution function, denoted c_1 and $f_1(x)$, respectively, and those available to A_2 are associated with c_2 and $f_2(x)$. Specifically, we
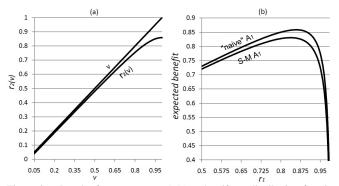
Figure 2. Results for $c_1 = c_2 = 0.01$ and uniform distribution functions over the interval $(0, 1)$: (a) A_2's reservation value as a function of the value v; (b) The expected benefit of A_1 as a function of the reservation value it uses (r_1) without constraining its values by A_2.
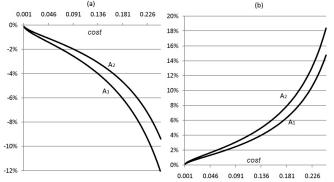

Figure 3. Differences (in percentages) between sequential multilateral search and parallel search: (a) when the agents follow the "naive" strategy simultaneously; (b) when the agents follow the equilibrium parallel strategy.

use a uniform distribution function defined over the interval $(0, m)$ (i.e., $f(x) = 1/m$, $\forall 0 \leq x \leq m$, otherwise $f(x) = 0$). We assume that the agents are not limited by a finite decision horizon (i.e., they can obtain as many values as requested). This latter assumption as well as the one regarding the infinite number of opportunities available to each agent are common in the search literature [18], [19].

Figure 2(a) depicts the reservation value that needs to be used by A_2, according to (2), as a function of the value v provided by A_1, for the case where $c_1 = c_2 = 0.01$ and both distributions are uniform over the interval $(0, 1)$. As expected, the value of r^2 increases as the value obtained, v, increases, and is always lower than v. Also, the difference between v and r^2 increases as v increases, as proven in Proposition 1. Figure 2(b) depicts the expected benefit of A_1 as a function of the reservation value it uses (r^1) in the same homogeneous environment, for both the case where the values obtained are bounded by the values returned by A_2 (denoted "S-M") and for the case where A_2 does not exist and A_1's benefit is not constrained at all (denoted "naive"). The first case describes the sequential multilateral model and the latter describes the problem of a single searcher as in [29]. As can be observed from the figure, the reservation value that maximizes the expected benefit of A_1 is different in the two cases.

The differences, in percentages, between the expected benefit of the agents when using the "naive" reservation values according to [29] and the sequential multilateral search strategy (constraining the values obtained in both cases) as a function of the search cost, are given in Figure 3(a). Here, again, the

distribution used for both agents is uniform over $(0, 1)$ and the search cost is identical for both agents ($c_1 = c_2 = c$). When applying the "naive" strategy, each agent sets its strategy according to [29], regardless of the search strategy of the other agent, thus the agents can search simultaneously. As can be observed from the graph, the use of the sequential multilateral search strategy actually worsened the agents' expected benefit compared to the use of the "naive" search strategies. While this may seem surprising, one should keep in mind that the "naive" set of strategies calculated by [29] is not in equilibrium, and each agent has an incentive to deviate from it. Figure 3(b) depicts the difference in the expected benefit of the agents when using the multilateral search strategies and when using the equilibrium strategy for the case of simultaneous searching (in percentages, as a function of the cost used, for the same environment as in Figure 3(a)). The equilibrium parallel strategies are those that guarantee that none of the agents has an incentive to deviate to using a different reservation-value strategy. As can be observed from Figure 3(b), the sequential multilateral search strategy has the potential to substantially improve both agents' expected benefit.

Furthermore, even under the permissive assumption that the "naive" reservation values are used, despite not being in equilibrium, the sequential multilateral search method has the potential to substantially improve both agents' performance, if the distribution of values associated with each agent or the costs of search are different: Figure 4(a) depicts the agents' individual and joint expected benefit, in a setting where A_1's distribution is defined over $(0, 1)$ and A_2's over $(0, 2)$, when using the sequential search analyzed in the above sections in comparison to "naive" search according to [29]. While A_2, when using the "naive" approach, suffers substantially from not considering in its search the constraint that is likely to be imposed by A_1's search (see lower curve), its expected benefit substantially increases in our model. The expected benefit of A_1 in this case does not change much, as the constraint that could be imposed by A_2's search is insignificant, due to the difference in the distribution interval. As the cost of search increases, the differences between the expected benefit of the two agents in both models increase, due to the improved efficiency achieved in the sequential search. Figure 4(b) is equivalent to Figure 4(a), except that both agents are fully cooperative. Here the joint (aggregative) expected benefit is greater, in comparison to the self-interested agents case. Since the search strategies according to the "naive" search are not influenced by the nature of the other agent (cooperative/non-cooperative), the difference between the expected joint benefit curves increases (in comparison to the non-cooperative case).

It is notable that the joint expected benefit does not necessarily improve when using the new method. For example, when using the reverse a setting, i.e., where A_1's distribution is defined over $(0, 2)$ and A_2's distribution is defined over $(0, 1)$, and search costs $c_1 = c_2 = 0.01$, the joint expected benefit is 1.11, while the use of the "naive" strategies yields 1.69. Still, the "naive" set of strategies will never hold in equilibrium.

Figure 5 illustrates the effect of the difference between the agents' search costs on their individual and joint expected benefit, in a setting where both agents' distribution functions
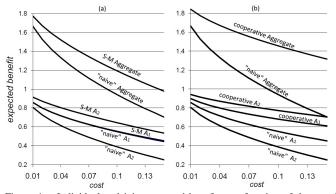
Figure 4. Individual and joint expected benefit as a function of the search cost ($c_1 = c_2$), when using the sequential multilateral model and the "naive" search strategy, where: (a) both agents are self-interested; (b) both agents are fully cooperative.
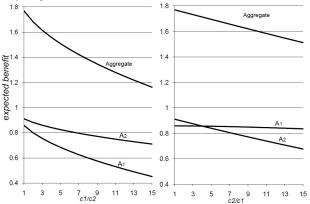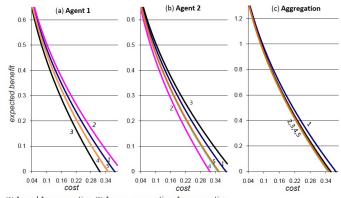

Figure 5. Expected benefit of the agents when using sequential multilateral search as a function of the costs ratios: (a) c_1/c_2 (taking $c_2 = 0.01$, fixed); and (b) c_2/c_1 (taking $c_1 = 0.01$, fixed).

(1) A_1 and A_2 cooperative. (2) A_1 non-cooperative; A_2 cooperative.
(3) A_1 cooperative; A_2 non-cooperative. (4) A_1 and A_2 non-cooperative. (5) A_1 and A_2 self-interesting.
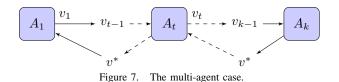
Figure 6. The case of using the cooperative strategy versus non-cooperative variations ($c_1 = c_2$): (a) Expected individual benefit for A_1; (b) Expected individual benefit for A_2; and (c) expected aggregate benefit.

are defined over $(0, 1)$. The difference between the search costs is captured by the ratios c_1/c_2 and c_2/c_1, keeping the denominator fixed. Therefore, the intersection point between parts (a) and (b) of the figure is in the value 1 over the horizontal axis. As expected, the increase in search costs is associated with a decrease in the expected benefit of both agents (as performance is also influenced by the reduction in the reservation value used by the other agent). The most interesting observation from Figure 5 is that the increase in c_1 substantially affects both agents' performance, while the increase in c_2 affects mostly A_2 and has a slight effect on A_1. This means that it is more beneficial to choose the agent associated with the greater search cost to act as A_1, if the goal is to maximize the joint expected benefit. When the two agents operate in a similar setting (i.e., when $c_1 = c_2$), agent A_2 will always benefit more than A_1 from using the sequential multilateral search. This can be easily proven by allowing A_2 to use a strategy similar to A_1's, except that when a value greater than v is obtained from A_1, A_2's search is terminated. In the latter case, A_2's accumulated cost is necessarily less than A_1's (while the value at the end of the process is similar to A_1's value).

Finally, Figure 6 describes the expected individual and joint benefit as a function of the search cost used by both agents, for the different variations of cooperation compared to the set of self-interested strategies. Here both distribution functions are

defined over $(0, 1)$. One of the curves on each graph represents the case in which both agents use the fully cooperative strategy. Two other curves represent the case in which one of the agents uses the cooperative strategy, while the other is self-interested and takes advantage of the fact that the first is being cooperative. The fourth curve represents the case where both agents are self-interested. Finally, the last curve represents the case where both agents are self-interested though each of them believes the other agent is cooperative. As expected, each agent benefits the most from acting non-cooperatively while the other agent is acting cooperatively (and vice versa, each agent suffers the most when acting cooperatively while the other acts non-cooperatively). Nevertheless, the joint expected benefit is maximized when both agents are cooperative. In the latter case, the joint expected benefit is substantially better compared to any of the other cases. The case where both agents defect from cooperation is associated with a decreased expected benefit for both agents (compared to acting cooperatively or using the self-interested strategy while assuming the other agent does the same), though it isn't as bad as when only one agent defects from cooperation.

V. EXTENSION TO MULTI-AGENT

The two-agents model can be extended to the general k-agents case, taking advantage of the analysis methodology described in the former sections. The multi-agent model considers a set of k agents $A = \{A_1, A_2, ..A_k\}$, all interested in forming an ad-hoc partnership from which they all benefit (e.g., a conference call). Each agent A_t ($t = 1, 2, .., k$) contributes a value x^t to the partnership. The common benefit of each agent from the partnership, denoted v_t, is the minimum of the values contributed, i.e., $v_t = min\{x^i | i = 1, 2, .., k\} = v^*$. We assume that the search of agent A_t starts only after agent A_{t-1} has finished its search (where A_1 is the first to search) and that the values obtained by agents that have already finished their search is common knowledge. All other assumptions used in Section II remain valid. The multi-agent settings are illustrated schematically in Figure 7.

The optimal search strategy for agent A_k is the same as the one used for A_2 in the two-agents model, i.e., given any value v_{k-1} received from agent A_{k-1}, it will extract its

Figure 7.   The multi-agent case.

reservation values according to (1). Now consider any agent A_t $(t = 2, .., k-1)$, receiving a value $v_{t-1} = min(x^{t-1}, v_{t-2})$ from agent A_{t-1}. The decision setting of agent A_t in this case, is similar to the one agent A_1 faces in the two-agents model, except that the value v_t it will transfer to A_{t+1} is bounded by v_{t-1}. Therefore, it uses the following modification of (4):

$$c_i^t = \int_{y=r_i^t}^{\infty} (E[v_{t+1}'|min(v_{t-1}, y)] - E[v_{t+1}'|r_i^t]) f_i^t(y) dy \quad (9)$$

where $E[v_{t+1}'|min(v_{t-1}, y)]$ is calculated similar to $E[v_2'|v_1]$ for the two-agents case, by substituting $v_2' = v_{t+1}'$, $v_1 = v_t$, $r_i^2 = r_i^{t+1}$ and $f_i^2(y) = f_i^{t+1}(y)$ in (3). Finally, the optimal strategy of agent A_1 is calculated similar to the strategy of A_1 in the two-agents case.

Similarly, we can also extend the cooperative model and the different defection variants, though for the latter the number of combinations of defecting agents is combinatorial.

VI. RELATED WORK

In many multi-agent environments autonomous agents may benefit from cooperating and coordinating their actions. Cooperation is mainly useful when an agent is incapable of completing a task by itself or when operating as a group can improve the overall performance [17]. Consequently, group-based cooperative behavior has been suggested in various domains [28], [30]. The recognition of the advantages encapsulated in teamwork and cooperative behaviors is the main driving force of many coalition formation models in the area of cooperative game theory and MAS [26]. Overall, the majority of cooperation and coalition formation MAS-related research tends to focus on the way coalitions are formed and consequently concerns issues such as the optimal division of agents into disjoint exhaustive coalitions [30], division of coalition payoffs [30] and enforcement methods for interaction protocols [20]. Only a few authors have considered the problem of determining the strategy of a group once formed [14], and no work to date considers search strategies for a cooperative search of the nature described in this paper.

The problem of a searcher operating in a costly environment, seeking to maximize its long-term utility, is widely addressed in classical economic search theory ([18], [19], [24] and references therein). Over the years, several attempts have been made to adopt search theory concepts for agent-based electronic trading environments associated with search costs [6], [16]. Despite the richness of search theory and its implications, most models, introduced to date, have focused on the problem of a single searcher that attempts to maximize its own expected benefit. Few studies have attempted to extend the search problem beyond a single search goal, e.g., attempting to purchase several commodities while facing imperfect information concerning prices [3], [5], [11]. Some even considered multi-agent cooperative search for multiple goals [25]. None of these works have applied any constraints on the values obtained along the search. The only constraint on the values obtained by a searcher that can be found in search theory works is the availability of recall (i.e., the ability to exploit formerly explored opportunities) [5], [19]. To date, to the best of our knowledge, a model of a multilateral search, in which one agent's search is constrained by the findings of other agents, as in the cognitive radio application, has not been introduced in the search theory research domain.

Multi-agent search can also be found in "two-sided" search models (where dual search activities are modeled) [4], [27]. The search in these models is used solely for the matching process between the different agents, i.e., for forming appropriate stable partnerships. The value of each agent from a given partnership depends on the partnership itself (e.g., the characteristics of the other agent with whom it partners). In our model the partnership is given a priori and the value of the partnership is derived from an external search process performed independently by each agent.

From the Dynamic Spectrum Access application point of view, various spectrum sensing approaches have been proposed, including a cooperative sensing scheme based on distributed detection theory [10], an adaptive MAC layer spectrum sensing [7] and a practical sensing technique that was evaluated in a testbed [23]. Several papers used game theory notions to compare the cooperative and non-cooperative behavior of spectrum sensing and sharing. In particular: [13] proposes a scheme in which users exchange "price" signals that indicate the negative effect of interference at the receivers, [2] analyzes the power control and channel selection problem as a Stackelberg game and [22] proposes a distributed approach, where devices negotiate local channel assignments aiming for a global optimum. Unlike our approach, most of the previous work in the area of Dynamic Spectrum Access and Cognitive Radio focuses on scenarios in which the SUs are in the same geographic area, sense the same set of channels and try to either agree on the same channel or on a set of non-interfering channels. To the best of our knowledge, searching for channels such that the overall performance is tied to the worst channel selected has not been studied before.

VII. DISCUSSION AND CONCLUSIONS

The sequential multilateral search model extends the legacy economic search to the case where the process involves several agents that need to engage in individual searches, and the value of each agent from the process depends on the minimum value found. As discussed throughout the introduction, such a setting arises in various real-life applications and particularly in Dynamic Spectrum Access Systems. The analysis given in this paper, proves that the optimal set of strategies to be used by the agents when using a sequential multilateral search protocol is reservation-value based. While this property aligns with a single agent's search, the equilibrium set of reservation values in the new model are different from those that ought to be used for the single agent's search model. This also implies that the sequence according to which the values of the different opportunities are obtained is often different from the one used in the sequential search. Moreover, a strategy derived according to the latter model can never be in equilibrium, as

the second agent always has an incentive to use a reservation value lower than the value it obtains. This should be taken into the first agent's considerations.

The sequential nature of the search protocol used enables the breaking down of the problem in order to find the optimal strategy of each agent as a function of the minimum value obtained by agents that have already searched and the search strategy that will be used by the other agents along the sequence. This enables the calculation of the equilibrium strategies using backward induction and a trivial extension of the analysis from the two-agent case to a general one.

The numerical illustration given in Section IV demonstrates the benefits of cooperative search behavior in the new model. Naturally, the cooperative setting is applicable whenever the agents have a joint goal (e.g., when considering family members). In other settings, this set of strategies is not stable, and, as expected, the worst expected joint benefit is obtained when each agent operates self-interestedly while believing that the other agent is cooperative. Another important observation from Section IV is the substantial effect of the order by which the sequential multilateral search process takes place over the individual and joint benefit. While this issue was left beyond the scope of the analysis, we believe that appropriate methods can be suggested for the agents to negotiate over the order by which they will search (and possibly come up with schemes for alternating orders in repeated settings) in order to improve the joint and individual expected benefit.

Additional directions for future research include the development of other multilateral search model variants, e.g., operating simultaneously (as used in a limited extent in the homogeneous-environment section), exchanging information throughout the search and even re-initiating searches by each agent based on the findings received from the other agent. Finally, applying the results to Dynamic Spectrum Access Networks will require taking into account several realistic considerations. These include the exchange of channel quality information between the SUs, the possible operation of a few interfering SUs in the same area (all searching for available channels) and channel mobility resulting from the arrival of PUs claiming back-channels used by the SUs.

ACKNOWLEDGMENTS

REFERENCES

[1] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt Generation/Dynamic Spectrum Access: A Survey," *Comput. Netw.*, vol. 50(13), pp. 2127–2159, 2006.

[2] M. Bloem, T. Alpcan, and T. Basar, "A stackelberg game for power control and channel allocation in cognitive radio networks," in *Proc. Valuetools'07*, Oct. 2007.

[3] K. Burdett and D. A. Malueg, "The theory of search for several goods," *J. of Economic Theory*, vol. 24, pp. 362–376, 1981.

[4] K. Burdett and R. Wright, "Two-sided search with nontransferable utility," *Review of Economic Dynamics*, vol. 1, pp. 220–245, 1998.

[5] J. A. Carlson and R. P. McAfee, "Joint search for several goods," *J. of Economic Theory*, vol. 32, pp. 337–345, 1984.

[6] S. Choi and J. Liu, "Optimal time-constrained trading strategies for autonomous agents," in *Proc. of MAMA'2000*, 2000.

[7] C.-T. Chou, S. S. N, H. Kim, and K. G. Shin, "What and How Much to Gain by Spectrum Agility?" *IEEE J. Sel. Areas Commun.*, vol. 25(3), pp. 576–588, 2007.

[8] FCC, "ET Docket No. 03-222, Notice of Proposed Rule Making and Order," Dec. 2003.

[9] FCC, "ET Docket No. 04-186, ET Docket No. 02-380, Second Report And Order And Memorandum Opinion And Order, FCC 08-260," Nov. 2008.

[10] M. Gandetto and C. Regazzoni, "Spectrum Sensing: A Distributed Approach for Cognitive Terminals," *IEEE J. Sel. Areas Commun.*, vol. 25(3), pp. 546–557, Apr. 2007.

[11] J. Gatti, "Multi-commodity consumer search," *Journal of Economic Theory*, vol. 86(2), pp. 219–244, 1999.

[12] A. Ghasemi and E. Sousa, "Spectrum Sensing in Cognitive Radio Networks: Requirements, Challenges and Design Trade-offs," *IEEE Commun.*, vol. 46(4), pp. 32–39, Apr. 2008.

[13] J. Huang, R. A. Berry, and M. L. Honig, "Spectrum Sharing with Distributed Interference Compensation," in *Proc. IEEE DySPAN'05*, Nov. 2005, pp. 88–93.

[14] T. Ito, H. Ochi, and T. Shintani, "A group-buy protocol based on coalition formation for agent-mediated e-commerce," *IJCIS*, vol. 3(1), pp. 11–20, 2002.

[15] K. Jaganathan, I. Menache, E. Modiano, and G. Zussman, "Non-cooperative Spectrum Access - The Dedicated vs. Free Spectrum Choice," in *Proc. ACM MobiHoc'11*, May 2011.

[16] J. Kephart and A. Greenwald, "Shopbot economics," *JAAMAS*, vol. 5(3), pp. 255–287, 2002.

[17] K. Lermann and O. Shehory, "Coalition formation for large scale electronic markets," in *ICMAS'2000*, 2000, pp. 216–222.

[18] S. Lippman and J. McCall, "The economics of job search: A survey," *Economic Inquiry*, vol. 14, pp. 155–189, 1976.

[19] J. McMillan and M. Rothschild, "Search," in *Handbook of Game Theory with Economic Applications*, R. Aumann and S. Hart, Eds., 1994, pp. 905–927.

[20] P. Michiardi and R. Molva, "Analysis of coalition formation and cooperation strategies in mobile ad hoc networks," *Ad Hoc Networks*, vol. 3, 2005.

[21] J. Mitola, "Cognitive Radio for Flexible Mobile Multimedia Communications," *ACM/Kluwer MONET*, vol. 6(5), pp. 435–441, 2001.

[22] C. Peng, H. Zheng, and B. Y. Zhao, "Utilization and Fairness in Spectrum Assignment for Opportunistic Spectrum Access," *ACM/Springer MONET*, vol. 11(4), pp. 555–576, Aug. 2006.

[23] H. Rahul, N. Kushman, D. Katabi, C. Sodini, and F. Edalat, "Learning to Share: Narrowband-Friendly Wideband Wireless Networks," in *Proc. IEEE SIGCOMM'08*, Aug. 2008.

[24] M. Rothschild, "Searching for the lowest price when the distribution of prices is unknown," *Journal of Political Economy*, vol. 82, pp. 689–711, 1974.

[25] D. Sarne, E. Manisterski, and S. Kraus, "Multi-goal economic search using dynamic search structures," *Autonomous Agents and Multi-Agent Systems*, vol. 21, pp. 204–236, 2010.

[26] O. Shehory and S. Kraus, "Methods for task allocation via agent coalition formation," *AIJ*, vol. 101(1-2), pp. 165–200, 1998.

[27] R. Shimer and L. Smith, "Assortative matching and search," *Econometrica*, vol. 68(2), pp. 343–370, 2000.

[28] N. Tsvetovat, K. Sycara, Y. Chen, and J. Ying, "Customer coalitions in electronic markets," *AMEC00*, 2000, pp. 121–138.

[29] M. L. Weitzman, "Optimal search for the best alternative," *Econometrica*, vol. 47, pp. 641–654, May 1979.

[30] J. Yamamoto and K. Sycara, "A stable and efficient buyer coalition formation scheme for e-marketplaces," in *Agents01*, 2001, pp. 576–583.