MULTIPLE CRITERIA DECISION MAKING
Vol. 12 2017

David M. Ramsey[*]

THE USE OF INITIAL FILTERS TO DIRECT SEARCH IN DECISION PROCESSES

DOI: 10.22367/mcdm.2017.12.10

Abstract

When looking for a valuable resource, many people use information from the Internet as a way of choosing a small number of offers to investigate in more detail. This paper considers strategies based on the filtering of initial information. A new model is presented according to which the goal of the decision maker is to maximise her expected reward from search taking into account the search costs. The effectiveness of strategies based on filtering is compared to sequential search based on a threshold and exhaustive search of a chosen number of items.

Keywords: filtering, job search problem, multiple signals, sequential search.

1 Introduction

When searching for a valuable good (e.g. a car or house), individuals very often use a two-stage process. In the first stage, the Internet or a specialist magazine may be used in order to find a number of offers that seem promising. This stage is normally characterised by the ability to compare general information regarding a large number of offers at relatively low cost. This stage may be thought of as wide, shallow search. On the basis of this initial information, the searcher may choose a number of offers to investigate more closely. It is assumed that the search costs for obtaining additional information at this stage are relatively large. Hence, the second stage may be called narrow, deep search. During (or after) this stage, the searcher may decide to purchase one of the offers and thus stop searching, or to return to the first stage of searching. Hence, the search process can be charac-

[*] Wrocław University of Science and Technology, Faculty of Computer Science and Management, Department of Operations Research, Wrocław, Poland, e-mail: david.ramsey@pwr.edu.pl.

terised as being partially parallel, since a number of offers can be considered at the same time. It is also partially sequential, since if the searcher returns to the first stage of search, then only new offers will be considered.

MacQueen and Miller (1960) presented a model of sequential search. A decision maker observes a (theoretically unlimited) sequence of offers whose values come from a known distribution and must at each stage decide whether to accept or reject an offer. A fixed cost is paid for observing each offer. Once an offer has been accepted, then the decision maker ceases searching. The optimal strategy for this problem is to accept the first offer whose value exceeds the expected reward from search, i.e. a threshold strategy. Various variations of such a problem have also been considered in the economics literature, in terms of a consumer searching for a job or product, and in the biological literature, particularly as models of mate choice in which just females are choosy. Stigler (1961) considered a similar model of consumer search. He assumed that the strategy of a consumer, the so called "best-of-n rule", is defined by the number of offers to observe, n. The consumer accepts the best of the n offers observed. The optimal strategy, which observes n^* offers, satisfies the condition that n^* is the largest integer n satisfying the condition that the expected marginal gain from observing the n-th offer (calculated before any offers are observed) is greater than the cost of observing the n-th offer. Hence, Stigler assumes that a consumer can return to a previously observed offer. However, given that the number of offers is unlimited and search costs are proportional to the number of objects seen, the threshold rule presented by MacQueen and Miller (1960) indicates that the recall of previous offers is of no value. Janetos (1980) compares various types of search rules, including threshold rules and best-of-n rules, in the context of mate choice. Based on his assumptions, the number of prospective mates was essentially fixed and search costs were not specifically modelled, Janetos concluded that best-of-n rules are the most effective. Real (1990) incorporates search costs into his model of mate choice and shows that when search costs are proportional to the number of prospective mates seen, then threshold rules are optimal. Hutchinson and Hałupka (2004) consider a model in which the population of males is patchy (i.e. search costs are not proportional to the number of males seen). They conclude that best-of-n type rules work best when males group together in so-called leks (i.e. recall of recently seen prospective partners is possible and cheap) and there is little spatial variation in the values of males. When males are more dispersed (i.e. recall of previously seen prospective partners is expensive or impossible) and/or there is spatial variation in the values of mates, then threshold rules are more successful.

The articles considered above assume that all the information regarding the value of an offer is observed at a single moment. MacQueen (1964) extended this approach to model sequential search in which the decision maker can gain additional information about an offer, which incurs additional inspection costs. The decision maker first observes a quantitative signal of the value of an object and on the basis of this observation chooses one of the following three options: a) immediately accept the offer and stop searching, b) gain additional information, in the form of a second signal, about the offer, c) reject the offer and continue searching. In case b), after observing the additional information, the decision maker must decide whether to accept or reject the offer. When the joint distribution of the two signals is known and the search costs are linear in both the number of offers seen and the number of times that the second signal is observed, then the optimal strategy is defined by a vector of three thresholds (t_1, t_2, t_3). If the quantitative measure of the first signal is less than t_1, then an offer should be immediately rejected. If this measure is greater than t_2, then the offer should be immediately accepted. Otherwise, additional information should be gathered about the offer. After the second signal has been observed, then the offer should be accepted if and only if the value of the object based on the two signals is at least t_3. Ramsey (2015) derives the specific form of this optimal policy when the two signals come from a joint normal distribution and the value of an offer is equal to a linear combination of the two signals. This approach is based on the concept of multi-attribute linear utility (MLU, see Keeney and Raiffa, 1993), which is one approach sometimes used to solve multi-criteria decision problems. Lim et al. (2006) extend this model to one where multiple signals describing the value of an offer are available. It is assumed that these signals must be observed in a particular order and that the expected value of an offer given the value of a signal is increasing in the quantitative measure of this signal. They define a dynamic programming procedure that derives the optimal policy for such a problem. Wiegmann et al. (2010) consider a similar problem within the framework of a mate choice problem. Bearden and Connolly (2007a, 2007b) consider an approach to such problems based on the concept of satisficing (see Simon, 1995). Based on such a strategy, the second signal is observed only when the quantitative measure of the first signal exceeds the appropriate threshold. Given that the second signal is observed, then the offer is accepted if and only if the quantitative measure of this signal exceeds the specified threshold. They show that such strategies can be close to optimal and on the basis of experiments find that individuals who are restricted to using satisficing strategies are on average as successful as individuals who are allowed to use strategies of the same form as the optimal strategy. Chun (2015) uses a similar approach, but assumes that the most important trait is optimised subject to satisficing other conditions.

Hogarth and Karelaia (2005) consider a problem in which a number of binary signals can be observed for each offer and offers can be observed in parallel. In accordance with the concept of satisficing, these signals can be interpreted as stating whether a given condition is satisfied or not. It is assumed that these conditions can be ranked according to importance. One approach to selecting an offer is the DEBA Procedure (Deterministic Elimination by Aspects). The decision maker starts with the most important condition and eliminates those offers that do not satisfy this condition. The decision maker then passes on to successively less important criteria and at each stage eliminates those offers that do not satisfy the present criterion. If only one offer is left, then that offer is chosen. If none of several remaining offers satisfy the present condition, or all of the conditions have been considered and several offers still remain, then an offer is chosen at random from those remaining. The effectiveness of such a procedure is compared to the so-called Equal Weighting Procedure (EW), according to which all the signals are observed for each offer and the offer that satisfies the most conditions is accepted (or a random choice made between the offers satisfying the most criteria). It is shown that when preferences are non-compensatory, i.e. the weight of a given condition is greater than the sum of the weights of less important criteria, then the DEBA procedure outperforms the EW procedure. Baucells et al. (2008) investigates the effectiveness of the DEBA procedure under various assumptions. It should be noted that in problems of this type the decision maker can choose in which order to observe the signals, unlike in the problem considered in this paper, where it is assumed that there is a natural order in which the signals must be observed. Also, these models do not explicitly model the costs of searching or making decisions. In the context of mate choice, Fawcett and Johnstone (2003) consider a similar problem where two binary traits are associated with the value of a prospective mate, but they define explicit costs for observing a trait. They derive conditions for the optimality of i) random mating, ii) mating based on one signal and iii) mating based on both signals. In the case that both signals should be considered, they derive a condition which determines which signal should be observed first.

Analytis et al. (2014) consider a two-stage search process, in which the information obtained in the first stage is used to decide in which order the offers should be investigated in the second round. Hence, in the first round comparison is made in parallel. The second round involves sequential inspection of the offers, starting with the offer that was adjudged to be the best in round one. At this stage, an offer is accepted and search ceases when the expected value of the offer given the second signal is greater than the expected reward from continued search (i.e. a threshold rule is used). This approach is developed in Analytis et al.

(2015), where a decision maker is searching for a good whose utility depends both on his/her preferences and (positively) on the popularity of that good (this might be appropriate for modelling the utility of e.g. a book or CD). Since Internet suppliers often order goods according to popularity, buyers use this as an initial filter in a sequential search procedure. Chhabra et al. (2014) look at a similar problem from the point of view of an information provider who wishes to set prices for services.

This article is intended as a starting point for building a model of searching for a valuable good which has unique (or near unique) characteristics, e.g. a house or second hand car. In the first round of search, information can be gained on a relatively large number of offers in parallel at low cost. The search costs incurred in the second round are assumed to be greater. Also, it is assumed that the information obtained in the first stage is of a more quantitative nature than the information gained in the second stage. For example, the description of a flat in the Internet will give specific information about the size, price and location of the flat, which can be evaluated in numerical terms. Visiting a flat will give additional information, which will tend to be of a more qualitative nature. Hence, the model assumes that in the first stage of search a quantitative measure of the value of an offer will be observed and after the second stage the searcher can only rank the offers.

Another of the goals of the article is to highlight conditions under which a strategy based on filtering can be optimal. Under such a strategy, on the basis of the first stage, the decision maker should choose a relatively small number of promising offers to observe more closely in the second round, before making a final decision. In particular, a comparison of the article of MacQueen (1964) with the research of Hutchinson and Hałupka (2004) is particularly useful in this regard. Sequential search using a threshold strategy, rather than a filtering strategy, will be optimal when the search costs are proportional to the number of offers observed and the searcher perfectly observes quantitative signals, as well as having perfect information about the distribution from which signals come (and their relation to the value of an offer). Hutchinson and Hałupka (2004) note that when the search costs are not linear and recall of previously seen mates is possible (this is the case when offers are observed in parallel), then best-of-n type rules may work very well. This is also true when decision makers do not have perfect information regarding the distribution of the value of offers. Also best-of-n type rules will work well when signals cannot be observed precisely, but can only be ranked according to attractiveness (for example, it is impossible to say anything about the attractiveness of the first offer observed, since it cannot be compared with anything). It follows that filtering strategies should be effective when one or more of the following conditions are satisfied:

1. In the first round a number of offers can be observed in parallel and the average search cost per offer is decreasing in the number of observations.
2. The decision maker does not have perfect information regarding the distribution of the signals or their relation to the value of offers.
3. Some of the signals cannot be precisely measured, but are comparable.

The model considered here assumes that the second signal cannot be precisely measured. However, search costs are assumed to be proportional to the number of offers investigated and it is implicitly assumed that the decision maker knows the distribution of the signals. Intuitively, if the amount of information contained in the second signal is sufficiently large, then the optimal strategy should be a filtering strategy.

Section 2 presents the model and describes how the optimal strategy is derived. Three types of strategy which are particularly important in this framework are highlighted: i) sequential search using a threshold based purely on the first signal, ii) full investigation of a number n of observations, iii) filtering strategies. Section 3 presents an example where the two signals have exponential distributions. This is used to illustrate the derivation of the optimal strategy. Section 4 presents the results and, in particular, concentrates on the form of the optimal policy according to the relative amount of information contained in the second signal. Conclusions and some directions for future research are given in Section 5.

2 A model with incomplete information

This section considers a new model of two-stage search in which the signal observed in the first round of search can be summarised using a quantitative variable, X_1, whereas the overall impression of an offer after the second round can only be ranked in comparison to the other offers observed within a batch. However, it is assumed that the overall value of the object is a quantitative value that comes from some distribution conditional on X_1. The cost of inspecting an offer in round i is defined to be c_i, $i = 1,2$.

The strategy of a searcher is defined by a vector of two variables (n, t), where n is the number of objects observed in a batch and t is the threshold used in the first round. Offer j, $1 \leq j \leq n$ passes through to round two if and only if $X_{1,j} \geq t$, where $X_{1,j}$ is the j-th observation of the variable X_1. If no offer from a batch exceeds this threshold, then the searcher observes another batch. If exactly one offer from a batch exceeds this threshold, then this offer is automatically accepted without incurring any search costs in the second round. If l offers from a batch pass through to round two, where $l > 1$, then these offers are compared and the appropriate search costs are incurred, i.e. $c_2 l$. Note that during the second round it suffices that the searcher remembers the best offer seen so far and compares this to each new offer. After observing all l offers, the searcher accepts the most highly ranked offer.

From the description of such strategies, a strategy with a batch size of one is a sequential search policy based purely on the first signal and defined by the appropriate threshold. Assuming that X_1 is a positive random variable, it follows that for $n > 1$, an individual following the strategy $(n, 0)$ will fully inspect all n objects in a batch and accept the best one seen. For these reasons, a strategy (n, t) will be called a filtering strategy when $n > 1$ and $t > 0$.

The goal of the searcher is to maximise the expected reward from search, which is assumed to be the value of the offer accepted minus the search costs incurred. Define $W(n, t)$ to be the random variable denoting the reward of a searcher using the strategy (n, t). Let $L \equiv L(n, t)$ be the number of offers that pass through to the second round. It follows that $L \sim \text{Bin}(n, p)$, where $p = P(X_1 > t)$. Conditioning on the number of offers that pass through to the second round, from the form of the strategy, we obtain:

$$E[W(n,t)|L = 0] = E[W(n,t)] - c_1 n \quad (1)$$

$$E[W(n,t)|L = 1] = E[V|X_1 > t] - c_1 n \quad (2)$$

$$E\left[\max_{1 \le j \le l}\{V_j|X_{1,j} > t\}\right] - c_1 n - c_2 l, \text{ for } 2 \le l \le n, \quad (3)$$

where V_j is the overall value of the j-th offer seen in round 2. It follows that the value function, $E[W(n,t)]$ may be derived using the adaptation of the law of total probability to expected values. A two-step procedure is used to find the optimal strategy. Firstly, the optimal threshold for a given batch size is found and then we optimize over the set of batch sizes. Note that we may obtain an estimate of how large the optimal batch size is by considering the expected number of items that should be seen under a policy based purely on the first signal. This is considered in the next section on the basis of an example. The optimal threshold for a given batch size is found using a program written in R using the *Optimx* package (see Kelley, 1999). Since the value function can have local optima, these functions were graphed in order to ensure that the global maximum was selected.

3 Example

It is assumed that the first signal, X_1, has an exponential distribution with mean 1 and the second signal, X_2, has an exponential distribution with mean μ, independently of the first signal. The value of an offer is given by $V = X_1 + X_2$. The value of the first signal is observed precisely, while the values of the offers can only be ranked. Since the standard deviation for an exponential distribution is equal to the mean, the proportion of the information about the value of an offer described by the value of the second signal (if it could be precisely observed) may be defined by $q = \frac{\mu}{1+\mu}$. The cost of observing the first signal is $c_1 = 0.01$ and the cost of closer inspection is $c_2 = 0.1\mu$. Hence, in the second round the cost of search relative to the amount of information gained is 10 times greater than in the first round.

In order to estimate the optimal batch size and the expected number of offers from a batch that pass through to round two when a filtering policy is optimal, we consider two auxiliary search problems. The first problem assumes that the searcher uses a sequential search policy based purely on observations of the first signal (i.e. uses a batch size of one). The second problem assumes that the searcher uses a policy of the form: compare k values of the second trait and choose the best one. From the linearity of both the search costs and the value of an offer as a function of X_1 and X_2, as well as the independence of these two variables, it is expected that the optimal batch size should be close to the expected number of offers seen under the optimal strategy in the first auxiliary problem and the expected number of offers passing through to the second round should be close to the optimal choice of k in the second problem.

In the first auxiliary problem, an offer should be accepted only when its value is greater than the future expected reward from search (ignoring any costs already incurred), i.e.:

$$\begin{aligned} t = E[\max\{X_1, t\}] - c_1 = E[X_1 | X_1 > t]P(X_1 > t) + tP(X_1 < t) - c_1 = \\ = (1+t)e^{-t} + t(1-e^{-t}) - c_1. \end{aligned} \tag{4}$$

The second line follows from the lack of memory property of the exponential distribution. Hence, the optimal threshold satisfies $e^{-t} = c_1 \Rightarrow t = \ln(c_1^{-1})$. Thus the probability of accepting an offer under the optimal policy is $P[X_1 > > \ln(c_1^{-1})] = c_1$ and the expected number of offers observed before one is accepted is c_1^{-1}.

In the second problem, the strategy is defined by the number of offers to be inspected. After observing these offers, the searcher chooses the best one. Note that it is better to observe k objects rather than just $k-1$ if and only if the expected increase in the value of the best object seen (based entirely on X_2) is at least as great as the inspection costs. The probability that the k-th offer is the best is $1/k$. Given that the k-th offer is the best, then the amount by which the value of this offer exceeds the value of the previously best offer has an exponential distribution with mean μ (again from the lack of memory property of the exponential distribution). It follows that the optimal choice of k is given by the largest value of k that satisfies $\frac{\mu}{k} \geq c_2 \Rightarrow k \leq \frac{\mu}{c_2}$.

Using these two auxiliary problems, under a strategy involving filtering, the size of a batch should be around 100 and the expected number of objects seen in the second round should be close to 10.

Now consider the search problem with two-step inspection. Under the filtering strategy (n, t), the probability of an offer passing through to the second round is $P(X_1 > t) = e^{-t}$. Hence, the number of offers observed in round 2, L, has a binomial distribution with parameters n and e^{-t}.

Let $X_{i,j}$ denote the numerical measure of the i-th signal for the j-th offer to be seen in round 2. Using the lack of memory property of the exponential distribution, the particular forms of Equations (2) and (3) giving the expected reward from search conditional on the number of offers observed in the second round are given by:

$$E[W(n,t)|L = 1] = 1 + t + \mu - c_1 n \tag{5}$$

$$E[W(n,t)|L = l] = t + \max_{1 \leq j \leq l}[X_{1,j} + X_{2,j}] - c_1 n - c_2 l, \text{ for } 2 \leq l \leq n. \tag{6}$$

From the form of Equation (6), it is clear that in order to solve this problem we should derive the distribution of $\max_{1 \leq j \leq l}[X_{1,j} + X_{2,j}]$. The density function of the value of an offer, $V = X_1 + X_2$, can be derived from the convolution of X_1 and X_2, i.e.:

$$f_V(v) = \int_0^v f_{X_1}(x) f_{X_2}(v - x) dx. \tag{7}$$

This gives:

$$f_V(v) = \begin{cases} \dfrac{\exp\left(-\frac{v}{\mu}\right) - \exp(-v)}{\mu - 1}, & \mu \neq 1 \\ ve^{-v}, & \mu = 1. \end{cases} \tag{8}$$

Since the values of the offers are independent, for $\mu \neq 1$ the distribution function of the maximum value of a set of l offers, U_l, is given by:

$$F_{U_l}(v) = P(U_l \leq v) = P(V \leq v)^l = \left[1 - \frac{\mu \exp\left(-\frac{v}{\mu}\right) - \exp(-v)}{\mu - 1}\right]^l. \tag{9}$$

Since U_l is a positive random variable, it follows that:

$$E(U_l) = \int_0^\infty [1 - F_{U_l}(v)] dv = \int_0^\infty 1 - \left[1 - \frac{\mu \exp\left(-\frac{v}{\mu}\right) - \exp(-v)}{\mu - 1}\right]^l dv. \tag{10}$$

Expanding the integrand in this expression, it follows that $E(U_l) = \sum_{k=1}^{l} I_k$, where:

$$I_k = (-1)^{k+1} \int_0^\infty \binom{l}{k} \left[\frac{\mu \exp\left(-\frac{v}{\mu}\right) - \exp(-v)}{\mu - 1} \right]^k dv. \tag{11}$$

These integrals can be derived by expanding in a similar manner. It follows that:

$$I_k = \frac{1}{(\mu - 1)^k} \sum_{i=0}^{k} (-1)^i \binom{k}{i} \frac{\mu^{k+1-i}}{k + i(\mu - 1)}. \tag{12}$$

Applying the analogue of the law of total probability to the expected reward from search, we have:

$$E[W(n,t)] = \sum_{l=0}^{n} E[W(n,t)|L = l]P(L = l). \tag{13}$$

Using Equations (1), (5) and (6) and rearranging, we obtain:

$$E[W(n,t)] = \frac{1}{1 - (1 - e^{-t})^n} [ne^{-t}(1 - e^{-t})^{n-1}(1 + t + \mu) + $$

$$+ \sum_{l=2}^{n} \left\{ (E[U_l] - c_2 l) \binom{n}{l} e^{-tl} (1 - e^{-t})^{n-l} \right\} - c_1 n. \tag{14}$$

For a given n, this function was optimised using the Optimx package in R and then the value function was globally maximised by allowing n to vary.

4 Results

The results are illustrated in Figures 1 to 6. Figures 1 to 3 depict how the form of the optimal policy changes as the proportion of information contained in the second signal changes (Figure 1 illustrates the optimal batch size, Figure 2 presents the corresponding optimal threshold and Figure 3 illustrates the expected number of objects inspected in the second round). When the proportion of information contained in the second signal is less than about 65%, then the optimal batch size is equal to one. This corresponds to the optimal sequential search policy based purely on the first signal. Figure 2 indicates that in this region a very high value of the first signal is necessary (and in this case sufficient) for an offer to be accepted. It might be surprising that the second signal must contain so much information in order to be considered, but this seems to result from the following three factors: a) the order in which the signals are observed promotes the importance of the first signal in the decision process, b) the first signal can be measured quantitatively, whereas the second signal is qualitative (at least in terms of assessment by the decision maker), c) the costs of inspecting the second signal are relatively more expensive.

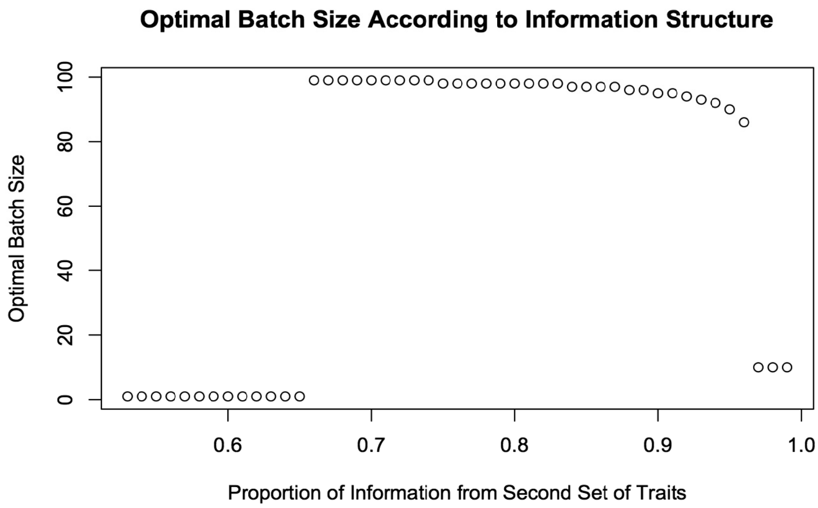

Figure 1. The optimal batch size according to the proportion of information contained in the second signal

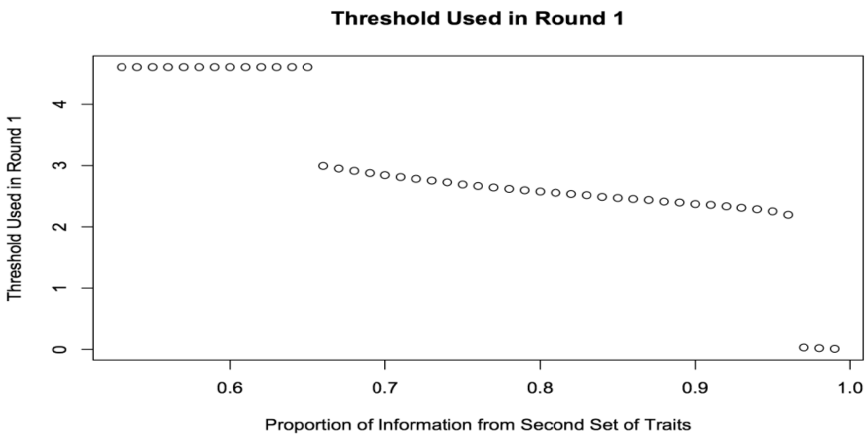

Figure 2. The optimal threshold (corresponding to the optimal batch size) according to the proportion of information contained in the second signal

When the proportion of information contained in the second signal is relatively high (between around 65% and 97%), then the batch size is close to 100. The threshold used in the first round is now quite high, rather than very high. This indicates that a filtering strategy is being used, such that the decision maker uses the first signal as a cheap way of restricting costly inspection only to offers that seem promising. As the second signal becomes more informative, in comparison to the first signal, the optimal batch size falls slowly. On the other hand, the expected number of offers passing through to round two steadily rises (see Figure 3), since the threshold used in the first round decreases. Hence, as one would expect, when the second signal becomes more informative, the searcher will place more importance on observing this signal, compared to observing the first signal. In all cases, the expected number of offers observed in round two is around five or more, which means that the probability of not accepting an offer from a batch is very small. Hence, the variance in the search costs incurred will be relatively small. This is important in the case of risk-averse searchers, who are willing to accept a lower expected reward than the optimal one, if the variance of this reward can be sufficiently reduced.

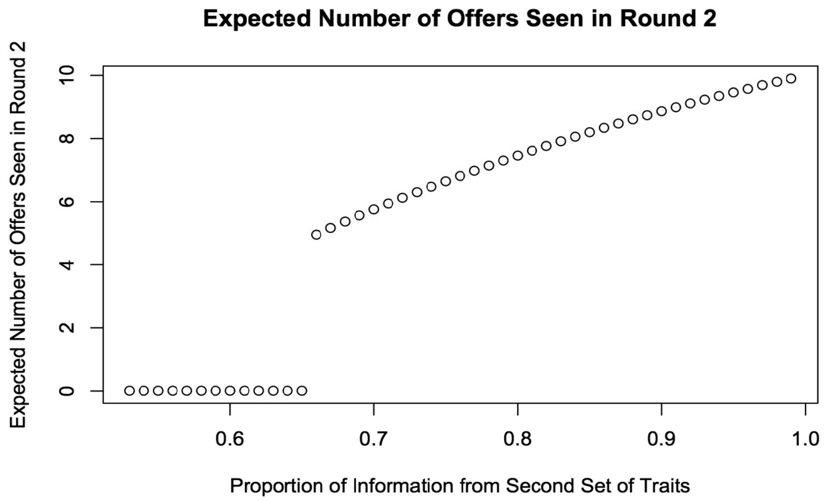

Figure 3. The expected number of offers seen in round 2 according to the proportion of information contained in the second signal

When the proportion of information contained in the second signal is very high (around 97% or more), then the optimal batch size is ten. The threshold used in round one is very close to zero, i.e. almost all the offers seen pass through to round two. This indicates that the first signal is of too little importance to be used as an initial filter to direct search. In this case, an optimally behaving searcher will essentially fully inspect the appropriate number of offers as indicated by the optimal strategy based purely on the second signal and accept the best offer seen.

Figure 4 illustrates the relative values of the expected rewards obtained under the three types of strategy described above: i) the optimal strategy based on filtering, ii) the optimal threshold strategy based purely on the first signal, iii) the strategy based on fully inspecting a batch of ten offers. The expected reward obtained under the optimal filtering policy is scaled to be equal to one. For the problem investigated here, filtering can lead to clear benefits when the second signal contains around 80% of the information about the value of an offer. In this case, the first signal contains enough information to direct search at relatively low cost, while the amount of information contained in the second signal ensures that it should be considered in decision making (even when no quantitative measure of the value of this signal is observable).

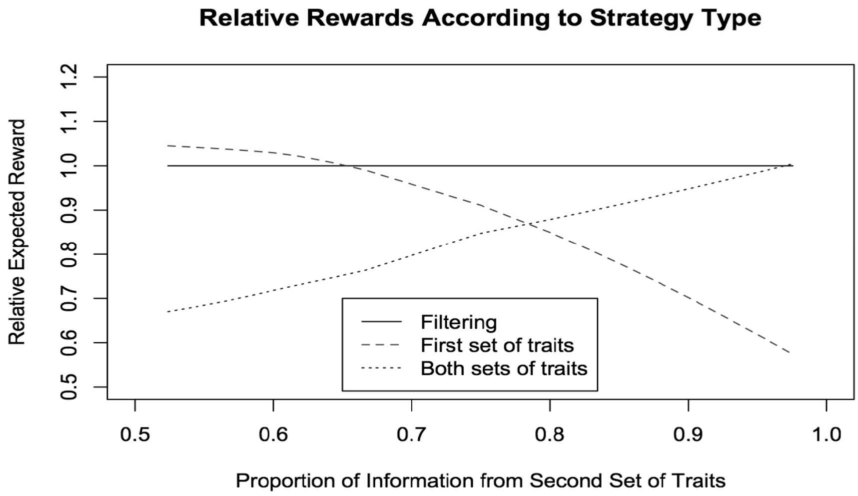

Figure 4. Relative rewards gained according to the type of strategy used and proportion of information contained in the second signal

Figure 5 illustrates how the choice of batch size affects the expected reward from search when the second signal contains 80% of the information. Note that the threshold used is optimally adapted to the batch size. As long as the chosen batch size is not radically different from the optimal batch size, then there is very little effect on the expected payoff. In fact, for the example considered, the expected reward obtained by using any batch size of between 60 and 149 and the corresponding optimal threshold is within 1% of the optimal expected reward. This is attained by using a batch size of 98.

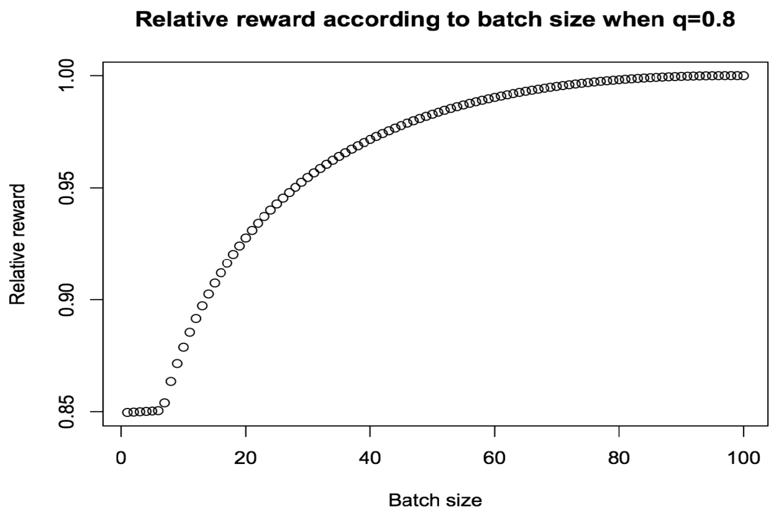

Figure 5. Effect of the batch size on the expected reward when 80% of the information is contained in the second signal

Figure 6 illustrates how the expected number of offers observed in the second round depends on the batch size used. It can be seen that as long as the batch size is not very small in relation to the expected number of offers seen under the optimal strategy based on the first signal (here 100), then the expected number of offers observed in the second round is almost constant. This suggests that a promising line of research would be to consider filtering strategies in which a fixed number of offers from a batch pass through to the second round. The author has some results for strategies of this type, but the solution of such problems can be combinatorically very complex. Hence, such strategies will be considered in a future paper.

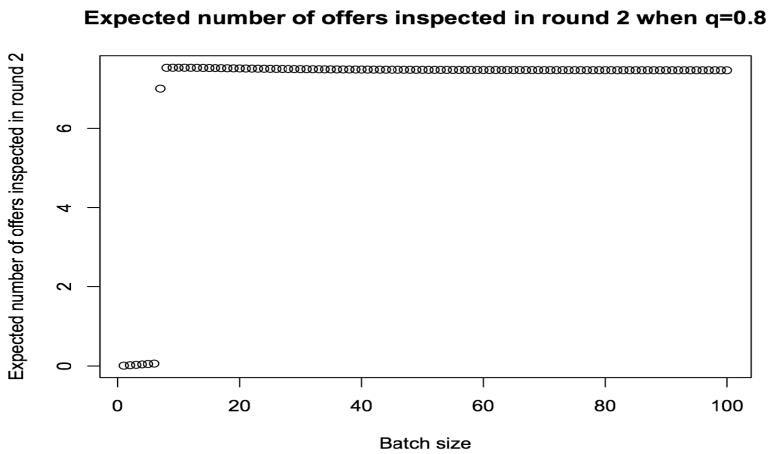

Figure 6. The expected number of offers seen in round 2 according to the batch size when 80% of the information is contained in the second signal

5 Conclusions

This paper has considered a model of searching for a valuable good in which additional information can be gained on chosen offers at some cost, before a final decision is made. In particular, this article has considered the concept of filtering strategies. These are strategies under which initial information on a number of offers is used to choose a smaller number of promising offers to investigate in more detail in the second stage. Such strategies may be beneficial when a number of offers can be observed in parallel and searchers do not have full information about the distribution of the value of offers and/or cannot precisely measure signals associated with the value of an offer. According to the model presented here, the initial filtering of offers is based on a quantitatively measurable initial signal that is cheap to observe. A set of offers which seem promising to the decision maker are then more closely observed on the basis of a second qualitative signal, such that the values of the offers can be ranked. The observation costs in the second round are relatively high compared to those incurred in the first round. These assumptions seem reasonable for the problem of searching for a flat using the Internet to gain initial information about offers before deciding which flats to visit. It is assumed that search costs are proportional to both the number of offers observed and the number of offers that are further investigated in the second round. On the basis of an example, it is shown that such a filtering process is optimal when the amount of information contained in the second signal is sufficiently high, but the first signal also contains some information. When the first signal contains very little information, the optimal strategy is very similar to a ”best-of-n” strategy, i.e. n offers are fully investigated and the best one is accepted. When the first signal contains a relatively large amount of information, then the optimal strategy is to use a threshold policy based on the first signal.

The assumption regarding the proportionality of search costs to the number of offers investigated is a simplification. As argued above, relaxing this assumption by considering the benefits of being able to observe offers in parallel would increase the benefits of using a filtering strategy. Another obvious adaptation of the model would be to assume that neither signal can be measured precisely, but offers can be compared at each stage. In this case, one should consider strategies under which the best k offers according to the first signal should be investigated in the second round. The author has some results regarding such a model. However, the solution of an example analogous to the one presented here is combinatorically complex and will be considered in future research. Another way of making the model more realistic would be to assume that the searcher does not choose the batch size, but this results from the rate at which new offers appear on the market and on the frequency with which an individual observes offers.

In terms of multiple criteria decision making, one interesting characteristic of filtering strategies is that they give very tight control over the search costs incurred while ensuring a valuable offer is obtained. It thus seems reasonable that filtering strategies give a high expected reward from search while keeping the variance of this reward low. This is an important feature, since individuals tend to be risk averse, and should be considered in future research.

One obvious practical problem with the model presented here is that the optimal solution depends on the distribution of the signals and the values of offers. The precise form of the optimal solution will depend on the form of these underlying distributions. In particular, the total variance of the value of an offer and the residual variance of the value of an offer given the first signal are very important in determining how many offers should be considered in each round. These variances depend on both the covariance matrix for the signals, and the relation of the value of an offer to the signals observed. One further avenue for future research would be to investigate heuristic methods of choosing an appropriate number of offers to compare based on the variation in the signals already seen. Also, it would be useful to investigate how robust such an approach would be to the precise form of the joint distribution of signals. For example, Connolly and Wholey (1988) note that this is a difficult problem to solve in practice, even when decision makers possess very good information about the distribution of signals and their relation to the value of an offer. However, it might be the case that an adaptation of the heuristic strategy proposed by Janetos (1980), i.e. ”choose five offers for further investigation and accept the best” might be a fairly robust strategy to implement in the second round of such a search procedure.

References

Analytis P.P., Kothiyal A., Katsikopoulos K. (2014), *Multi-attribute Utility Models as Cognitive Search Engines*, Judgment and Decision Making, 95, 403-419.

Analytis P.P., Stojic H., Moussaïd M. (2015), *The Collective Dynamics of Sequential Search in Markets for Cultural Products* [in:] Analytis P.P. (2015), *Psychological Process Models and Aggregate Behavior*, 81-124.

Baucells M., Carrasco J.A., Hogarth R.M. (2008), *Cumulative Dominance and Heuristic Performance in Binary Multiattribute Choice*, Operations Research, 56(5), 1289-1304.

Bearden J.N., Connolly T. (2007a), *Optimal Satisficing*, Management Science, 1(1), 1-14.

Bearden J.N., Connolly T. (2007b), *Multi-attribute Sequential Search*, Organizational Behavior and Human Decision Processes, 103(1), 147-158.

Chhabra M., Das S., Sarne D. (2014), *Expert-mediated Sequential Search*, European Journal of Operational Research, 234(3), 861-873.

Chun Y.H. (2015), *Multi-attribute Sequential Decision Problem with Optimizing and Satisficing Attributes*, European Journal of Operational Research, 243(1), 224-232.

Connolly T., Wholey D.R. (1988), *Information Mispurchase in Judgment Tasks: A Task-driven Causal Mechanism*, Organizational Behavior and Human Decision Processes, 42(1), 75-87.

Fawcett T.W., Johnstone R.A. (2003), *Optimal Assessment of Multiple Cues*, Proceedings of the Royal Society of London B: Biological Sciences, 270, 1637-1643.

Hogarth R.M., Karelaia N. (2005), *Simple Models for Multiattribute Choice with Many Alternatives: When It Does and Does Not Pay to Face Trade-offs with Binary Attributes*, Management Science, 51(12), 1860-1872.

Hutchinson J.M.C., Hałupka K. (2004), *Mate Choice When Males Are in Patches: Optimal Strategies and Good Rules of Thumb*, Journal of Theoretical Biology, 231(1), 129-151.

Janetos A.C. (1980), *Strategies of Female Mate Choice: A Theoretical Analysis*, Behavioral Ecology and Sociobiology, 7(2), 107-112.

Keeney R.L., Raiffa H. (1993), *Decisions with Multiple Objectives: Preferences and Value Trade-offs*, Cambridge University Press, Cambridge, UK.

Kelley C.T. (1999), *Iterative Methods for Optimization*, Society for Industrial and Applied Mathematics.

Lim C., Bearden J.N., Smith J.C. (2006), *Sequential Search with Multiattribute Options*, Decision Analysis, 3(1), 3-15.

MacQueen J.B. (1964), *Optimal Policies for a Class of Search and Evaluation Problems*, Management Science, 10(4), 746-759.

MacQueen J., Miller Jr R.G. (1960), *Optimal Persistence Policies*, Operations Research, 8(3), 362-380.

Ramsey D.M. (2015), *On a Sequential Decision Process Where Offers Are Described by Two Traits*, Multiple Criteria Decision Making, 10, 141-154.

Real L. (1990), *Search Theory and Mate Choice. I. Models of Single-sex Discrimination*, The American Naturalist, 136(3), 376-405.

Simon H.A. (1995), *A Behavioral Model of Rational Choice*, The Quarterly Journal of Economics, 69(1), 99-118.

Stigler G.J. (1961), *The Economics of Information*, The Journal of Political Economy, 69(3), 213-225.

Wiegmann D.D., Weinersmith K.L., Seubert S.M. (2010), *Multi-attribute Mate Choice Decisions and Uncertainty in the Decision Process: A Generalized Sequential Search Strategy*, Journal of Mathematical Biology, 60(4), 543-572.