Applied Intelligence (2022) 52:7529–7549
https://doi.org/10.1007/s10489-021-02856-8

A new algorithm for modeling online search behavior and studying ranking reliability variations

Debora Di Caprio[1] · Francisco J. Santos-Arteaga[2] · Madjid Tavana[3,4]

Accepted: 16 September 2021 / Published online: 30 September 2021

Abstract
We design an information retrieval algorithm that mimics the stochastic behavior of decision-makers (DMs) when evaluating the alternatives displayed by an online search engine. The algorithm consists of a decision tree that incorporates all the 1024 decision nodes that may arise from the information retrieval process of DMs. We calibrate the behavior of the algorithm to the one observed from online users and run several sets of 1,000,000 queries. Each query lets DMs decide which subset of the ten alternatives composing the initial page of results to click, allowing us to evaluate their behavior as ranking reliability is assumed to decrease when DMs decide not to click on an alternative. We compare the click-through rates (CTRs) obtained when modifying the degree of ranking reliability derived from the alternatives displayed on the first page of search results. We illustrate how the stability of the CTR prevails among the top-ranked alternatives within relatively reliable scenarios while it drops when imposing large initial decrements in reliability. The resulting consequences regarding the importance of relative ranking positions are analyzed, the top three alternatives exhibiting a generally contained decrease in their CTRs that contrasts with the cumulative pattern arising from the fourth position onwards.

Keywords Online search algorithm · Information retrieval · Ranking reliability · Decision tree · Click-through rate

1 Introduction and literature review

Decision-makers (DMs) trust the rankings provided by online search engines despite their ignorance about the selection process used to generate them [1–3]. The use of eye-tracking technology has allowed researchers to illustrate how DMs scan alternatives in the order displayed by the engine while focusing on the highest-ranked ones [4, 5]. Empirical analyses dealing with the behavior of DMs describe a disproportionate concentration of clicks on the first two alternatives, with values decreasing as DMs proceed further down the ranking [6, 7]. Seminal studies on the behavior of search engine users concluded that an average of two pages was clicked per query [8, 9].

In addition to the academic evidence provided above, institutional bodies such as the European Commission [10] have performed opinion polls analyzing the trust of their citizens in different information technology applications. Figures 1 and 2 present the results from one of these polls, highlighting the trust of European users in the rankings displayed by online platforms and the effect that these rankings have on their behavior.

Formal information retrieval models are constrained by the cognitive limits of DMs, who are unable to base their behavior on the 3,628,800 permutations derived from the ten initial results provided by the engine [11, 12]. Thus, it is generally assumed that DMs have limited information assimilation capabilities, requiring the introduction of heuristic mechanisms to evaluate the corresponding alternatives [13]. Figure 3 describes the average number of terms defined

✉ Madjid Tavana
tavana@lasalle.edu

Debora Di Caprio
debora.dicaprio@unitn.it

Francisco J. Santos-Arteaga
fsantosarteaga@unibz.it

1 Department of Economics and Management, University of Trento, Trento, Italy

2 Faculty of Economics and Management, Free University of Bolzano, Bolzano, Italy

3 Business Systems and Analytics Department, La Salle University, Philadelphia, PA 19141, USA

4 Business Information Systems Department, Faculty of Business Administration and Economics, University of Paderborn, Paderborn, Germany

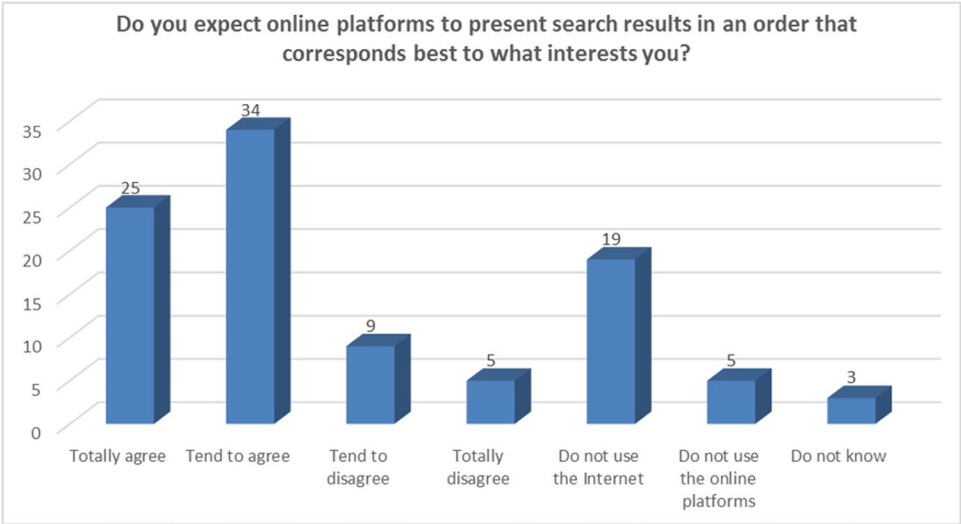


Fig. 1 Source: European Commission [10]. TNS Opinion, April 9, 2016, to April 18, 2016; 27,969 respondents; 15 years and older

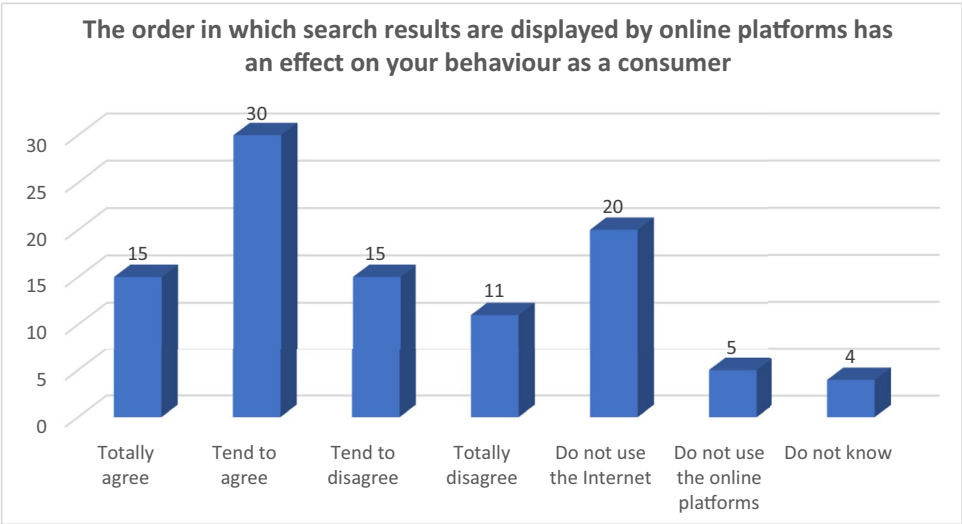


Fig. 2 Source: European Commission [10]. TNS Opinion, April 9, 2016 to April 18, 2016; 27,969 respondents; 15 years and older

per search query in the United States in 2017 and 2020. The queries performed by DMs are not particularly complex, though their complexity has increased over time from one to two words, with a three-word upper limit constraining most searches in both periods.

Therefore, the only requirement that should be imposed on the evaluation process of DMs is their trust in the ranking delivered by the engine, a constraint conditioning potential policy implications [10]. This intuition contrasts with the approach followed by the systems literature, which generally proceeds in the exact opposite way. That is, most papers focus on instances dealing with a particular information retrieval problem and then compare the results obtained with those available from online search engines [14–16]. A similar drawback can be associated with one of the main research branches dealing with applied online information retrieval processes, namely, electronic commerce [17–19].

While the approaches defined in the literature are obviously valid and extremely helpful, they are subject to the observer bias inherent to most experimental frameworks. More precisely, users are aware of the fact that their behavior is being monitored, which conditions their actual behavior. We bypass this potential problem by following a simulation approach to the analysis of online search behavior [20, 21]. In particular, we focus on the sequential behavior of users directly retrieved from their online searches as the objective to stimulate and mimic. The structure of the resulting algorithm can then be modified according to the phenomena being analyzed, which constitutes an important requirement in the systems literature [22, 23].

We now describe the results from different empirical studies that analyze the information retrieval behavior of users—and the subsequent click-through rates (CTRs)—when performing online searches. The results obtained present very

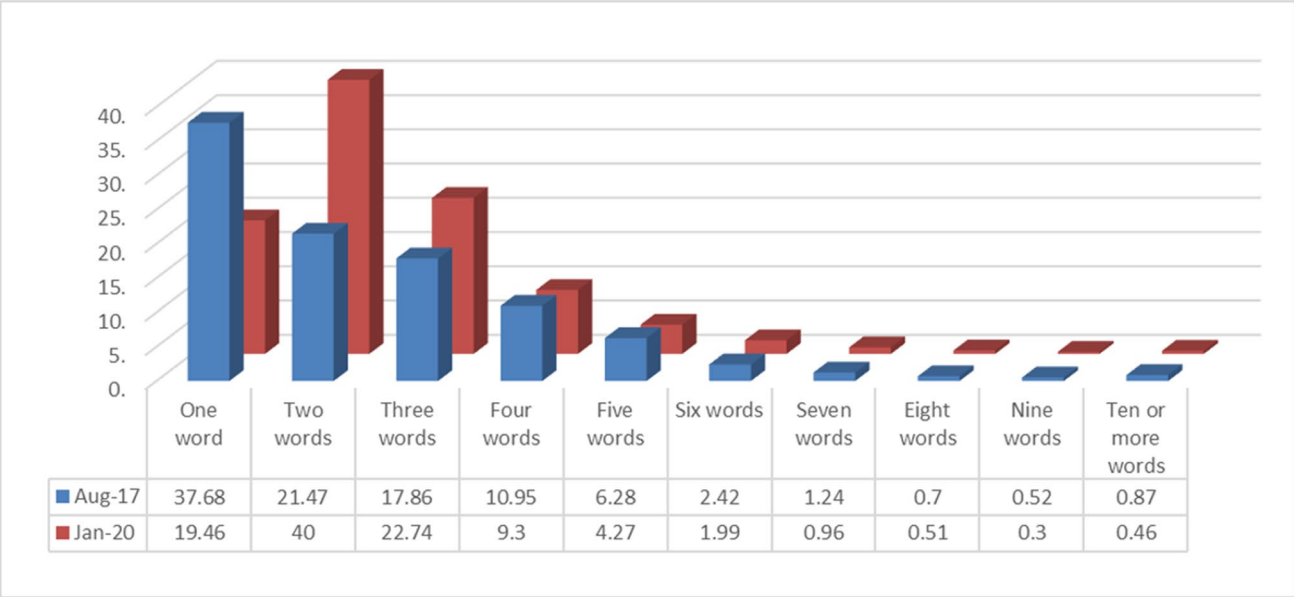


Fig. 3 The average number of terms defining online search queries in the United States: August 2017 vs. January 2020. Data retrieved from Statista (https://www.statista.com/statistics/269740/number-of-search-terms-in-internet-research-in-the-us/). Original source: keyworddiscovery.com.

Table 1 Average traffic shares computed by Chitika [6] and Dean [7]

Google Result Page Rank	Chitika [6] (%)	Dean [7] (%)
1	32.5	26.96
2	17.6	21.00
3	11.4	15.90
4	8.1	11.56
5	6.1	8.08
6	4.4	5.27
7	3.5	3.49
8	3.1	2.64
9	2.6	2.55
10	2.4	2.55
11	1.0	–
12	0.8	–
13	0.7	–
14	0.6	–
15	0.4	–

similar CTRs profiles—concentrated on the alternatives displayed within the first page of search results—that remain consistent through time.

Two main empirical studies dealing with average traffic shares and CTRs, performed by Chitika [6] and Dean [7], respectively, describe similar search patterns displayed by online users. The results are summarized in Table 1. Note that we have converted the CTRs presented by Dean [7] into average traffic shares to compare both sets of results. Besides their similarities, the results described by Chitika [6] have been included to highlight the early emphasis placed on the initial page of results provided by search engines.

In addition, online platforms such as the Advanced Web Ranking website (https://www.advancedwebranking.com/ctrstudy/) provide monthly updated values of the CTRs generated by users, allowing to analyze the consistency of their search behavior through time. Table 2 presents the annual organic click-through rates for international desktop searches performed since 2015, while also including the most recent results corresponding to April 2021. As can be observed, the trends displayed by online users remain consistent across time while exhibiting similar retrieval patters to those described by Chitika [6] and Dean [7]. It should be highlighted that the trends observed remain consistent when considering particular countries such as Australia, UK and USA.

The consistent focus of DMs on the first three alternatives composing the ranking and its prevalence through time are evident throughout the whole set of data. We conclude by noting that the small percentage differences across time or countries can be easily accounted for by adapting the probabilities of our model so as to incorporate modifications in any of these values.

Table 2 Annual organic CTRs from international desktop searches

Google Result Page Rank	2015	2016	2017	2018	2019	2020	April 2021
1	33.62	33.62	35.24	35.28	34.04	34.7	34.6
2	18.33	17.45	17.81	17.13	16.91	16.77	16.36
3	11.36	10.84	10.81	10.57	10.2	10.15	9.71
4	7.94	7.62	7.42	7.05	6.78	6.83	6.43
5	5.74	5.52	5.34	5.01	4.8	4.87	4.49
6	4.25	4.11	3.99	3.69	3.54	3.6	3.27
7	3.28	3.15	3.07	2.82	2.69	2.74	2.46
8	2.62	2.48	2.41	2.21	2.11	2.16	1.92
9	2.09	1.96	1.92	1.78	1.7	1.74	1.53
10	1.7	1.59	1.61	1.54	1.43	1.46	1.29
11	–	1.4	1.57	1.71	1.36	1.35	1.17
12	–	1.38	1.66	1.94	1.58	1.45	1.2
13	–	1.34	1.66	1.89	1.65	1.5	1.22
14	–	1.26	1.59	1.78	1.54	1.46	1.22
15	–	1.14	1.52	1.69	1.42	1.39	1.2
16	–	1.07	1.43	1.62	1.32	1.3	1.12
17	–	0.96	1.32	1.54	1.2	1.22	1.05
18	–	0.86	1.21	1.45	1.12	1.14	0.97
19	–	0.78	1.12	1.36	1.03	1.06	0.9
20	–	0.68	1.01	1.29	0.95	0.97	0.82

Source https://www.advancedwebranking.com/ctrstudy/

2 Contribution

In the current paper, we mimic the CTR behavior of DMs when evaluating the alternatives displayed by a search engine. We define a sequential information retrieval algorithm describing the incentives of a DM to click on the alternatives ranked on the first page of results after running a search query. The only assumption that will be imposed to define the evaluation process is the trust placed by DMs in the order of the items displayed by the engine. The probability of clicking on a given alternative through the decision-tree structure of the algorithm is then equated to the CTR exhibited by online users and sets of 1,000,000 queries are simulated and validated according to these values.

After validating the behavior of the algorithm, we illustrate the effects from a decrease in the reliability of the ranking as the results provided by the engine do not match the characteristics expected by DMs [24, 25]. That is, if after a quick assessment the DM decides not to click on a given alternative, he will experience a loss of confidence in the ranking and a subsequent decrease in the probability of clicking on the remaining alternatives as he proceeds down the ranking [26, 27]. The resulting effect on the intention of DMs to click and evaluate the different alternatives given their relative ranking positions will be analyzed.

To the best of our knowledge, the decision-tree algorithmic framework presented is completely novel, while being solely based on the sequential behavior displayed by online users. In this regard, the algorithm can be easily extended to incorporate different behavioral constraints determined by the specific information retrieval phenomenon being studied. All in all, the main objective of the paper is to define an algorithmic benchmark upon which to validate and extrapolate the potential behavior of users through different information retrieval scenarios. In addition, when formalizing the behavior of DMs, we will define a technical model describing the interactions that arise as DMs retrieve information to evaluate the characteristics of the alternatives displayed by a search engine.

The following key points summarize and highlight the main contributions of the algorithm introduced through the paper:

1. The design of the algorithm is based on a decision model that mimics the sequential search behavior of DMs when browsing through a set of ranked alternatives. DMs read the snippets, interpret them, and decide on which ones to click based on how close the characteristics of the alternatives are to their preferred ones.
2. The algorithm defines each step of the information retrieval process and the subsequent evaluation paths that DMs may follow as they proceed through the ranking delivered by the engine. This feature allows DMs to go back to previously evaluated alternatives after pro-

ceeding further through the ranking or going forward and backward to retrieve information before deciding what links to click.

3. One of the main advantages of our approach is that entire retrieval processes can be generated according to the subjective behavioral characteristic of users. In the current paper, we calibrate the algorithm so that the probabilities defining the sequential decisions are determined by the empirical values of the corresponding CTRs.
4. For comparative purposes, we define basic algorithmic structures designed to mimic the CTR behavior observed empirically but not to account for the interactions across alternatives arising through the different search paths that DMs may generate as they retrieve information.
5. Frictions have been introduced to illustrate the interrelated sequential decision structure of the algorithm as the effect from behavioral modifications is carried over through the different paths generated. This type of pattern cannot be analyzed by algorithms that do not incorporate the interactions across alternatives arising through DMs' sequential information retrieval processes.

The rest of the paper proceeds as follows. Section 3 highlights the main differences between the standard approach of computer scientists to formalize information retrieval processes and the decision theoretical one implemented in the current paper. Section 4 presents a formal decision theoretical model that provides intuition on the sequential evaluation behavior of DMs leading to the actual CTRs observed. Section 5 summarizes the basic applications of decision trees to formalize sequential decision processes. Section 6 illustrates in detail the main properties of the decision-tree structure defining the information retrieval process of DMs and the algorithmic framework designed to incorporate them. Section 7 describes the information retrieval algorithm, whose behavior is simulated and validated empirically in Sect. 8. Section 9 presents the main managerial implications derived from the implementation of the benchmark algorithm. Section 10 suggests potential extensions and concludes.

3 Computer science versus decision theory

The approach implemented in the current paper to formalize the behavior of DMs differs from those commonly applied in the literature on computer science and artificial intelligence. The models considered by these research areas tend to focus on applications of deep and machine learning techniques to extrapolate the behavior of DMs given some predetermined characteristics selected from their online purchases. The CTRs are estimated by applying an artificial intelligence technique to categorize DMs based on observed behavior and the subsequent categories assigned.

More precisely, researchers focus on extracting particular qualities of the users from the data so as to enhance the categorization capacity of the corresponding deep-learning techniques when predicting CTRs. For instance, a recent tendency within the literature consists of incorporating into the analysis the interest of users estimated from their behavioral data [28, 29]. Moreover, computer scientists have recently emphasized the sequential quality of the retrieval process, namely, the importance of the CTR behavior displayed by users at different points in time as a major categorization feature [30].

We develop a model built on decision theory and expected utility postulates. In particular, we do not categorize DMs based on their observed characteristics but simulate their actual sequential decision processes. By doing so, we are able to design an information retrieval algorithm that mimics the observed behavior of DMs. The algorithm is calibrated so that the probabilities determining the sequential decisions reflect the actual behavior of DMs when deciding which alternatives to click. That is, we are generating artificial DMs whose searches give place to the information retrieval data used by computer scientists to categorize their behavior via neural networks.

The significance and effectiveness of the proposed approach are described below.

1. The decision theoretical model relates the cutoff values conditioning the evaluation behavior of DMs to the similarities existing between the linguistic terms observed in the snippets and the ideal ones defined by the user.
2. In particular, the cutoff values are determined by the utility that may be obtained from retrieving information on a given alternative relative to the potential utility derived from acquiring information on the next alternative composing the ranking.
3. The flexibility of the model and the subsequent algorithm allows to modify the subjective behavior of DMs at any exact point through the retrieval process and evaluate the resulting consequences in terms of CTRs.
4. The model can incorporate signals or any strategic incentive into the behavioral profiles of different types of DMs and describe how they should be expected to behave given their modified characteristics.

Given these features, the output generated by the algorithm constitutes an important addition to the analysis of information retrieval processes. In particular, we must emphasize that the algorithm provides two different strings of data.

On the one hand, it describes the pages clicked by the DMs.
On the other hand, and more importantly, the algorithm provides a numerical representation for each of the evaluations determining the retrieval behavior of DMs. As stated above, this information is not generally available, with most empirical studies aiming to extrapolate the preferences and evaluations of DMs from the retrieval data observed.

Finally, one of the main advantages of our approach is the fact that we can generate entire retrieval processes based on the subjective behavioral characteristic of users. In this regard, we are able to simulate the potential consequences derived from any modification affecting the main postulates that define the retrieval behavior of users. This result allows for the implementation of machine learning techniques to categorize potential information retrieval processes using the results derived from the algorithm as a benchmark. This quality is particularly useful when relating the cutoff values defined through the algorithm to the actual characteristics of DMs. The following section presents a simple decision theoretical model that incorporates the main behavioral attributes of DMs into the definition of these cutoff values.

4 A simple sequential evaluation model based on the characteristics of the alternatives

Given the evidence presented by the empirical literature on the information retrieval behavior of users, we will focus our analysis on the ten alternatives composing the initial page of results delivered by search engines. The current section introduces a simple decision theoretical model that provides intuition on the behavior of DMs leading to the actual CTRs observed.

The model is simple enough to define a sequential framework where rational DMs lacking information on the characteristics defining the alternatives must assess the descriptions provided by the search engine before deciding on which links to click. That is, given the information presented in the snippets, the DMs must decide whether to click on an alternative so as to gather additional information or proceed with the next alternative composing the ranking.

At the same time, the model illustrates the complexity inherent to any basic decision framework, with DMs having to account for the realizations of the different characteristics and their subsequent expected values, together with the evaluation results obtained from the ten potential alternatives that must be considered before making a final decision.

4.1 Formalization and technical assumptions

Assume that the alternatives are composed by two main sets of characteristics that take values in two nonempty sets, X_1 and X_2. The alternatives evaluated by DMs are therefore characterized by a pair $(x_1,\ x_2) \in X_1 \times X_2$. More precisely, the alternatives are composed by characteristics that can be directly observed and evaluated by reading the snippets, X_1, and those that require clicking on the corresponding link to be analyzed and evaluated, X_2. The intuition on which such an analytical framework is based follows from the search and experience attributes defined by Nelson [31] when evaluating alternatives.

A *preference relation* $\succeq$ on X_i, $i = 1, 2$, is a reflexive, complete, and transitive binary relation on X_i. A utility function $u_i : X_i \to \mathrm{R}$ *representing* $\succeq$ *on* X_i satisfies the following order-preserving condition, $\forall x', x'' \in X_i,\ x' \succeq x'' \Leftrightarrow u_i(x') \geq u_i(x'')$.

X_1 and X_2 will be identified with a closed real subinterval of $[0, +\infty)$, $X_i = [x_i^m,\ x_i^M]$, for $i = 1, 2$, such that $0 < x_i^m < x_i^M$. We endow X_i with the standard Euclidean topology and assume that the standard linear order $<$ constitutes the preference relation defined on X_i. Both features allow us to assume strictly increasing and continuous u_i functions. As a result, the function $u : X_1 \times X_2 \to \mathrm{R}$ defined by $u(x_1, x_2) = u_1(x_1) + u_2(x_2),\ \forall (x_1, x_2) \in X_1 \times X_2$, is increasing. Moreover, $u(x_1, x_2)$ induces and represents an additive preference relation on $X_1 \times X_2$.

Abusing notation, we will interpret X_i as a continuous random variable with associated probability densities $f_i : X_i \to [0,\ 1]$. Intuitively, $f_i(x_i)$ describes the subjective probability assigned to the ith characteristic of a randomly evaluated alternative being $x_i \in X_i$. We will assume that the densities are independent, though this assumption can be modified to incorporate correlated characteristics to the analysis.

4.2 Expected evaluation utilities

In this section, we introduce two functions describing the expected utilities obtained from the information retrieval processes implemented by DMs. Consider the behavior of a DM when evaluating an alternative denoted by J. After observing the value of the initial characteristics of J, the DM must decide between either clicking on the link and continuing the evaluation of J, or proceeding with the next alternative composing the ranking. The corresponding decision is determined by the value $x_1 \in X_1$ observed for J.

An important remark follows. In order to simplify the presentation, we study the case where the alternatives are defined by a unique characteristic within both X_1 and X_2.

Considering a unique $x_i \in X_i$, for $i = 1, 2$, implies that we will not analyze potential developments of the model incorporating, for instance, the convolution of probability distributions, but assume that the DM assigns a utility to the alternatives based on the relative distance of each characteristic from the best potential evaluation that could be observed. We elaborate on the consequences from this assumption at the end of the current section.

The first evaluation function describes the utility derived from the characteristics observed in the snippets, x_1, relative to their best potential values, x_1^M, together with the set of relative potential realizations of the second characteristic, $x_2 \in [x_2^m,\ x_2^M]$, that may be observed when acquiring additional information on the alternative

$$U(x_1, x_2) \stackrel{def}{=} u_1(x_1) + \int_{x_2 \in X_2} \mu_2(x_2)\big(u_2(x_2)\big)\, dx_2 \tag{1}$$

with

$$u_1(x_1) = \frac{x_1 - x_1^m}{x_1^M - x_1^m}, \quad x_1 \in X_1 \tag{2}$$

$$E_2 = \int_{x_2 \in X_2} \mu_2(x_2)\big(u_2(x_2)\big)\, dx_2 = \int_{x_2 \in X_2} f(x_2)\left(\frac{x_2 - x_2^m}{x_2^M - x_2^m}\right) dx_2, \quad x_2 \in X_2 \tag{3}$$

Note that we have simplified the analysis by defining linear utilities to formalize the evaluation of the relative distances existing within characteristics. This assumption can be relaxed to account for non-linear relationships determined by the relative distances, which would require defining the corresponding certainty equivalent values, $ce_i = u_i^{-1}\big(E_i\big)$, instead of the expected utility ones, E_i, for $i = 1, 2$, when incorporating a reference alternative to the analysis.

Given the lack of information regarding the distribution of characteristics, we will assume that the DM assigns a uniform distribution to its potential realizations within the interval $[x_i^m,\ x_i^M]$, for $i = 1, 2$. Uniform distributions account for the highest information entropy that follows from the complete uncertainty faced by the DM regarding the set of potential realizations that may be observed [32].

The utility received by the DM when evaluating the first characteristic from a given alternative is therefore given by

$$U(x_1, x_2) \stackrel{def}{=} \left(\frac{x_1 - x_1^m}{x_1^M - x_1^m}\right) + \int_{x_2 \in X_2} f(x_2)\left(\frac{x_2 - x_2^m}{x_2^M - x_2^m}\right) dx_2 \tag{4}$$

On the other hand, the expected utility derived from randomly evaluating any of the alternatives composing the ranking is defined as follows

$$C(x_1, x_2, \varphi) \stackrel{def}{=} \int_{x_1 \in X_1} \mu_1(x_1)\big(u_1(x_1)\big)\, dx_1 + \varphi \int_{x_2 \in X_2} \mu_2(x_2)\big(u_2(x_2)\big)\, dx_2 \tag{5}$$

with $\varphi \in [1, 2]$ representing the compensation, in utility terms, required by the DMs for the uncertainty faced when evaluating an alternative whose second set of characteristics is unknown at the time the decision is being made [33]. Given the trust exhibited by DMs regarding the rankings provided by search engines, it can be assumed that the value of φ increases as DMs proceed through the alternatives composing ranking.

Substituting the values of the corresponding expressions within Eq. (5) we obtain

$$\begin{aligned} C(x_1, x_2, \varphi) \stackrel{def}{=} & \int_{x_1 \in X_1} f(x_1)\left(\frac{x_1 - x_1^m}{x_1^M - x_1^m}\right) dx_1 \\ & + \varphi \int_{x_2 \in X_2} f(x_2)\left(\frac{x_2 - x_2^m}{x_2^M - x_2^m}\right) dx_2 \end{aligned} \tag{6}$$

For illustrative purposes, consider a scenario where the characteristics defining the alternatives are uniformly distributed within the interval [0, 10]. In this case, the normalized expected values of both X_1 and X_2 equal 0.5. Clearly, $U(x_1, x_2) \in [0.5, 1.5]$ and $C(x_1, x_2, \varphi) \in [1, 1.5]$, with the value of $\varphi \in [1, 2]$ determined by the trust placed on the alternative evaluated being able to satisfy the subjective preferences of the DM based on its relative ranking position.

The model has been designed to compute the value of $\varphi \in [1, 2]$ such that $U(x_1, x_2)$ and $C(x_1, x_2, \varphi)$ cross at the value of x_1 required to generate the CTRs observed for each alternative based on their relative ranking positions. In this regard, the expression $(2 - \varphi)$ describes the degree of trust placed by DMs on the alternative being evaluated as a function of its relative ranking position. In other words, the cut-off values determining the information retrieval behavior of DMs through the different algorithmic structures defined in the paper correspond to the x_1 generated by the subjective φ assigned by DMs to each alternative based on its ranking position.

Consider, for instance, the CTRs provided by Dean [7] and presented in the second column of Table 5. The value of φ obtained when $x_1 = 0.68$ is equal to 1.36. Similarly, a value of $x_1 = 0.75$ implies $\varphi = 1.5$, while $x_1 = 0.97$ requires $\varphi = 1.94$. Figure 4 illustrates the limit evaluation thresholds defined by $U(x_1, x_2)$ and $C(x_1, x_2, \varphi)$, which are based on the values of x_1 determined by φ as DMs proceed through the ranking.

We must highlight that if we were to introduce the convolution of several random variables by assuming that multiple characteristics are included within both X_1 and X_2 that the

main retrieval setting and analyses would remain qualitatively unchanged, only quantitative modifications regarding the value φ would arise. We are, however, not trying to develop an exhaustive model describing how the acceptance probabilities are generated but simply assuming that the similar behavior observed in the different analyses of CTRs allows for a direct extrapolation into a formal setting. However, even within a simplified decision setting such as the one presented in this section, we must assume that DMs account for the effects of multiple variables determining their information retrieval behavior within each decision node.

5 Decision trees and information retrieval

Decision trees constitute a standard instrument of analysis within the operational research literature [34, 35]. Their relatively simple structure and malleability make them extremely useful when formalizing sequential decision processes, evolving into the random forest techniques currently applied in machine learning environments [36, 37]. The design of decision trees specific to the online information retrieval behavior of DMs is surprisingly sparse. That is, despite the empirical evidence illustrating that DMs follow a decision-tree-like evaluation structure when acquiring information online, researchers have not aimed at its structural formalization.

The apparent simplicity of the task may be a reason for the prevalence of this research gap. DMs could be assumed to follow a decision-tree retrieval process easily formalized given the heuristic behavior generally assumed on bounded rational agents with a limited capacity to observe and assimilate information [38, 39]. Indeed, absent cognitive frictions, the retrieval formalization may seem straightforward from a computational viewpoint, requiring the simulation of ten independent random variables, each of them assigned a different cutoff value. This type of simulation would produce processes consisting of ten potential evaluations—from a tree composed of 21 nodes—absent any relationship or interactions among the alternatives. We could only claim the mimicking quality of the model while being unable to illustrate the effects following any evaluation decision made through the sequential retrieval process.

In other words, the empirical analyses of the online search behavior of users, such as those of Dean [7] and the Advanced Web Ranking website, describe the CTRs obtained from a given number of search queries, which are defined as follows

$$\text{CTR of alternative } i = \frac{\textit{Number of users clicking on the link to alternative } i}{\textit{Number of users performing a search}}, \quad i = 1, \ldots, 10. \tag{7}$$

Given the fact that each run of the algorithm can be naturally considered to equate a search query by a DM, Eq. (7) implies that a simple algorithm such as the one presented in Fig. 5 suffices to generate the CTRs observed in real-life environments. Clearly, the algorithm assumes that DMs consider evaluating all ten alternatives based on a predetermined set of probabilities while omitting the potential interactions arising from the information retrieved. We elaborate on this important feature and its consequences regarding formalizing information retrieval processes throughout the next section.

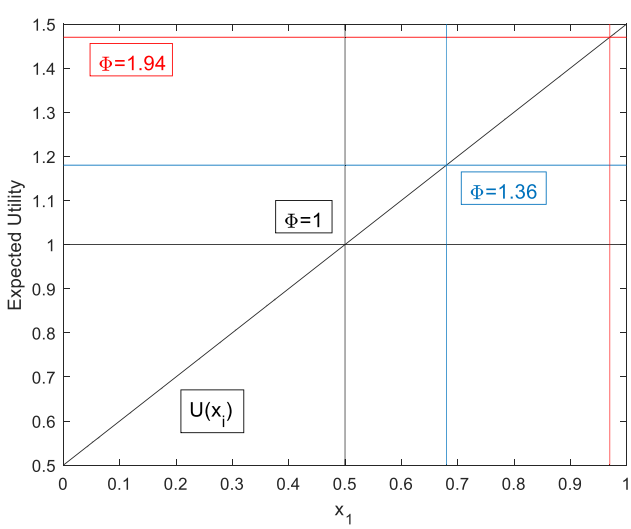


Fig. 4 Limit evaluation thresholds defined for different φ requirements determining x_1

6 Mimicking algorithms and interactions across alternatives

We proceed with an important remark involving the results generated by the main algorithms analyzed through the paper. The simplest algorithm, where DMs are assumed to perform ten independent searches—with a predetermined cutoff value defined for each alternative –, and the complete framework where DMs aim at observing ten satisfying alternatives (out of a total of ten available ones) both lead to the same stochastic structure and deliver the same CTRs. That is, both algorithms are capable of generating a set of CTRs that mimics the actual behavior observed among online users.

The intuition for this result is that aiming to evaluate the whole set of alternatives composing the ranking implies that the probability of considering each one of them equals one. This fact suggests that the complete algorithm generates the same stochastic evaluation framework as the one

```
nrows = 20;
ncols = 1000000;
A = zeros(nrows,ncols);
for i = 1
   for j = 1:ncols
        A(i,j)= rand(1);
        if A(i,j)>0.68
        A(11,j)=1;
        end
          A(i+1,j)= rand(1);
          if A(i+1,j)>0.75
          A(12,j)=2;
          end
            A(i+2,j)= rand(1);
            if A(i+2,j)>0.81
            A(13,j)=3;
            end
              A(i+3,j)= rand(1);
              if A(i+3,j)>0.86
              A(14,j)=4;
              end
                A(i+4,j)= rand(1);
                if A(i+4,j)>0.90
                A(15,j)=5;
                end
                  A(i+5,j)= rand(1);
                  if A(i+5,j)>0.94
                  A(16,j)=6;
                  end
                     A(i+6,j)= rand(1);
                     if A(i+6,j)>0.96
                     A(17,j)=7;
                     end
                         A(i+7,j)= rand(1);
                         if A(i+7,j)>0.97
                         A(18,j)=8;
                         end
                            A(i+8,j)= rand(1);
                            if A(i+8,j)>0.97
                            A(19,j)=9;
                            end
                               A(i+9,j)= rand(1);
                               if A(i+9,j)>0.97
                               A(20,j)=10;
                               end
    end
end
```

Fig. 5 Code of the simple algorithm delivering the CTRs observed empirically but lacking interactions across alternatives and evaluations

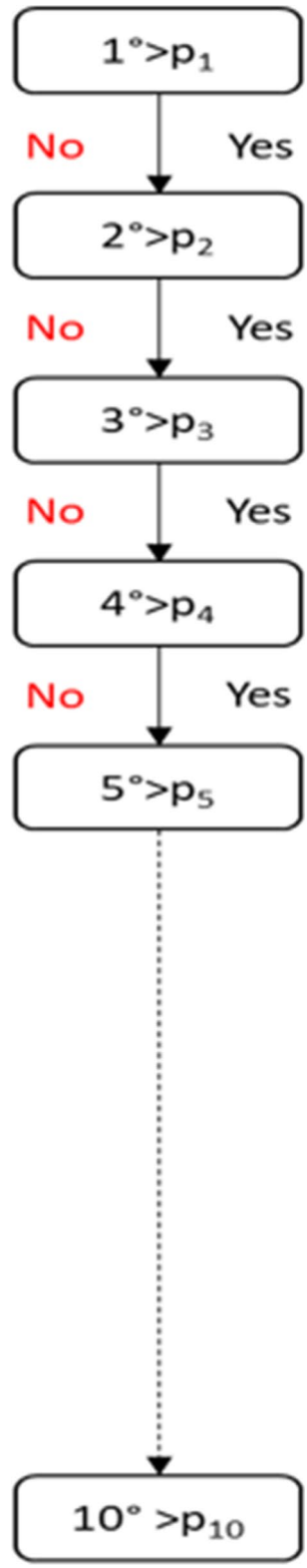


Fig. 6 Structure of the simple algorithm delivering the CTRs observed empirically but lacking interactions across alternatives and evaluations

defined by the simplest algorithm, where 10 alternatives are independently considered with probability one, each clicking decision determined by a predetermined cutoff value. However, in the latter case, the DM cannot condition his retrieval behavior on the evaluations of the alternatives observed previously.

Figure 6 describes the algorithmic structure of the simplest retrieval process that may be defined to replicate the CTR behavior observed. This basic model describes realizations from ten independent random variables, each alternative assigned a cutoff value, $p_i, i = 1, ..., 10$, determining the behavior of the DM. A sufficiently large number of realizations would reflect the exact value of the cutoff, as stated by the law of large numbers. However, as can be intuitively understood, the retrieval behavior determining the structure of the simplest and complete algorithms is substantially different. The latter scenario is closer to an actual sequential evaluation process.

Table 3 CTRs derived from the simple mimicking algorithm lacking interactions across alternatives and evaluations

CTR	Simple Mimic
Avg	1.1893
1	0.3194
2	0.2503
3	0.1901
4	0.1396
5	0.0998
6	0.0599
7	0.0400
8	0.0298
9	0.0302

```
nrows = 20;
ncols = 1000000;
A = zeros(nrows,ncols);
for i = 1
   for j = 1:ncols
        A(i,j)=rand(1);
        if A(i,j)>0.5
        A(11,j)=1;
        end
    end
end
```

Fig. 8 Code for the simple algorithm when DMs observe only one alternative

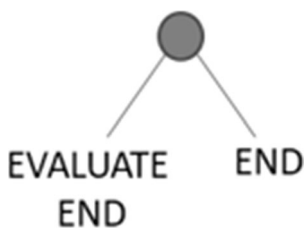


Fig. 7 Decision tree when DMs observe one alternative

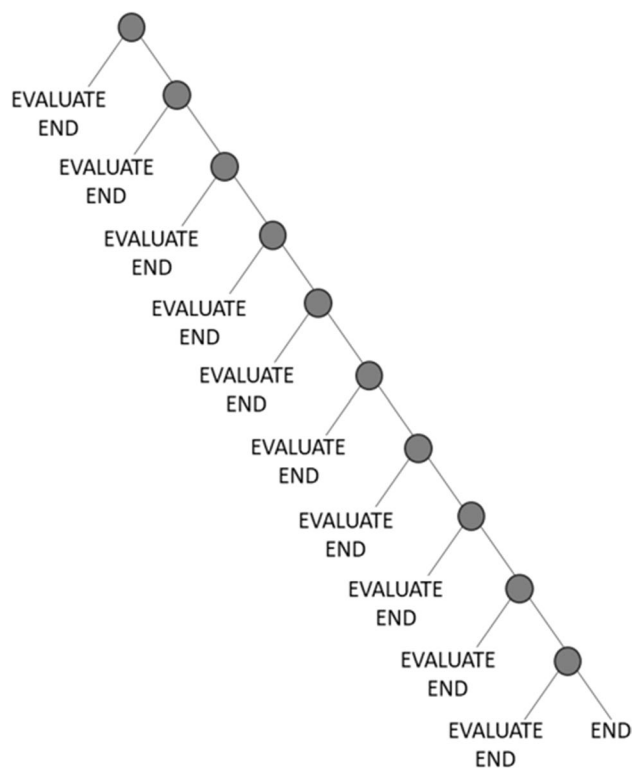


Fig. 9 Decision tree when DMs aims at finding one satisfying alternative out of ten

The CTRs obtained and the ranking positions of the alternatives being evaluated are reported in Table 3. The similarities with the complete algorithm introduced in the current paper will become evident in Sect. 8. Thus, while generating identical CTRs to those observed in real-life evaluation scenarios is an almost trivial problem, incorporating the set of potential interactions across variables requires the design and implementation of a much more complex algorithmic structure.

6.1 Formalizing the set of potential paths

We provide additional intuition illustrating the main differences between both types of algorithms by describing the case where DMs must either consider evaluating one alternative or find a satisfying alternative within the first ten provided by a search engine. A significant conceptual difference exists between the basic setting described in Figs. 7 and 8 and the approach defined in the current paper. DMs consider evaluating every potential alternative until they find one that satisfies their subjective preferences. In the former case, DMs are constrained to consider a unique alternative since assuming that the alternative evaluated is not necessarily the first one observed requires defining the whole set of paths that may be followed by the DM depending on the realizations observed. Figures 9 and 10 present the decision tree and the corresponding code describing the complete evaluation structure for a cutoff value of 0.5.

As can be expected, the differences in the CTRs generated by both algorithms are considerable. The results are summarized in the first two columns of Table 4 for a total of 1,000,000 queries per algorithmic framework. The CTRs derived from the algorithm defining the complete retrieval process are significantly more complete than those from the basic setting. DMs observe the alternatives in the order displayed by the engine and evaluate the first one that satisfies their predetermined requirements. The alternative selected is not necessarily the one located in the first ranking position,

Fig. 10 Code for the complete algorithm when DMs aim at finding one satisfying alternative within the first ten

```
nrows = 20;
ncols = 1000000;
A = zeros(nrows,ncols);
for i = 1
   for j = 1:ncols
      A(i,j)=rand(1);
      if A(i,j)>0.5
      A(11,j)=1;
         else
         A(i+1,j)=rand(1);
         if A(i+1,j)>0.5
          A(11,j)=2;
           else
           A(i+2,j)=rand(1);
          if A(i+2,j)>0.5
           A(11,j)=3;
            else
            A(i+3,j)=rand(1);
            if A(i+3,j)>0.5
            A(11,j)=4;
                else
                A(i+4,j)=rand(1);
                if A(i+4,j)>0.5
                A(11,j)=5;
                  else
                  A(i+5,j)=rand(1);
                  if A(i+5,j)>0.5
                  A(11,j)=6;
                     else
                     A(i+6,j)=rand(1);
                     if A(i+6,j)>0.5
                     A(11,j)=7;
                          else
                          A(i+7,j)=rand(1);
                          if A(i+7,j)>0.5
                          A(11,j)=8;
                             else
                             A(i+8,j)=rand(1);
                             if A(i+8,j)>0.5
                             A(11,j)=9;
                                 else
                                 A(i+9,j)=rand(1);
                                 if A(i+9,j)>0.5
                                 A(11,j)=10;
                                 end
                               end
                             end
                         end
                     end
                 end
             end
          end
         end
      end
   end
end
end
```

implying that DMs may eventually evaluate any of the ten alternatives composing the ranking.

An important remark follows. The fact that DMs trust the rankings displayed by the engines does not mean that they follow them blindly. It simply means that DMs proceed through the ranking in the order displayed and apply their corresponding selection criteria when deciding on which alternatives to click. The alternatives located in the initial positions of the ranking are clicked with a higher probability but do not constitute the whole set of evaluations defining the CTRs. If this were the case, and given the fact that users evaluate an average of two alternatives per search query [8, 9], all clicks would be concentrated on the first two alternatives, which would each display a CTR of 100%.

6.2 Returning to previous alternatives through the ranking

An important characteristic of our algorithmic setting is that it allows DMs to consider the whole set of alternatives composing the initial page of results over different iterations. The enhanced structure of the algorithm, designed to reflect the whole set of evaluation paths that may be generated in a sequential search process, allows simulating the actual behavior of DMs who decide not to evaluate a given alternative but return to it after proceeding further through the ranking. Including such a possibility requires the algorithm to consider the whole set of potential paths that the DMs may follow. An example of this type of behavior is

Table 4 CTRs obtained from the simple, complete, return, and return with adjusted probability algorithms when DMs aims at finding one satisfying alternative within the first ten

CTR	Simple	Complete	Return	Return adjusted
Avg	0.5005	0.9990	0.9995	0.9991
1	0.5005	0.4992	0.4999	0.5002
2	–	0.2504	0.2452	0.2504
3	–	0.1256	0.1500	0.1259
4	–	0.0625	0.0749	0.0626
5	–	0.0313	0.0150	0.0304
6	–	0.0155	0.0075	0.0152
7	–	0.0078	0.0037	0.0076
8	–	0.0038	0.0019	0.0038
9	–	0.0019	9.5350e−04	0.0019
10	–	9.9000e−04	4.8100e−04	9.6900e−04

```
A(i+1,j)=rand(1);
if A(i+1,j)>0.6
A(11,j)=2;
   else
   A(i+2,j)=rand(1);
   if A(i+2,j)>0.5
   A(11,j)=3;
     else
     A(i+3,j)=rand(1);
     if A(i+3,j)>0.5
     A(11,j)=4;
       else
       if A(i+3,j)>0.2
       A(11,j)=2;
         else
```

(a) Return code modification

```
A(i+1,j)=rand(1);
if A(i+1,j)>0.6
A(11,j)=2;
   else
   A(i+2,j)=rand(1);
   if A(i+2,j)>0.58
   A(11,j)=3;
     else
     A(i+3,j)=rand(1);
     if A(i+3,j)>0.64
     A(11,j)=4;
       else
       if A(i+3,j)>0.35
       A(11,j)=2;
         else
```

(b) Return code with adjusted probability modification

Fig. 11 Code modification for the complete algorithm when DMs aims at finding one satisfying alternative within the first ten: returning to the second alternative after observing the realization of the fourth one and deciding to proceed through the ranking

provided in Fig. 11, where DMs reconsider evaluating the second alternative after reaching the fourth one.

Figure 11 describes the modifications implemented to the code of the complete algorithm described in Fig. 10 to allow DMs to return to the second alternative after observing the realization of the fourth one and deciding to proceed through the ranking. The framework analyzed corresponds to a DM who aims at observing one satisfying alternative under complete uncertainty. That is, the cutoff value required to evaluate each alternative equals 0.5. Figure 11a and b describe a scenario where the second alternative is evaluated with a lower probability than the original benchmark, namely, 40% in contrast to the previous 50%. Thus, after deciding whether or not to click on the fourth alternative, the DM returns to the second one and considers, once again, evaluating it. Note that, to preserve the same CTRs as in the original setting, we must adapt the evaluation probabilities and account for the fact that the number of alternatives remaining to be evaluated in the fourth position is lower than those available when considering the second alternative the first time.

In particular, the third column of Table 4 illustrates how the CTR of the third alternative increases relative to the initial benchmark, given the larger number of evaluations of the second alternative declined by the DM. Further modifications are required to compensate for this fact, as illustrated in Fig. 11b—where all the probabilities have been adjusted accordingly—, and the fourth column of Table 4. We should note that the probabilities assigned to revisiting the second alternative have been computed within a relatively simple evaluation scenario, that is, finding one alternative that satisfies a predetermined criterion under complete uncertainty. In this case, the required probabilities can be intuitively computed by applying the law of large numbers to the alternatives that must be considered at each decision node. The evaluation framework and subsequent compensation schemes become increasingly complex as we require DMs to find additional satisfying alternatives.

6.3 Introducing frictions

We have designed our information retrieval algorithm as an increasingly complex tree incorporating all the decision nodes that may arise from the information retrieval process of DMs [40]. In particular, the inclusion of frictions requires defining every potential retrieval path that may be followed by a DM when proceeding through the ranking delivered by the search engine. Below, we provide an intuitive, though extended, description of the retrieval and evaluation process faced by a DM.

The main characteristics of the decision tree defined to replicate the CTR behavior of DMs can be summarized as

follows. The tree is composed of $2^{10} - 1 = 1023$ binary decision nodes and has a total of $\sum_{n=0}^{10} 2^n = 2047$ nodes. Incorporating an additional alternative, namely, accounting for eleven potential evaluations, would require defining $2^{11} - 1 = 2047$ decision nodes within a tree composed of 4095 nodes. As the empirical average traffic shares and CTR percentages described in Tables 1 and 2 illustrate, the gain obtained in terms of the algorithm's explanatory capacity is marginal compared to the computational complexity derived from the inclusion of the additional nodes required to simulate the behavior of DMs.

Consider now the initial decision that must be made by the DM, namely, retrieving information from the first alternative ranked by the engine. The DM observes the information provided by the engine and must decide whether or not to click on the alternative displayed. If the DM considers the information in line with his preferences, he clicks on the alternative and evaluates the corresponding linked website. After evaluating the website, the DM faces the second alternative displayed by the engine absent frictions.

Before deciding whether to click on the first alternative, two potential paths open for the DM, if clicking, the credibility of the ranking remains unchanged. Given his trust in the engine, the DM believes that the alternative ranked second is less likely to align with his preferences than the first place. On the other hand, if he does not click on the first alternative due to a misalignment with his preferences, friction may be introduced into the retrieval process, emphasizing the decrease in ranking credibility.

These two paths must be incorporated into the subsequent retrieval process. Note that, within each of them, the DM must consider two potential outcomes—depending on the alignment of the second alternative with his preferences—and the corresponding frictions added to the resulting paths. The process must be repeated for each of these four paths, the successive eight ones, and so on until the last ranking alternative is reached.

It should be emphasized that the intensity of the frictions incorporated into the retrieval process depends on the number of alternatives misaligned with the preferences of the DM and their respective ranking positions. That is, prevailing frictions have different effects depending on the point they are introduced in the retrieval process. A compensating mechanism could be introduced as the DM observes alternatives aligned with his preferences, an extension that can be easily incorporated within the structural framework of the model.

The algorithm becomes increasingly complex as additional paths arise from each potential node composing the tree, determined by the different combinations of misaligned alternatives and their ranking positions. All in all, DMs generate a decision tree with multiple paths and nodes, each accounting for the potential frictions arising through the retrieval process. The absence-of-frictions framework can be collapsed into a simple decision tree algorithm, lacking all the interactions that may arise through the information retrieval process of DMs.

An important remark is due here. It could be assumed that the characteristics of the websites linked affect the willingness of the DM to continue acquiring information. That is, after clicking on an alternative and evaluating the content of the linked website, the DM verifies whether or not it aligns with his preferences. If it does not, we may assume that misalignment leads to friction, a feature that can be incorporated into the retrieval algorithm. This modification would increase its complexity, but the resulting framework could be calibrated so that its mimicking quality prevails. In particular, this extension would double the set of paths that must be analyzed, one relating to whether to click on an alternative and the other to the content of the linked website.

A similar intuition follows from the searches performed by academic scholars when gathering information on a given research topic. After inputting the keywords, we obtain a rank of documents from which we can observe the title, authors, journal, some lines from the abstract, and basic metrics. We decide whether to click on the link and retrieve additional information on the corresponding document based on this information. If we observe several unrelated documents, we may consider redefining the keywords used to perform the search. We can also filter the results to some extent, but consumers browsing through the alternatives displayed by a search engine are generally unable to do so.

7 The information retrieval algorithm

A basic description of the benchmark information retrieval algorithm is provided in Fig. 12, while Fig. 13 outlines its codified structure. As was the case in the examples described through the previous sections, the sequential structure described in these figures considers each alternative in the order displayed by the engine. A random value drawn from a uniform distribution defined within [0, 1] is assigned to each of the ten alternatives composing the first page of results per query.

A cutoff value denoted by $p_i, i = 1, ..., 10$, is defined for each alternative depending on its ranking position. The behavior of the DM is determined by the randomly defined fact of whether or not the realization assigned to each alternative is higher or lower than the respective cutoff values. The DM clicks on the alternative and proceeds with the next one, absent any search frictions if it is higher. If it is lower, the DM does not click on the alternative but still considers evaluating the next one; he does so; however, with friction added to the subsequent cutoff values.

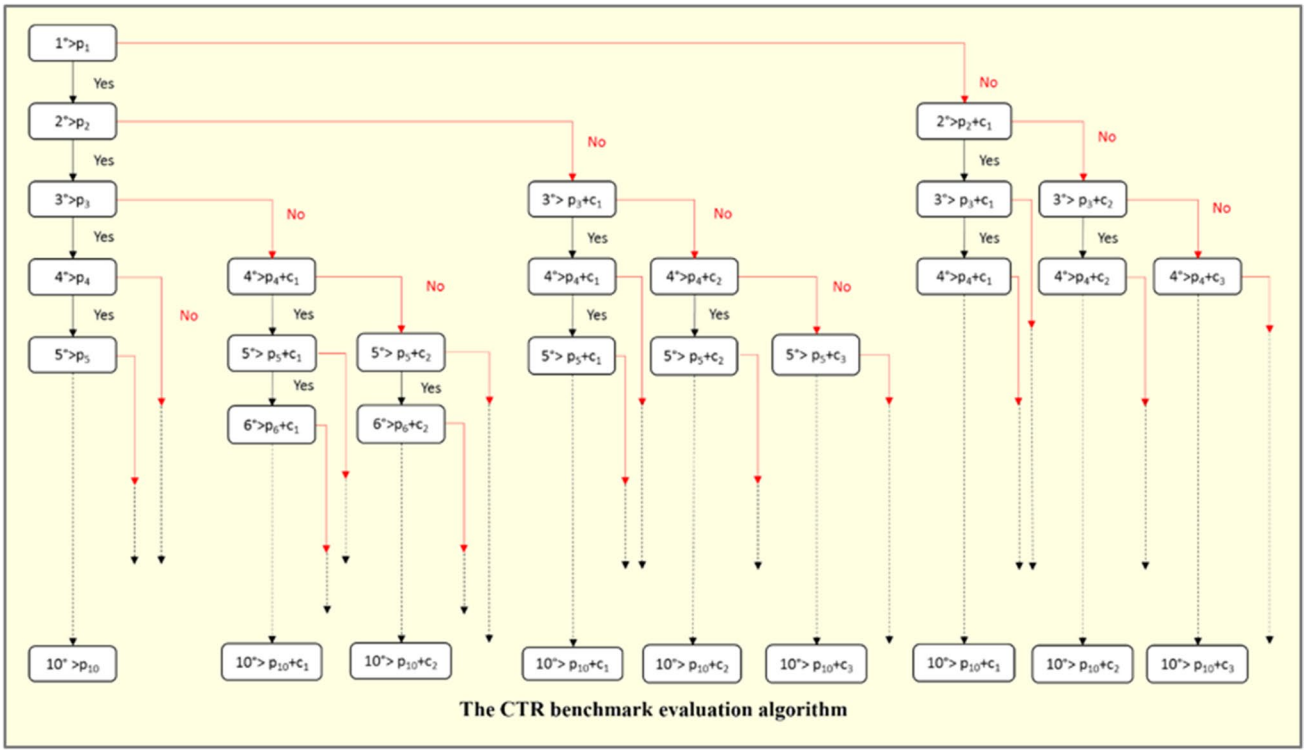


Fig. 12 Structure of the CTR benchmark evaluation algorithm

Frictions arise when a given evaluation does not align with the DM's preferences, decreasing the user's confidence in the ranking provided by the engine. That is, search frictions, denoted by $c_i, i = 1, ..., 9$, with $c_i < c_{i+1}$, arise after deciding not to click on a given alternative. Frictions, which increase as users lose confidence in the ranking, describe the path followed by DMs in terms of the number of preference misalignments and their relative ranking positions. Thus, each query may lead to multiple paths depending on which alternatives are clicked by the DM, their ranking positions, and the value of the frictions.

Note that all the interactions taking place among the potential evaluations must be completely defined. Each time the DM makes an actual decision, he is following one of two potential paths defined at each node. All the potential paths arising from each node must be defined and accounted for as part of the process. The effect from three evaluations that are not aligned with the preferences of DMs differs substantially—regarding the outcome obtained—depending on where they are located within the search process. This fundamental subtlety defines the main problem at hand.

The numerical analysis assesses the capacity of the model to replicate the actual behavior observed among online users, while different friction values are introduced depending on the performance of the evaluation process relative to the requirements of the DM. The algorithm has been coded and runs using Matlab software.

8 Numerical simulations

The intuition on which the algorithm is built follows directly from the sequential behavior observed among online users. We calibrate the algorithm using the empirical analysis of Dean [7], who computed the CTRs of the organic results ranked in the first Google page from a sample of five million queries. The corresponding CTRs are presented in the second column of Table 5. As illustrated in Table 1, though not identical, the average traffic shares of the Google ranking results described by Chitika [6] are quite similar.

We have also developed a battery of scenarios designed to evaluate the CTR behavior of DMs when their confidence in the ranking provided by the engine decreases, affecting their willingness to click on the alternatives displayed. The sets of cutoff values and search frictions defining the different evaluation scenarios analyzed are described in Table 6. Table 5 compares the CTRs obtained by Dean [7] with those derived from implementing these evaluation scenarios.

Each column describes the results obtained from simulating 1,000,000 runs over 10 potential alternatives. The number of runs has been chosen to guarantee the converge of the CTRs to the cutoff values defining the respective frictionless settings. More precisely, the columns denoted by 'None' assume zero search frictions and are used to validate the behavior of the algorithm as it converges to the cutoff values defined by the respective benchmarks. In this regard, the

Fig. 13 Codified structure of the CTR benchmark evaluation algorithm

```
generate random uniform_i
        if random uniform_i > p_i
        page is visited and evaluated;
        random uniform_{i+1};
                        if random uniform_{i+1} > p_{i+1}
                        page is visited and evaluated;
                        random uniform_{i+2};
                                        if random uniform_{i+2} > p_{i+2}
                                        ......
                        else
                        if random uniform_{i+2} > p_{i+2} + c_1
                        page is visited and evaluated;
                        random uniform_{i+3};
                                        if random uniform_{i+3} > p_{i+3} + c_1
                                        page is visited and evaluated;
                                        random uniform_{i+4};
                                        ....

                                        else
                                        random uniform_{i+4};
                                        if random uniform_{i+4} > p_{i+4} + c_2
                                        page is visited and evaluated;
                                        random uniform_{i+5};
                                        ....
                        else
                        if random uniform_{i+3} > p_{i+3} + c_2
                        page is visited and evaluated;
                        random uniform_{i+4};
                        ....
        else
        if random uniform_{i+1} > p_{i+1} + c_1
        page is visited and evaluated;
        random uniform_{i+2};
                if random uniform_{i+2} > p_{i+2} + c_1
                page is visited and evaluated;
                random uniform_{i+2};
                ....
                        end
                                        end
                        end
        end
end
```

'None' column within the $[p_1,\ldots, p_{10}] = [68,\ldots, 97]$ scenario represents the calibration benchmark based on the results reported by Dean [7].

The second and third scenarios considered constitute intuitive validation sequences, with the probability of clicking decreasing from 90 to 5% and from 50 to 5% as the DM proceeds down the ranking, respectively. These scenarios provide higher flexibility when analyzing the cumulative effects that frictions have on the information retrieval behavior of DMs. In addition, they could be assumed to reflect different degrees of trust in the rankings delivered by the search engines, ranging from an initially confident to a more neutral approach.

Regarding search frictions, the first two subsets define small progressive increments in the probability of rejection. They have been specifically introduced to test the behavior of the algorithm within the $[p_1,\ldots, p_{10}] = [68,\ldots, 97]$ scenario. The same intuition applies to the three remaining subsets, with the final one constituting the largest cumulative increment in rejection probabilities. To provide additional intuition regarding the output from the simulations, Table 7 presents the stochastic realizations and the corresponding

Table 5 Comparing CTRs across different scenarios

Scenario		$[p_1,\ldots,p_{10}]=[68,\ldots,97]$						$[p_1,\ldots,p_{10}]=[10,\ldots,95]$				$[p_1,\ldots,p_{10}]=[50,\ldots,95]$			
		$[c_1,\ldots,c_9]$						$[c_1,\ldots,c_9]$				$[c_1,\ldots,c_9]$			
CTR	Dean [7]	None	0.2–1.8	0.5–4.5	5–10	10–15	5–45	None	5–10	10–15	5–45	None	5–10	10–15	5–45
Avg		1.19	1.12	1.02	0.78	0.62	0.68	4.55	4.16	3.92	3.96	2.75	2.18	1.86	1.82
1	31.7	32.02	32.07	32.01	31.96	31.91	31.98	90.00	89.99	90.06	89.99	50.07	50.02	50.02	49.95
2	24.7	24.98	24.85	24.72	21.61	18.15	21.56	80.00	79.47	79.04	79.55	45.07	42.48	40.00	42.44
3	18.7	19.03	18.73	18.20	13.82	9.23	11.66	69.96	68.61	67.15	68.44	40.05	35.81	31.94	34.65
4	13.6	13.98	13.58	12.93	7.76	2.73	2.76	60.07	57.38	54.92	56.89	34.96	29.69	25.01	26.32
5	9.5	9.99	9.41	8.46	2.75	0	0.14	50.00	46.15	42.64	44.73	30.01	23.77	18.57	17.68
6	6.2	6.02	5.17	3.96	0	0	0	39.97	35.05	30.79	32.00	25.03	17.86	12.75	8.59
7	4.1	3.99	2.97	1.51	0	0	0	30.01	24.04	19.19	18.62	20.05	11.99	6.77	1.96
8	3.1	2.98	1.84	0.17	0	0	0	19.96	13.10	8.03	5.45	15.04	6.21	0.98	0.17
9	3	3	1.63	0	0	0	0	10.02	2.18	0	0.18	10.02	0.58	0	0
10	3	3	1.43	0	0	0	0	5.01	0	0	0	4.96	0	0	0

All numbers indicate percentages, except the Avg row, which describes the average number of pages clicked per search

Table 6 Cutoff values and search frictions defining the evaluation scenarios

Evaluation scenarios	
$[p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9, p_{10}]$	$[c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8, c_9]$
[68, 75, 81, 86, 90, 94, 96, 97, 97, 97]	[0.2, 0.4, 0.6, 0.8, 1, 1.2, 1.4, 1.6, 1.8]
[10, 20, 30, 40, 50, 60, 70, 80, 90, 95]	[0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5]
[50, 55, 60, 65, 70, 75, 80, 85, 90, 95]	[5, 6, 7, 8, 9, 10, 10, 10, 10]
	[10, 11, 12, 13, 14, 15, 15, 15, 15]
	[5, 10, 15, 20, 25, 30, 35, 40, 45]

All numbers indicate percentages

Table 7 Simulation sample from the $[p_1,\ldots,p_{10}]=[10,\ldots,95]$ and $[c_1,\ldots,c_9]=$None scenario

	Search Queries									
Stochastic realizations	0.815	0.158	0.656	0.706	0.439	0.276	0.751	0.841	0.352	0.076
	0.906	0.971	0.036	0.032	0.382	0.680	0.255	0.254	0.831	0.054
	0.127	0.957	0.849	0.277	0.766	0.655	0.506	0.814	0.585	0.531
	0.913	0.485	0.934	0.046	0.795	0.163	0.699	0.244	0.550	0.779
	0.632	0.800	0.679	0.097	0.187	0.119	0.891	0.929	0.917	0.934
	0.098	0.142	0.758	0.823	0.49	0.498	0.959	0.350	0.286	0.130
	0.278	0.422	0.743	0.695	0.446	0.960	0.547	0.197	0.757	0.569
	0.547	0.916	0.392	0.317	0.646	0.340	0.139	0.251	0.754	0.469
	0.958	0.792	0.655	0.950	0.709	0.585	0.149	0.616	0.380	0.012
	0.965	0.959	0.171	0.034	0.755	0.224	0.258	0.473	0.568	0.337
Pages clicked	1	1	1	1	1	1	1	1	1	3
	2	2	3	6	2	2	2	2	2	4
	4	3	4	9	3	3	3	3	3	5
	5	4	5	0	4	7	4	5	4	0
	9	5	6	0	0	0	5	0	5	0
	10	8	7	0	0	0	6	0	7	0
	0	10	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0

pages clicked by the DM for ten queries of the $[p_1,\ldots, p_{10}]=[10,\ldots, 95]$ and $[c_1,\ldots, c_9]=$None scenario.

8.1 Analysis of the click-through rates

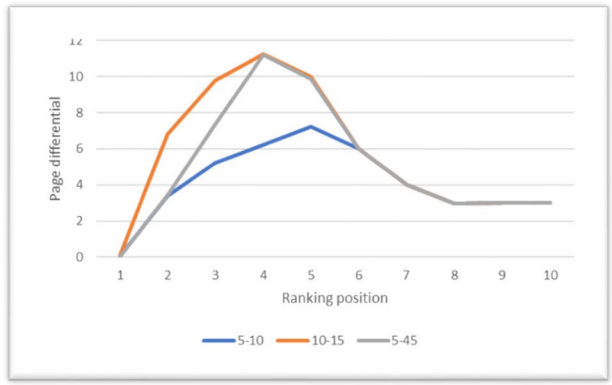
The numerical results illustrate that the worst performances in all scenarios arise from the $[c_1,\ldots, c_9]=[10,\ldots, 15]$ friction framework. This result follows from the significant initial increment in the probability of rejection, an effect that dominates the higher cumulative rejection probability defined by $[c_1,\ldots, c_9]=[5,\ldots, 45]$. Also, note the substantial decrease in the average number of pages clicked within these friction frameworks, a consistent effect throughout all the scenarios as described in the fourth row of Table 5.

The $[p_1,\ldots, p_{10}]=[68,\ldots, 97]$ and $[p_1,\ldots, p_{10}]=[50,\ldots, 95]$ scenarios perform considerably worse than $[p_1,\ldots, p_{10}]=[10,\ldots, 95]$ due to the higher rejection probabilities derived from identical search frictions. In other words, the $[p_1,\ldots, p_{10}]=[10,\ldots, 95]$ scenario allows for a larger set of pages to be clicked with higher probability throughout the evaluation process, preserving similarities relative to the frictionless case halfway through its rankings.

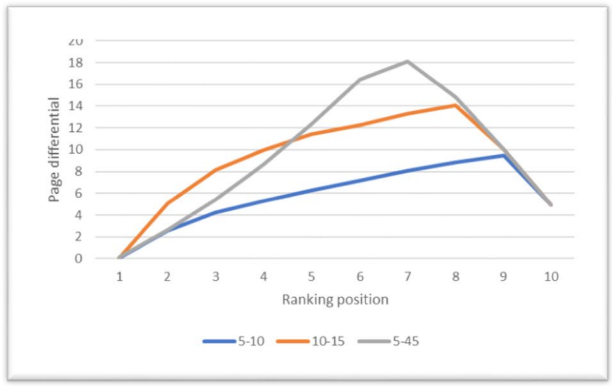
We now analyze the $[p_1,\ldots, p_{10}]=[68,\ldots, 97]$ scenario and highlight the importance of being placed within the first five ranking positions when accounting for search frictions. Note that $[c_1,\ldots, c_9]=[0.5,\ldots, 4.5]$ performs relatively worse than $[c_1,\ldots, c_9]=[0.2,\ldots, 1.8]$ only within the second half of ranked alternatives. Thus, despite the larger search frictions, the number of initial pages clicked remains almost unchanged, as also illustrated in the other scenarios when comparing the $[c_1,\ldots, c_9]=[5,\ldots, 10]$ and $[c_1,\ldots, c_9]=[5,\ldots, 45]$ frameworks, emphasizing the importance of the initial increments in rejection probability—while noting the actual cumulative divergences arising as the DM proceeds further down the set of ranked alternatives.

In order to provide additional intuition, Fig. 14 illustrates the percentage differences in the number of pages clicked relative to the frictionless framework for all scenarios. The horizontal axis represents the ranking position of the page analyzed, while the vertical axis corresponds to the number of pages of difference in percentage terms. The analysis focuses on the tendencies displayed by the $[c_1,\ldots, c_9]=[5,\ldots, 10]$, $[c_1,\ldots, c_9]=[10,\ldots, 15]$ and $[c_1,\ldots, c_9]=[5,\ldots, 45]$ search friction frameworks within the different evaluation scenarios.

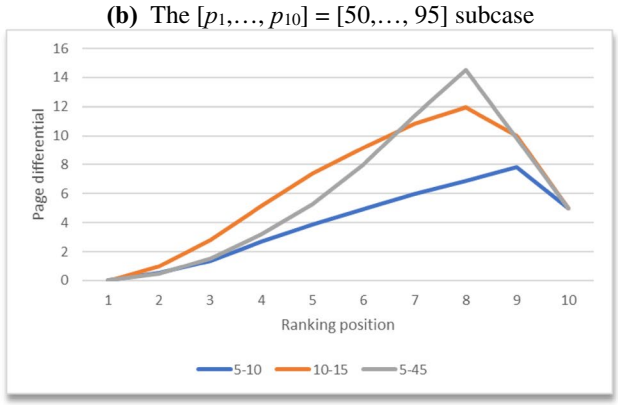
Figures 14(a)-14(c) emphasize the inferior performance of the $[c_1,\ldots, c_9]=[10,\ldots, 15]$ framework in all scenarios, as well as the cumulative divergent process following from $[c_1,\ldots, c_9]=[5,\ldots, 45]$. These figures also illustrate the smoother tendency displayed by the $[c_1,\ldots, c_9]=[5,\ldots, 10]$ framework in all scenarios, a feature that can be validated through the evolution of $[c_1,\ldots, c_9]=[0.2,\ldots, 1.8]$ and $[c_1,\ldots, c_9]=[0.5,\ldots, 4.5]$ within the $[p_1,\ldots, p_{10}]=[68,\ldots, 97]$



(a) The $[p_1,\ldots, p_{10}] = [68,\ldots, 97]$ subcase



(b) The $[p_1,\ldots, p_{10}] = [50,\ldots, 95]$ subcase



(c) The $[p_1,\ldots, p_{10}] = [10,\ldots, 95]$ subcase

Fig. 14 Differences in the number of pages clicked relative to the frictionless setting

scenario. That is, even within the most restrictive scenario, defined by the rejection probabilities that follow from the actual CTRs observed in real-life settings, the increase in CTR differences triggered by $[c_1,\ldots, c_9]=[5,\ldots, 10]$ remains low within the initial alternatives composing the ranking.

The intuition is similar when considering the CTR differences triggered by $[c_1,\ldots, c_9]=[5,\ldots, 10]$, which defines a lower cumulative increase in the rejection probabilities than $[c_1,\ldots, c_9]=[5,\ldots, 45]$, through the different evaluation scenarios. Both friction frameworks deliver a smoother

increment in CTR differences within the initial ranking positions when compared to $[c_1,\ldots, c_9]=[10,\ldots, 15]$. Note also how the initial increment in CTR differences displayed by this latter framework is cumulatively surpassed by $[c_1,\ldots, c_9]=[5,\ldots, 45]$.

Thus, the initial value of the frictions and their cumulative persistence determine the CTR differences exhibited through the information retrieval process of DMs. The simple evaluation algorithm described in Fig. 5 cannot account for any of these features, particularly the cumulative ones, since any increment in the rejection probabilities would have to be defined individually for each decision node, omitting any potential interaction across them. More importantly, the differences between algorithms and their capacity to account for potential modifications in the behavior of DMs would be more substantial when considering retrieval processes involving less than ten alternatives, as the intuition derived from the analyses performed in Figs. 7, 8, 9 and 10 and Table 4 suggests.

Finally, when considering the average number of clicks per search, we observe that relatively few pages are clicked within the $[p_1,\ldots, p_{10}]=[68,\ldots, 97]$ scenario. Similar values to the ones obtained by Jansen et al. [9] and Baeza-Yates [8] follow from the $[p_1,\ldots, p_{10}]=[50,\ldots, 95]$ scenario, while too many pages are clicked on average within $[p_1,\ldots, p_{10}]=[10,\ldots, 95]$. Thus, a less restrictive evaluation scenario than the one validating the results reported by Dean [7] is required for DMs to click on two pages per search query.

9 Managerial implications

The benchmark decision-tree algorithm introduced in the current paper applies to any sequential information retrieval framework conditioned by the subjective characteristics of DMs. The algorithm is sufficiently flexible to incorporate any information retrieval requirements within its main structure. Each node within the algorithm has been designed to consider a simple random event to implement a direct decision rule. Each node provides managers with a highly malleable setting within the large structure composing the algorithm. That is, the decision rule can be extended to incorporate different strategic qualities inherent to most decision and information retrieval processes currently analyzed in the literature.

The intensity and duration of frictions can be easily modified through the retrieval process and adapted to any characteristic of the DM. For instance, we could bind the cumulative effect of frictions by imposing an upper limit on their sequential increments. DMs lose confidence in the ranking after facing a number of misalignments, but the increments do not spread through the whole set of alternatives. We introduce below three scenarios allowing for a maximum of up to three cumulative misalignments between the preferences of the DM and the alternatives. That is, frictions increase for a maximum of three misalignments, remaining constant afterward. The numerical scenarios compared are given by:

- the addition of unique friction, $c_1 = 0.1$, after the first misalignment to all remaining cutoff values;
- the inclusion of two frictions, namely, $c_1 = 0.1$ after the first misalignment, and $c_2 = 0.2$ after the second one. c_2 is added to all the cutoff values remaining after the second misalignment;
- a more complex cumulative scenario with three frictions following the same intuition as the previous ones, that is, $c_1 = 0.1$, $c_2 = 0.2$, and $c_3 = 0.3$, the latter added after the third misalignment to the remaining cutoff values.

The value of frictions remains constant after the respective upper limits are reached, independently of the number of additional misalignments. We have integrated these scenarios into the $[p1,\ldots, p10]=[10,\ldots, 95]$ benchmark framework for illustrative purposes. The CTRs obtained are described in Table 8, while Fig. 15 illustrates the CTR differences in percentage terms between each scenario and the frictionless benchmark framework.

Three main results follow from the simulations. Firstly, the top three ranked alternatives display almost identical CTRs in all the scenarios. Secondly, the differences between the second and third scenarios are concentrated on the sixth and seventh alternatives, the remaining CTRs being almost identical. However, these differences are significant when comparing the first and second scenarios. That is, increments in frictions have different effects on CTRs depending on the total number of frictions considered and the ranking position of the corresponding alternatives. Finally, the non-linear

Table 8 CTRs without and with up to three levels of friction

Scenario	$[p_1,\ldots, p_{10}]=[10,\ldots, 95]$			
CTR	None	$c_1=0.1$	$[c_1=0.1; c_2=0.2]$	$[c_1=0.1; c_2=0.2; c_3=0.3]$
Avg	4.55	3.97	3.70	3.62
1	90.00	89.95	89.98	90.00
2	80.00	79.01	78.91	78.99
3	69.96	67.15	66.91	66.90
4	60.07	55.08	53.66	53.59
5	50.00	43.04	39.87	39.02
6	39.97	31.58	26.04	23.60
7	30.01	20.66	13.01	8.21
8	19.96	10.20	1.15	1.15
9	10.02	0.04	0.03	0.04
10	5.01	0	0	0

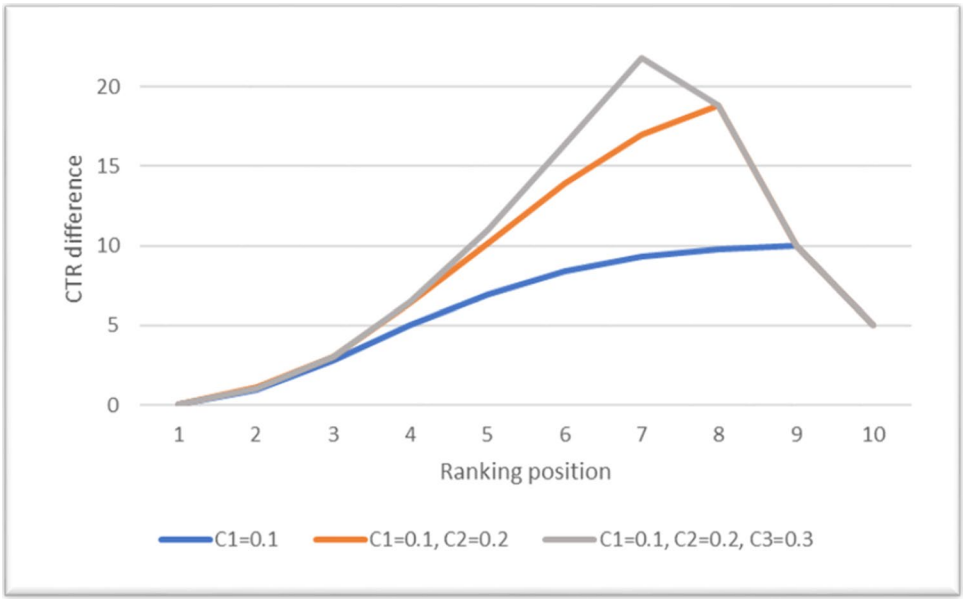


Fig. 15 CTR differences relative to the frictionless framework within $[p_1,\ldots, p_{10}]=[10,\ldots, 95]$

behavior exhibited by the CTR differences describes an incremental effect concentrated on the alternatives composing the lower half of the ranking.

The systems literature has consistently emphasized—and validated empirically—the importance of being located within the first two positions of the ranking provided by a search engine [41, 42]. Extrapolating the potential effects derived from the introduction of any type of friction on the set of ranked alternatives is essential to evaluate the willingness to pay firms for an improvement in their relative positions.

We conclude by highlighting the algorithm's capacity to accommodate any assumption regarding the behavior of DMs through the information retrieval process, providing managers with a benchmarking tool to analyze the effects of any potential modification to the preferences of DMs. In this regard, the algorithm can be easily extended to evaluate sequential games [43] and strategic social interactions [44], since its structure allows for the implementation of any potential behavioral strategy within each decision node.

10 Conclusion and future research directions

We have defined a computable benchmark that can simulate the behavior of different types of users as they proceed through the alternatives ranked by a search engine. The framework accounts for the subjective preferences defining the behavior of DMs, together with the characteristics of the alternatives and the subsequent evaluations and decisions made by the DMs. We have illustrated the substantial differences between a trivial algorithm, providing the same CTRs but ignoring the interactions across alternatives and the one introduced in the current paper.

The main contribution of the current algorithmic framework can be summarized as follows:

1. The design of the algorithm is based on a decision model that mimics the sequential search behavior of DMs, defining the information retrieval processes observed in real-life settings. Behavioral modifications can be introduced at any point through the process, and the consequences of CTRs analyzed.
2. The algorithm provides a numerical evaluation for each characteristic observed by the DMs and the resulting action in terms of clicking behavior. Most empirical studies aim at extrapolating the evaluations of DMs from the behavior observed, lacking an exact numerical value that can be assigned to represent their preferences.

One of the main qualities of the mimicking algorithm follows from the (almost) complete absence of behavioral constraints imposed on its design, increasing its structural malleability and capacity to incorporate any type of frictions or information retrieval strategies. To illustrate this point, we have analyzed the consequences derived from reliability modifications, though more complex scenarios could be incorporated and studied using the algorithm as a benchmark. For instance, the information retrieval effects of impatience, quality signals from recommender systems, or shifts in the subjective beliefs of DMs can be easily analyzed by adapting the algorithm accordingly.

Finally, while representing a sequential decision-making process that could be defined as rational, the current algorithm is based on empirical observations, leaving

aside the complexities of mathematical formalization. In this regard, future research should examine the optimality of the behavior followed by DMs relative to the combinatorial optima that would result from evaluating different subsets of alternatives among those displayed in the first page of results delivered by a search engine.

Acknowledgements Dr. Madjid Tavana is grateful for the partial support he received from the Czech Science Foundation (GAČR19-13946S) for this research.

Declarations

Conflict of interest The authors declared that they have no conflict of interest.

References

1. Epstein R, Robertson RE (2015) The Search Engine Manipulation Effect (SEME) and its possible impact on the outcomes of elections. Proc Natl Acad Sci USA 112:E4512–E4521
2. Jansen MBJ, Spink A, Saracevic T (2000) Real life, real users, and real needs: a study and analysis of user queries on the Web. Inf Process Manage 36:207–227
3. Luo W, Cook D, Karson EJ (2011) Search advertising placement strategy: exploring the efficacy of the conventional wisdom. Inf Manag 48:404–411
4. Lewandowski D, Kammerer Y (2020) Factors influencing viewing behaviour on search engine results pages: a review of eye-tracking research. Behav Inf Technol. https://doi.org/10.1080/0144929X.2020.1761450
5. Lorigo L, Haridasan M, Brynjarsdóttir H, Xia L, Joachims T, Gay G, Granka L, Pellacini F, Pan B (2008) Eye tracking and online search: lessons learned and challenges ahead. J Am Soc Inf Sci Technol 59:1041–1052
6. Chitika (2013) The value of Google result positioning. Chitika Insights, June 7, 2013. Chitika, Westborough. perma.cc/7AGC-HTDH
7. Dean B (2019) We analyzed 5 million Google search results. Here's what we learned about organic click through rate. https://backlinko.com/google-ctr-stats. Accessed 23 Sep 2021
8. Baeza-Yates R (2005) Applications of web query mining. In: Losada DE, Fernández-Luna JM (eds) Advances in information retrieval. information systems and applications, incl. Internet/Web, and HCI, vol 3408. Springer, Berlin, pp 7–22
9. Jansen MBJ, Spink A, Bateman J, Saracevic T (1998) Real life information retrieval: a study of user queries on the web. ACM SIGIR Forum 32:5–17
10. European Commission (2016) Online platforms. Special Eurobarometer No. 447, April 2016. European Union
11. Hendahewa C, Shah C (2017) Evaluating user search trails in exploratory search tasks. Inf Process Manage 53:905–922
12. Yu H, Jatowt A, Blanco R, Joho H, Jose JM (2017) Decoding multi-click search behavior based on marginal utility. Inf Retrieval J 20:25–52
13. Simon HA (1997) Administrative behavior. The Free Press, New York
14. Schneider M, Deck C, Shor M, Besedeš T, Sarangi S (2019) Optimizing choice architectures. Decis Anal 16:2–30
15. Speier-Pero C (2019) Using aggregated data under time pressure: a mechanism for coping with information overload. J Decis Syst 28:82–100
16. Victorelli EZ, Dos Reis JC, Hornung H, Prado AB (2020) Understanding human-data interaction: literature review and recommendations for design. Int J Hum Comput Stud 134:13–32
17. Preibusch S, Kübler D, Beresford AR (2013) Price versus privacy: an experiment into the competitive advantage of collecting less personal information. Electron Commer Res 13:423–455
18. Sun H, Fan M, Tan Y (2020) An empirical analysis of seller advertising strategies in an online marketplace. Inf Syst Res 31:37–56
19. Yoo B, Jeon S, Han T (2016) An analysis of popularity information effects: field experiments in an online marketplace. Electron Commer Res Appl 17:87–98
20. Bell D, Mgbemena C (2018) Data-driven agent-based exploration of customer behavior. SIMULATION 94:195–212
21. Dunke F, Nickel S (2020) Neural networks for the metamodeling of simulation models with online decision making. Simul Model Pract Theory 99:102016
22. Mahony C, Sammon D, Heavin C (2016) Design guidelines for online resources: a longitudinal analysis of information processing. J Decis Syst 25:329–342
23. Zhang J, Adomavicius G, Gupta A, Ketter W (2020) Consumption and performance: understanding longitudinal dynamics of recommender systems via an agent-based simulation framework. Inf Syst Res 31:76–101
24. Unkel J, Haas A (2017) The effects of credibility cues on the selection of search engine results. J Am Soc Inf Sci 68:1850–1862
25. Wang J, Yang Z, Brocato ED (2018) An investigation into the antecedents of prepurchase online search. Inf Manag 55:285–293
26. Power DJ, Cyphert D, Roth RM (2019) Analytics, bias, and evidence: the quest for rational decision making. J Decis Syst 28:120–137
27. Yi C, Jiang Z, Li X, Lu X (2019) Leveraging user-generated content for product promotion: the effects of firm-highlighted reviews. Inf Syst Res 30:711–725
28. Li H, Duan H, Zheng Y, Wang Q, Wang Y (2020) A CTR prediction model based on user interest via attention mechanism. Appl Intell 50:1192–1203
29. Wang Q, Liu F, Xing S, Zhao X (2019) Research on CTR prediction based on stacked autoencoder. Appl Intell 49:2970–2981
30. Qin J, Zhang W, Wu X, Jin J, Fang Y, Yu Y (2020) User behavior retrieval for Click-Through Rate prediction. In: Proceedings of the 43rd international ACM SIGIR Conference on research and development in information retrieval (SIGIR' 20). Association for Computing Machinery, New York, pp 2347–2356
31. Nelson P (1970) Information and consumer behavior. J Polit Econ 78:311–329
32. Tavana M, Di Caprio D, Santos Arteaga FJ, O'Connor A (2015) A novel entropy-based decision support framework for uncertainty resolution in the initial subjective evaluations of experts: the NATO enlargement problem. Decis Support Syst 74:135–149
33. Herrmann JW (2015) Engineering decision making and risk management. Wiley, Hoboken
34. Pei S, Hu Q (2018) Partially monotonic decision trees. Inf Sci 424:104–117
35. Sagi O, Rokach L (2020) Explainable decision forest: transforming a decision forest into an interpretable tree. Inf Fusion 61:124–138
36. Herce-Zelaya J, Porcel C, Bernabé-Moreno J, Tejeda-Lorente A, Herrera-Viedma E (2020) New technique to alleviate the cold start problem in recommender systems using information from social media and random decision forests. Inf Sci 536:156–170

37. Rokach L (2016) Decision forest: twenty years of research. Inf Fusion 27:111–125
38. Lim S (2013) College students' credibility judgments and heuristics concerning Wikipedia. Inf Process Manage 49:405–419
39. Tavana M, Santos-Arteaga FJ, Di Caprio D, Tierney K (2016) Modeling signal-based decisions in online search environments: a non-recursive forward-looking approach. Information & Management 53:207–226
40. Schulz F (2008) Trees with exponentially growing costs. Inf Comput 206:569–578
41. Gao R, Shah C (2020) Toward creating a fairer ranking in search engine results. Inf Process Manag 57:102138
42. Pan B, Hembrooke H, Joachims T, Lorigo L, Gay G, Granka L (2007) In Google we trust: users' decisions on rank, position, and relevance. J Comput Mediated Commun 12:801–823
43. Brihaye T, Geeraerts G, Hallet M, Le Roux S (2020) On the termination of dynamics in sequential games. Inf Comput 272:104505
44. van Benthem J (2018) Computation as social agency: what, how and who. Inf Comput 261:519–535