Information Sharing Under Costly Communication in Joint Exploration

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

Abstract

This paper studies distributed cooperative multi-agent exploration methods in settings where the exploration is costly and the overall performance measure is determined by the minimum performance achieved by any of the individual agents. Such an exploration setting is applicable to various multi-agent systems, e.g., in Dynamic Spectrum Access exploration. The goal in such problems is to optimize the process as a whole, considering the tradeoffs between the quality of the solution obtained and the cost associated with the exploration and coordination between the agents. Through the analysis of the two extreme cases where coordination is completely free and when entirely disabled, we manage to extract the solution for the general case where coordination is taken to be costly, modeled as a fee that needs to be paid for each additional coordinated agent. The strategy structure for the general case is shown to be threshold-based, and the thresholds which are analytically derived in this paper can be calculated offline, resulting in a very low online computational load.

1 Introduction

In many settings, the benefits from the different alternatives available to an agent are uncertain. For example, in eCommerce, a shopbot does not know a priori the pricing of a requested product among the different merchants' web-sites. Similarly, in the Mars exploration rover mission, the rover does not know a priori the terrain conditions in the different locations it can potentially visit. In both examples, the agents can explore the alternatives (denoted "opportunities" onwards) available to them, revealing the actual benefit ("value") with which they are associated, however incurring a cost (e.g., consuming some of their resources) as such exploration is inherently costly. The goal of the agent is not necessarily to find the opportunity associated with the maximum value, but rather to maximize the overall benefit, defined as the value of the opportunity eventually picked minus the costs accumulated along the exploration process.

This exploration process becomes more complex whenever conducted cooperatively by several agents. For example: when the agents are robots that need to evaluate several potential locations for mining a certain mineral on the face of Mars (Hazon, Aumann, and Kraus 2009); a group of buyers that need to evaluate several potential sellers for buying different products (Sarne, Manisterski, and Kraus 2010); and secondary users in Dynamic Spectrum Access applications that need to evaluate different connections to a central server in order to establish a common communication link (Akyildiz et al. 2006). The cooperative exploration is more complex in the sense that the agents' exploration is now affected also by findings of other agents in the group, calling for inter-agent coordination. Alas, coordination is inherently costly. Therefore the agents' cooperative exploration strategy also needs to take into consideration the overhead associated with the coordination between them along the process.

In this paper, we formally introduce a cooperative multilateral exploration model with costly coordination. Specifically, we focus on a type of problems where the benefit of each agent is the minimum among the best results obtained in any of the individual exploration efforts. For example, in Dynamic Spectrum Access applications the quality of service experienced by all of the agents depends on the lowest-quality channel selected for the communication by any of the agents. A similar setting arises in coordination-management applications (e.g., DARPA's project) where the quality of performance of a task is commonly defined by a quality accumulation *minimum* function over the sub-tasks (Smith et al. 2007; Atlas and Decker 2010). The overall system's performance, which the agents attempt to maximize, is thus the sum of the resulting individual benefits minus the costs (of exploration and coordination) accumulated by the different agents along the process. Therefore, the agents may choose to have only some of them coordinate their exploration efforts, while others explore individually with no coordination with the other group members. To the best of our knowledge, this is the first attempt to analyze a model of cooperative exploration constrained by individual findings with communication costs, where the agents can decide on the extent of coordination they will employ throughout their exploration.

2 Related Work

In many multi-agent environments, autonomous agents may benefit from cooperating and coordinating their actions (Rosenfeld et al. 2008). 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 (Lermann and Shehory 2000). Consequently, group-based cooperative behavior has been suggested in various domains (Talukdar et al. 1998; Dias 2004; Tsvetovat et al. 2000). 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 multi-agent systems (MAS) (Li et al. 2003; Shehory and Kraus 1998). 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 (Sandholm et al. 1999), division of coalition payoffs (Yamamoto and Sycara 2001) and enforcement methods for interaction protocols (Michiardi and Molva 2005).

The problem of an agent engaged in exploration in a costly environment, seeking to maximize its long-term utility, is widely addressed in classical search theory (e.g., (McMillan and Rothschild 1994; Smith 2011) and references therein). Over the years, several attempts have been made to adopt search theory concepts for agent-based electronic trading environments associated with exploration costs (Sarne, Manisterski, and Kraus 2010; Kephart and Greenwald 2002). Despite the richness of search theory and its implications, most models introduced to date have focused on the problem of a single agent that attempts to maximize its own expected benefit. Few studies have attempted to extend the exploration problem beyond a single goal, e.g., attempting to purchase several commodities while facing imperfect information concerning prices (Hazon, Aumann, and Kraus 2009; Burdett and Malueg 1981). Some even considered multi-agent cooperative exploration for multiple goals (Sarne, Manisterski, and Kraus 2010) and statistical methods (Reches, Gal, and Kraus 2013) for determining how much exploration is needed. However, none of these works applies any constraints on the values obtained along the exploration process. The only constraint prior work has considered on the values obtained by an agent is the availability of recall (i.e., the ability to exploit formerly explored opportunities) (Carlson and McAfee 1984). Furthermore, none of these works considered costly coordination and its different aspects. Multi-agent exploration constrained by the findings of the other agents can be found in our prior work (2011; 2012) however the models presented there constrain the exploration scheme, either in the sense that the agents are arbitrarily ordered and each agent can explore only after the other agents ordered before it have fully completed their exploration process (hence the coordination question becomes irrelevant) or by binding all agents to the same opportunity at any given exploration step (hence full coordination is mandatory). These constraints preclude the use of a hybrid exploration schemes of the type presented in our paper and imply different exploration strategies, substantially complicating the analysis.

3 The Model

We consider a setting where a group of k cooperating agents attempt to achieve a shared goal. In order for the goal to be achieved each agent needs to engage in a costly exploration. The individual exploration processes involve the evaluation of different opportunities which values are a priori uncertain. This uncertainty is modeled through a probability distribution function $f(x)$, i.e., the value of each opportunity in any of the individual exploration processes is drawn from $f(x)$. The exploring agent can eliminate the uncertainty associated with the value of any given opportunity at a cost c (expressed in terms of opportunity values), as the agent needs to consume some of its resources as part of this process. The model assumes that the agents are not limited by the number of opportunities they can evaluate. This individual costly exploration model is standard and the above assumption are common in prior work (Kephart and Greenwald 2002).

Once all agents have completed their individual exploration, the benefit of each of the k agents from the resulting cooperative exploration process is the minimum among the best results obtained in any of the individual exploration processes, denoted v^*. The overall benefit is thus kv^* minus the costs accumulated along the individual explorations. Since the agents are cooperative, their goal is to maximize the overall expected benefit.

Taking the Dynamic Spectrum Access application domain as an example, each agent represents a terminal and all terminals are interested in establishing a connection between them (e.g. for a conference call, document/video sharing or a multi-player game). The terminals are located in different geographical locations and each terminal can use different wireless channels to connect to a central server supporting the requested application. Each terminal senses different channels of different qualities until it selects a specific channel with a specific quality. The sensing is costly in that the terminal needs to allocate some of its resources for the task (e.g., energy or delay other transmissions). The quality of service provided by the application depends on the lowest quality channel selected by any of the terminals (e.g., if one of the terminals ends up with a low quality channel, the experience of all of the users will be negatively affected).

The agents are assumed to be fully rational and acquainted with the distribution function $f(x)$, the number of agents k and the exploration cost c. Their decision whether to evaluate additional opportunities along their exploration thus needs to take into consideration the tradeoff between the marginal improvement in the value of v^* and the cost incurred.

The model assumes the agents can share their findings along their exploration process through costly communication. Specifically, we assume that in order to communicate, the agents need to lease communication services from an external operator. While there are numerous cost schemes that can be used, our model assumes that the cost of unlimited communication throughout the exploration between $k' \leq k$ agents is $k'c_m$, where c_m is the cost of subscribing an additional agent to the service.

4 Analysis

The analysis considers the case where the communication cost is substantially high, thus the agents prefer to avoid it completely, and the case where the communication is free (i.e., $c_m = 0$), thus fully adopted. These facilitate the analysis of the general case where $c_m > 0$, wherein not all the agents necessarily employ coordination.

4.1 Non-Coordinated Exploration

If the agents cannot communicate whatsoever, then their exploration takes place separately and can be perceived as performed in parallel, where the value of v^* is revealed only after all individual explorations have come to an end.

The optimal cooperative strategy in this case is based on having each agent use a reservation value (i.e., a threshold) for its individual exploration. Each agent will keep exploring as long as the best value found is below the reservation value it uses. The use of a reservation-value strategy results from the fact that the agent's state depends solely on the best value found so far. More formally, the optimality (in terms of overall expected-benefit-maximization) of the reservation-value-based strategy in this case derives from the fact that if the agent finds it beneficial to explore when the best value obtained so far is v, then it is inevitably beneficial to explore if the best value found so far is v' for any $v' < v$ (and vice versa).

Theorem 1. *The cooperative expected-benefit-maximizing exploration strategy with no communication between the agents is to have each agent A_i use a reservation value r that satisfies:*

$$c=k\int_{y=r}^{\infty}f(y)\Big(\int_{x=-\infty}^{\infty}(min(y,x)-min(r,x))\bar{f}(x)dx\Big)dy \quad (1)$$

where $\bar{f}(x)$ is the probability distribution function of the minimum among the best values obtained by all other agents, i.e, $min(x_1,..,x_{i-1},x_{i+1},..,x_k)$.

Proof. The overall expected benefit of the system from the cooperative exploration is the expected value with which the agents end up, denoted $E[v^*]$, multiplied by k, minus the accumulated costs along the exploration. Now consider the effect of agent A_i's exploration over $E[v^*]$. Given the distribution of the minimum among the best values obtained by all other agents, $\bar{f}(x)$, the value of $E[v^*]$ when A_i uses a reservation value r is given by:

$$E[v^*]=\int_{y=-\infty}^{r}E[v^*]f(y)dy+\int_{y=r}^{\infty}f(y)\int_{x=-\infty}^{\infty}min(y,x)\bar{f}(x)dxdy \quad (2)$$

The first term above relates to the case where a value $y < r$ is obtained through an additional exploration, in which case agent A_i resumes its exploration process according to the reservation value rule. Since A_i is not limited by the number of opportunities it can evaluate, it now faces the exact same decision problem as before, resulting in an expected value $E[v^*]$. The second term relates to the case where the value obtained is above r, in which case the agent terminates its exploration and the expected value $E[v^*]$ is the minimum value found among the value obtained by A_i and the values obtained by the other agents (captured by the distribution function $\bar{f}(x)$). The expected cost accumulated along the exploration of A_i when using r is given by $\frac{c}{1-F(r)}$, as this becomes a Bernoulli sampling process with a success probability of the value obtained being greater than the threshold used, $1 - F(r)$. Therefore, the overall expected benefit of the system as a function of the reservation value r used by A_i, denoted $B(r)$, is given by (after isolating $E[v^*]$ in Equation 2):

$$B(r) =\frac{k\int_{y=r}^{\infty}f(y)\int_{x=-\infty}^{\infty}min(y,x)\bar{f}_i(x)dxdy}{1-F(r)} - \frac{c}{1-F(r)} - C \quad (3)$$

where C denotes the expected cost accumulated along the other agents' exploration. In order to find the expected-benefit-maximizing reservation value r, we take the first derivative of Equation 3 and set to zero, resulting in:

$$c =k\int_{y=r}^{\infty}f(y)\Big(\int_{x=-\infty}^{\infty}min(y,x)\bar{f}(x)dx\Big)dy - k(1-F(r))\int_{x=-\infty}^{\infty}min(r,x)\bar{f}(x)dx \quad (4)$$

which after some mathematical manipulations becomes Equation 1. □

Equation 1 has an intuitive interpretation that is commonly found in other exploration-based models (Chhabra, Das, and Sarne 2011; McMillan and Rothschild 1994) — the overall expected-benefit-maximizing reservation value, r, is the value where the agent is precisely indifferent: the expected marginal benefit from obtaining the value of the opportunity (represented by the right-hand term of the equation) exactly equals the cost c of obtaining that additional value.

Since all agents face a similar setting (characterized by $f(x)$ and c) they all use the same reservation value r. This enables a simple formulation of the function $\bar{f}(x)$:

$$\bar{f}(x) = \frac{d(1-(1-F^{returned}(x))^{k-1})}{dx}, \quad (5)$$

where:

$$F^{returned}(x) = \begin{cases} 0 & x \le r \\ \frac{F(x)-F(r)}{1-F(r)} & x > r \end{cases} \quad (6)$$

The function $F^{returned}(x)$ returns the probability that the best value with which an agent that uses a reservation value r ends up when terminating its exploration will be below x. The term $(1-(1-F^{returned}(x))^{k-1})$ is thus the probability that the minimum among the results of the other $k-1$ agents' explorations will turn out to be below x, and therefore its derivative is the probability distribution function of the minimum among the best values obtained by all other agents.

Using Equation 1 we can now calculate r, and since all agents use the same reservation value, the probability distribution function of the minimum among the best values found by *all* agents (unlike with $\bar{f}(x)$ which apply to all agents except one) is given by: $\frac{d(1-(1-F^{returned}(x))^k)}{dx}$. Therefore:

$$E[v^*] = \int_{x=-\infty}^{\infty} \big(x \cdot \frac{d(1-(1-F^{returned}(x))^k)}{dx}\big)dx \quad (7)$$

The accumulated cost along the exploration process of each of the agents is given by $\frac{c}{1-F(r)}$, as the success probability is $(1-F(r))$, hence the system's overall expected benefit is:

$$EB = k\big(E[v^*] - \frac{c}{1-F(r)}\big) \quad (8)$$

4.2 Fully Coordinated Exploration

If the agents can communicate without incurring a cost, then their exploration strategy should take into consideration, at each step of the process, the values found by any of the other agents. Furthermore, since the exploration is costly, it is advantageous for the agents to execute their exploration sequentially (having one agent explore at a time) rather than in parallel. Since the state of the agents now depends on the vector of best values found by the different agents, the overall expected-benefit-maximizing strategy is no longer based on a single reservation-value. Instead, as we prove in this section, it assigns a different reservation value for each state and applies it to (i.e., compares it with) the minimum among the set of best values found by the different agents.

We represent the system's state as a vector $V = (v_1, .., v_k)$, in the k-dimensional space, where v_i ($1 \leq i \leq k$) is the best value found so far by agent A_i.[1] We use $S(V) \rightarrow \{i, terminate\}$ to denote the agents' strategy, where i ($1 \leq i \leq k$) suggests that agent A_i needs to execute an exploration step next and *terminate* means the exploration as a whole should be terminated. If the exploration is terminated, then the value of v^* is determined according to the minimum value of V. For convenience we use V_{min} to denote the minimum value in V (i.e., $V_{min} = min(v_1, .., v_k)$). Due to the nature of v^* it is obvious that given a state V, if the optimal strategy is to resume the exploration, then the agent who should be evaluating an additional opportunity is the one whose highest value is V_{min}.[2] This is simply because any increase in the best value obtained by any other agent A_i associated with the best value $v_i > V_{min}$ can affect v^* only if it is accompanied by findings greater than v_i of agents currently associated with the best values lower than v_i. The agents' strategy can therefore be expressed as: $S(V) \rightarrow \{resume, terminate\}$.

Proposition 1. *For any state* $V = (v_1, .., v_k)$*, if* $S(V) = resume$ *then inevitably* $S(V') = resume$ *for any* V' *differing from* V *only in the value of its minimum, where* $V'_{min} < V_{min}$ *(formally:* $V' = (v_1, ..., v_{i-1}, v'_i < v_i, v_{i+1}, ..., v_k)$ *where* $v_i = V_{min}$*).*

Proof. Assume otherwise, i.e., $S(V') = terminate$. Since $S(V) = resume$, then once a state V is reached, the agent associated with V_{min} will resume exploring until a better value is obtained. From the system's point of view, this is preferable over terminating the exploration, i.e., over ending up with a value V_{min}. Now consider the option to resume exploration by the agent associated with value V'_{min} when starting from a state V' until obtaining a value greater than V_{min}. The expected exploration cost and the distribution of the value with which the agent will end up (i.e., above V_{min}) is equal in both cases. Therefore, since terminating the exploration process when in state V' yields $V'_{min} < V_{min}$, the strategy $S(V') = terminate$ cannot be optimal (in terms of expected-benefit-maximization). □

The main implication from Proposition 1 is that for all states that differ only in the value of their minimum element there is a single reservation value for determining whether to resume exploration. Whenever reaching a new state, each agent needs to determine if the best value it had obtained so far is V_{min}. The agent associated with V_{min} will calculate the reservation value according to the current state V, denoted $r(V)$, and resume the exploration if its value is below $r(V)$.

We use $\sigma(V, y) \rightarrow V'$ to denote the new state to which the system transitions after the agent associated with V_{min} has obtained a value y in its exploration, if it was initially in state V. If v_i is the minimum value in V (i.e., $v_i = V_{min}$) then:

$$\sigma(V, y) = \begin{cases} V & y \leq V_{min} \\ (v_1, ..., v_{i-1}, y, v_{i+1}, ..., v_k) & \text{otherwise} \end{cases} \quad (9)$$

The system's expected benefit if acting according to the expected-benefit-maximizing strategy $S(V)$ is thus given by the following recursive equation:

$$EB(V) = \begin{cases} kV_{min} & r(V) \leq V_{min} \\ -c + \int_{y=-\infty}^{r(V)} EB(V)f(y)dy + & \text{otherwise} \\ \int_{y=r(V)}^{\infty} EB(\sigma(V, y))f(y)dy & \end{cases} \quad (10)$$

Theorem 2. *The expected-benefit-maximizing exploration strategy with full communication, when in state* V*, is to terminate the exploration if* $V_{min} \geq r(V)$ *and otherwise resume exploration, where* $r(V)$ *is the solution to:*

$$c = \int_{y=r(V)}^{\infty} (EB(\sigma(V, y)) - kr(V))f(y)dy \quad (11)$$

The value $r(V)$ *is the same for all states differing only by their minimum value (thus can be calculated only once for these states).*

Proof. In order to find the expected-benefit-maximizing reservation value $r(V)$ we take the first derivative of Equation 10 and set to zero, resulting in:

$$c = \int_{y=r(V)}^{\infty} EB(\sigma(V, y))f(y)dy - (1 - F(r(V)))kr(V) \quad (12)$$

which after some mathematical manipulations becomes Equation 11. □

The overall expected benefit of the system, denoted EB, can be calculated using Equation 10 when starting from the state where none of the agents has engaged in exploration yet, i.e., $EB = EB(0, .., 0)$.

[1] If agent A_i has not yet engaged in an exploration then $v_i = 0$.

[2] In the case where two or more agents hold a value equal to V_{min} there is no importance to the selection of which of them will explore next.

4.3 Coordination with Costly Communication

Finally, we analyze the general case where $c_m > 0$, wherein not all of the agents necessarily employ coordination. In this case, it is possible that only a subset of agents will use coordination or that the agents will divide into groups that coordinate their exploration separately. Since the communication cost is linear in the number of agents, the overall expected-benefit-maximizing strategy is to have at most one group of agents coordinate their exploration and to have the remaining agents execute their exploration in isolation. This is because whenever two groups of agents merge and coordinate their exploration, the same performance as with the two separate groups can be achieved simply by asking each agent to follow the exploration strategy it would have used if operating in its original group (hence the division into groups cannot possibly improve the performance of a single unified group).

Consider the case where k' agents use costly coordination and the remaining $k - k'$ agents explore with no coordination. The system's expected value (i.e., the minimum value with which the agents end up) when the agents exploring individually use reservation value r and the agents using the coordination service use a reservation value function $r(V)$, given that a state $V = \{v_1, ..., v_{k'}\}$ was reached in the coordinated exploration, denoted $EV(V)$, is given by:

$$EV(V)=\begin{cases}\int_{x=-\infty}^{\infty} min(x, V_{min})\bar{f}_p(x)dx & r(V)\leq V_{min}\\ \int_{y=-\infty}^{r(V)} EV(V)f(y)dy+ & \text{otherwise}\\ \int_{y=r(V)}^{\infty} EV(\sigma(V,y))f(y)dy & \end{cases} \quad (13)$$

where $\bar{f}_p(x) = \frac{d(1-(1-F^{returned}(x))^{k-k'})}{dx}$ is the probability distribution function of the minimum among the values obtained by the agents exploring individually with no coordination service (i.e., the equivalent of $\bar{f}(x)$ for $k-k'$ agents). The above calculation method of $EV(V)$ resembles the one used in Equation 10, however it also takes into consideration the results of the agents that manage their exploration individually.

Similarly, the expected cost thereafter of those agents exploring through the coordination service when using a reservation value function $r(V)$, given that a state V was reached in the coordinated exploration, is given by:

$$EC(V)=\begin{cases}0 & r(V)\leq V_{min}\\ -c+\int_{y=-\infty}^{r(V)} EC(V)f(y)dy+ & \text{otherwise}\\ \int_{y=r(V)}^{\infty} EC(\sigma(V,y))f(y)dy & \end{cases} \quad (14)$$

The overall expected benefit of the system onwards, given state V of the coordinated exploration, denoted $EB(V)$, can be calculated using Equations 13 and 14:

$$EB(V) = kEV(V) - EC(V) - \frac{(k-k')c}{1-F(r)} - k'c_m \quad (15)$$

and, in particular, when starting from the state where none of the agents has engaged in exploration yet: $EB{=}EB(0,..,0)$.

Before introducing the optimal exploration strategy in this case, we introduce two complementary notations. First, $f_c(V, x)$ is the probability distribution function of the minimum among the best values obtained by the k' agents exploring with coordination, given that they start from state V:

$$f_c(V,x)=\begin{cases}0 & x<V_{min}\vee(V_{min}\neq x\wedge r(V)\leq V_{min})\\ 1 & V_{min}=x\wedge r(V)\leq V_{min}\\ \int_{y=-\infty}^{r(V)} f_c(V,x)f(y)dy+ & \text{otherwise}\\ \int_{y=r(V)}^{\infty} f_c(\sigma(V,y),x)f(y)dy & \end{cases}$$

The calculation method of $f_c(V, x)$ resembles the calculation of $EV(V)$ according to Equation 13. The probability of obtaining a value lesser than or equal to x by the k' agents, denoted $F_c(x)$, is thus given by: $F_c(x) = \int_{y=-\infty}^{x} f_c((0,..,0), y)dy$.

Second, we denote the probability distribution function of the minimum among the best values obtained by all agents, except for one agent that explores individually, by $\bar{f}_l(x)$. This latter term facilitates the analysis of the effect of r on the system's expected benefit and can be calculated as:

$$\bar{f}_l(x)=\frac{d(1-(1-F_c(x))(1-F^{returned}(x))^{k-k'-1})}{dx} \quad (16)$$

The term $(1 - (1 - F_c(x))(1 - F^{returned}(x))^{k-k'-1})$ is in fact the probability that the minimum among the results of the other $k - 1$ agents' explorations is below x, and therefore its derivative is the probability distribution function of the minimum among the best values obtained by all other agents. The function $F^{returned}(x)$ can be calculated according to Equation 6.

At this point, we have everything needed to introduce Theorem 3, which specifies the optimal exploration strategy for the case where $c_m > 0$.

Theorem 3. *The optimal exploration strategy where k' agents coordinate their exploration and $k - k'$ agents explore separately is to set a reservation value r for the agents exploring separately and a reservation value function $r(V)$ for the k' agents that use coordination, according to the solution of the set of equations:*

$$c = k\int_{y=r(V)}^{\infty}(EV(\sigma(V,y)) - EV(r(V)))f(y)dy \quad (17)$$

$$c=k\int_{y=r}^{\infty} f(y)\Big(\int_{x=-\infty}^{\infty}(min(y,x)-min(r,x))\bar{f}_l(x)dx\Big)dy \quad (18)$$

Proof. The proof relies in large on the proofs given for the two previous cases (Theorems 1 and 2), therefore we only detail the differences. Equation 17 augments Equation 11 in a way that considers the effect of the minimum best value found by any of the agents that explore separately in parallel (i.e., without communication). It is obtained by taking the first derivative of Equation 15 according to $r(V)$, equating it to zero and applying some standard mathematical manipulations. Equation 18 augments Equation 1 in a way that takes into consideration in $f_c(x)$ the minimum value found by the agents exploring in coordination in addition to the minimum among those exploring individually in parallel. □

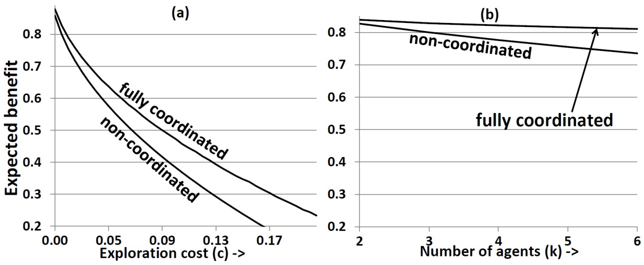

Figure 1: (a) Expected benefit of 3-agents' exploration as a function of the exploration cost. (b) Expected benefit (per agent) of k agents' exploration as a function of k (using $c = 0.01$). Both settings use $c_m = 0$.

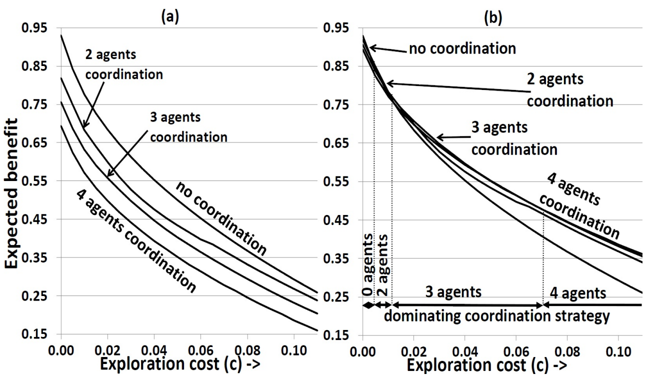

Figure 2: Expected benefit (per agent) as a function of the exploration cost, for $k = 4$, where the cost of coordination services is: (a) $c_m = 0.25$; (b) $c_m = 0.1$.

In order to find the optimal exploration strategy, we need to extract the overall expected benefit for any number of agents exploring with coordination (i.e., for k' =0,2,3,..., k) according to Theorem 3. The number of agents exploring in coordination for which the highest expected benefit, EB, is obtained is the one by which the agents should operate.

5 Numerical Illustration

In order to illustrate the performance achieved with the different methods, we use a tractable synthetic setting that simplifies calculations yet enables demonstrating the main solution characteristics. The setting uses a uniform distribution function defined over the interval $(0, 1)$ (i.e., $f(x) = 1$, $\forall 0 \leq x \leq 1$ and zero otherwise).

We first demonstrate the effect of the increase in the exploration cost c and the increase in the number of agents exploring cooperatively over the individual expected benefit using $c_m = 0$ (see Figure 1). As expected, the fully coordinated case dominates exploration with no coordination, as far as expected benefit is concerned, and both the increase in exploration costs and in the number of agents exploring cooperatively result in a decrease in the expected benefit (per agent). The correlation between the expected benefit and the number of agents is explained by the fact that as the number of agents increases, the expected minimum of the obtained values decreases and more exploration is required.

Finally, Figure 2 depicts the expected benefit as a function of the exploration cost for $k = 4$, when each curve depicts a different number of agents using coordination services. The fee for coordination services is $c_m = 0.25$ (Figure 2(a)) and $c_m = 0.1$ (Figure 2(b)), and the agents can choose to have $0, 2, 3, 4$ of them operate in coordination. As expected, when the coordination cost is high, agents will explore in parallel with no coordination (Figure 2(a)), and when low, coordination is preferred to different extents (Figure 2(b); the number of agents using coordination is depicted at the bottom). From Figure 2(b) it is notable that the choice of how many agents will use coordination depends also on the exploration cost — when the exploration cost is low, agents can compensate over the lack of costly communication with cheap extended exploration. As the exploration cost increases, the value of coordination increases as the saving achieved in repeated exploration becomes substantial.

6 Discussion and Conclusions

The importance of coordination in cooperative exploration is in enabling the agents to refine their exploration strategy based on the findings of others. Yet, since coordination (and in particular communication) is inherently costly, the agents should carefully reason about its extent of use. In many settings, the agents may find it more beneficial to have only some of them (if at all) coordinate their exploration. The analysis given in this paper facilitates the decision regarding the amount of coordination they should apply.

The analysis proves that the optimal strategy that should be used by the agents in cooperative exploration settings with partial coordination where the value of each agent from the process depends on the minimum value found is reservation-value based. Agents exploring in isolation (i.e., with no communication with the others) will use a stationary reservation value, whereas those that use the coordination service will be using a state-based reservation value. The sequential nature of the exploration process used enables some level of separation in the analysis. The resulting set of equations that needs to be solved is actually based on each agent's best-response strategy given the distribution of the minimum value resulting from the exploration strategies of the other agents. By proving that the optimal exploration scheme is to have a sub-group of agents using the coordination service and all other agents exploring with no communication with the others, one only needs to solve for the expected-benefit-maximizing exploration strategy, for a number of exploration schemes that is linear in the number of agents.

Finally, we propose several possible extensions of this work for future work. The first considers the case where the cost of communication is not linear in the number of agents. In this case, it is possible that the expected-benefit-maximizing strategy will be based on having the agents divide themselves into several coordinated-groups, each using the coordination service separately. The second considers other charging schemes set by a self-interested communication provider, e.g., a charge per communication message. In this case the exploration itself will be of a different structure, in comparison to the reservation-value based scheme used in this paper.

Acknowledgment

This work was partially supported by the Israeli Ministry of Industry and Trade under project RESCUE.

References

Akyildiz, I. F.; Lee, W.-Y.; Vuran, M. C.; and Mohanty, S. 2006. NeXt Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey. *Comput. Netw.* 50:2127–2159.

Atlas, J., and Decker, K. 2010. Coordination for uncertain outcomes using distributed neighbor exchange. In *Proc. AAMAS*, 1047–1054.

Burdett, K., and Malueg, D. A. 1981. The theory of search for several goods. *JET* 24:362–376.

Carlson, J. A., and McAfee, R. P. 1984. Joint search for several goods. *JET* 32:337–345.

Chhabra, M.; Das, S.; and Sarne, D. 2011. Expert-mediated search. In *Proc. AAMAS*, 415–422.

Dias, M. 2004. *TraderBots: A New Paradigm for Robust and Efficient Multirobot Coordination in Dynamic Environments*. Ph.D. Dissertation, Robotics Institute, CMU.

Hazon, N.; Aumann, Y.; and Kraus, S. 2009. Collaborative multi agent physical search with probabilistic knowledge. In *Proc. IJCAI*, 167–174.

Kephart, J., and Greenwald, A. 2002. Shopbot economics. *JAAMAS* 5:255–287.

Lermann, K., and Shehory, O. 2000. Coalition formation for large-scale electronic markets. In *Proc. ICMAS*, 167–174.

Li, C.; Chawla, S.; Rajan, U.; and Sycara, K. 2003. Mechanisms for coalition formation and cost sharing in an electronic marketplace. In *Proc. EC03*, 68–77.

McMillan, J., and Rothschild, M. 1994. Search. In *Proc. Handbook of Game Theory with Economic Applications*, 905–927.

Michiardi, P., and Molva, R. 2005. Analysis of coalition formation and cooperation strategies in mobile ad hoc networks. *Ad Hoc Networks* 3:193–219.

Reches, S.; Gal, Y. K.; and Kraus, S. 2013. Efficiently gathering information in costly domains. *Decision Support Systems*.

Rochlin, I.; Sarne, D.; and Laifenfeld, M. 2012. Coordinated exploration with a shared goal in costly environments. In *Proc. ECAI*, 690–695.

Rochlin, I.; Sarne, D.; and Zussman, G. 2011. Sequential multilateral search for a common goal. In *Proc. IAT*, 349–356.

Rosenfeld, A.; Kaminka, G. A.; Kraus, S.; and Shehory, O. 2008. A study of mechanisms for improving robotic group performance. *Artif. Intell.* 172:633–655.

Sandholm, T.; Larson, K.; Andersson, M.; Shehory, O.; and Tohme, F. 1999. Coalition structure generation with worst case guarantees. *AIJ* 111:209–238.

Sarne, D.; Manisterski, E.; and Kraus, S. 2010. Multi-goal economic search using dynamic search structures. *JAAMAS* 21:204–236.

Shehory, O., and Kraus, S. 1998. Methods for task allocation via agent coalition formation. *AIJ* 101:165–200.

Smith, S.; Gallagher, A.; Zimmerman, T.; Barbulescu, L.; and Rubinstein, Z. 2007. Distributed management of flexible times schedules. In *Proc. AAMAS*, 472–479.

Smith, L. 2011. Frictional matching models. *Annual Reviews in Economics* 3:319–338.

Talukdar, S.; Baerentzen, L.; Gove, A.; and de Souza, P. S. 1998. Asynchronous teams: Cooperation schemes for autonomous agents. *Heuristics* 4:295–321.

Tsvetovat, N.; Sycara, K.; Chen, Y.; and Ying, J. 2000. Customer coalitions in electronic markets. In *Proc. AMEC*, 121–138.

Yamamoto, J., and Sycara, K. 2001. A stable and efficient buyer coalition formation scheme for e-marketplaces. In *Proc. Agents'01*, 576–583.