An online advertisement's clickthrough rate provides a fundamental measure of its quality, which is widely used in ad selection strategies. Unfortunately, ads placed in contexts where they are rarely viewed — or where users are unlikely to be interested in commercial results — may receive few clicks regardless of their quality. In this paper, we model the variability of a user’s browsing behavior for the purpose of click analysis and prediction in sponsored search. Our model incorporates several important contextual factors that influence ad clickthrough rates, including the user’s query and ad placement on search engine result pages. We formally model these factors with respect to the list of ads displayed on a result page, the probability that the user will initiate browsing of this list, and the persistence of the user in browsing the list. We incorporate these factors into existing click models by augmenting them with appropriate query and location biases. Using expectation maximization, we learn the parameters of these augmented models from click signals recorded in the logs of a commercial search engine.

To evaluate the performance of the models and to compare them with state-of-the-art performance, we apply standard evaluation metrics, including log-likelihood and perplexity. Our evaluation results indicate that, through the incorporation of query and location biases, significant improvements can be achieved in predicting browsing and click behavior in sponsored search. In addition, we explore the extent to which these biases actually reflect varying behavioral patterns. Our observations confirm that correlations exist between the biases and user search behavior.

Categories and Subject Descriptors: H.3.5 [Information Storage and Retrieval]: Online Information Services; H.2.8 [Database Applications]: Data Mining

General Terms: Algorithms, Experimentation, Measurement

Additional Key Words and Phrases: Click model, sponsored search, contextual factors, clickthrough, Bayesian inference, query log

1. INTRODUCTION

Consider an ad as it appears in the context of a search engine result page. This context may strongly influence whether or not the user will click on the ad, and includes the display location of the ad, the rank of the ad, the user's query, the organic search results, and the other ads displayed along with it. A SERP (search engine result page) represents a result page that the search engine returns as the result of a query. When a user clicks on an ad displayed on a SERP, they are transferred to the ad's landing page, and the context and click information are typically stored in a log by the search engine. An ad impression represents a pair of ad and SERP where the ad appears on the SERP as a result of a query. Commercial search engines generate and store this information as part of their normal operations, and historical clickthrough rates for...
ads can be computed directly from this log information. By modeling user browsing behavior, this context may be exploited to compute measures of ad quality.

The main goal of this paper is to gain insight into user browsing and click behavior in sponsored search, which in turn improves our ability to infer the probability of clicks on the advertising links. An intuitive strategy for ad selection is to display better performing ads earlier in the result list (i.e., at higher rank positions) as measured by their individual clickthrough probabilities. However, unrelated ads, no matter how popular they are, may annoy the user, have an adverse effect, and produce no economic benefit to the advertiser [Abrams and Schwarz 2007; Broder et al. 2008]. Moreover, ads that are rarely viewed by users — perhaps because they are often displayed at lower rank positions on the page — may be inappropriately penalized if the context in which they were shown is not considered.

To address the above issues, an ad should be considered in the context of the ads appearing prior to it on a result page. The rate at which the ad is viewed and clicked is therefore assumed to depend both on its own quality and on the quality of the other ads that are displayed above it on the result page. This concept is introduced by Ghosh and Mahdian [2008] as the externality effects of ads in advertising auctions, and it is based on the assumption that users visually scan the ad list from top to bottom. Once an ad is examined by the user, ad-specific factors will determine the click decision and continuation probability. This user behavior can be also explained by the cascade model [Craswell et al. 2008]; i.e., an ad is examined only if the user scans over all the previously displayed ads. This model assumes the user is infinitely persistent in continuing to examine ads until they click on one; hence, it only applies to the first click. There have been efforts to extend the cascade model through various hypotheses and assumptions, all aiming to model the user browsing and click behavior in a more realistic fashion, in which multiple clicks are permitted.

In this paper, we model a user’s browsing and click behavior in sponsored search by extending the cascade model to include the effects of context. Our work builds on an advanced cascade model, known as the dynamic Bayesian network (DBN) model [Chapelle and Zhang 2009]. We define the attraction probability of an ad with respect to a query, as the probability that a user who issued the query and examined the ad, perceives the ad to be relevant to their commercial need [Ashkan and Clarke 2009; Dai et al. 2006; Hu et al. 2009] and therefore clicks on it. We also define the satisfaction probability for each ad with respect to a query. This probability represents the chance that a user who issued the query, clicked on the ad, and viewed its landing page, finds that the landing page satisfies their commercial need.

We augment context by adding biases that are based upon known properties of sponsored search, and upon the differences that exist in user browsing behavior over advertising links as opposed to organic links. According to Jansen and Resnick [2006], users exhibit a strong bias against ads, as opposed to organic links. Thus, users have a stronger tendency to consider the organic links rather than ads. As for the ads themselves, users are known to pay more attention to the top-listed ads (i.e., those appearing above the organic results) as opposed to the side-listed ads (i.e. those appearing to the right of the organic results) [Jansen and Resnick 2006]. The nature of the query may also influence the probability that the user will initiate browsing of the ad list, and continue browsing the ad list once they start. A user who issues a commercially oriented query may be assumed to have a greater tendency to purchase or utilize a commercial service, and thus they click on ads more frequently [Ashkan and Clarke 2009].

We formally model a notion of location bias in this work, in order to account for top-listed and side-listed ads separately. Furthermore, query biases are introduced and parameterized in order to account for the probability that the user will initiate brows-
ing of the ad list, and for their persistence (patience) in continuing to browse through the list. The initiation probability with respect to a particular query is defined as the chance that the user who issues this query will eventually initiate browsing the ad list. The persistence or transition probability with respect to a query is defined as the chance that the corresponding user who examined a particular ad at rank $i$ will continue on to examine the ad at rank $i + 1$. Both the initiation probability and the transition probability are determined separately for different display locations of ads on result pages.

In this work, a group of existing models are re-visited and adapted to the sponsored search domain based on these assumptions, with their parameters learned from the click signals recorded for the advertising links appearing in a log of search result pages. We use a group of standard evaluation metrics in order to evaluate the performance of the models and compare them with prior state-of-the-art performance. The evaluation results indicate that significant improvements are achieved in the prediction of the browsing and click behavior in sponsored search. To sum up, this paper addresses the following issues:

- In measuring the historical performance of an ad, we must take into account the overall quality of all the preceding ads displayed on the result pages, rather than simply computing a clickthrough rate, as if the ad appeared in isolation. To reflect this property, we build our work upon the cascade model.
- We define an initiation probability for each location that ads can be displayed (i.e., top and side), which indicates the probability that a user will initiate browsing of ads in that location. The stronger the commercial intent of a user, the more likely the user will initiate browsing of ads.
- We assume that transition probabilities between the ads can vary for different queries and locations. In related work, the transition probability is often assumed to be constant across all ads, locations, and queries. By incorporating variable transition probabilities, we are able to better cope with actual user behavior.
- Initiation probabilities and transition probabilities are assumed to depend on the placement of the ad list on the page. If there is a tendency to look at the ads, the top-listed ones are assumed to be examined more often and more persistently by the user when compared to the side-listed ads. We reflect this property by employing a different probabilistic model for the top and the side.
- Contextual factors are modeled through query-, ad-, and page-dependant parameters. We learn and update these parameters in an online fashion.
- Variability in user behavior and search intent is explained with respect to the introduced biases.

The remainder of the paper is organized as follows: A background overview of related work is provided by Section 2. The location and query biases, along with the procedure for parameter inference, are introduced in Section 3. In Section 4, the results of an empirical study are presented. Section 5 reports a study that explores whether or not the introduced biases actually reflect varying behavioral patterns for different users. We provide a summary and discuss implications of this work in Section 6.

2. BACKGROUND AND RELATED WORK

Online advertising provides the major source of revenue for commercial search engines. In its simplest form, for a given query, a set of candidate ads are obtained through an ad retrieval algorithm that matches the query terms with the ads’ bid
These ads are then ranked in decreasing order based on the expected revenue estimated for each individual ad, \( a \), typically as: \( \text{REV}(a) = b_a \times Q_a \) where \( b_a \) is the bid placed by an advertiser, and \( Q_a \) is the quality score of the ad as a notion of its performance predicted by the search engine.

The cost-per-click model [Fain and Pedersen 2006] is among the most common revenue models, and it charges an advertiser for the amount they need to pay per click. This amount is computed based on the value of the bids placed through the generalized second price (GSP) auction [Edelman et al. 2005]. Given \( a \) ranked right above \( a' \) in the result list, the amount that \( a \) has to pay upon click is computed based on the bid placed by the lower ad (i.e., \( a' \)) and on the quality score of the two ads, as follows: \( b_{a'} \times Q_{a'}/Q_a \). It can be seen that the ad quality score has a crucial effect in determining the expected revenue for the search engine. From the user’s point of view, a good quality ad should be trustworthy and relevant to their interests. Hence, the expected ad clickthrough rate (CTR) is the classic measure of ad quality, such that a click on an ad is viewed as a potential purchase opportunity for the service or product that is offered by the advertiser. In this section, we review several ad click studies, which we follow with a discussion of browsing models in Web search.

2.1. Ad Click Analysis

Information obtained from logs and other implicit feedback resources, such as queries and clicks [Richardson 2008], have been widely used to interpret and predict future user behavior. Richardson et al. [2007] describe a prediction model based on logistic regression using statistics of existing ads. They incorporate features from ads, such as their bid terms, the length of ads, the landing pages, and statistics concerning related ads. Debmbczynski et al. [2008] approximate the title and the body of an ad by combining all queries for which the ad was displayed. They use these ad features along with the ad’s target URL to build a prediction model based on decision rules, generating recommendations on how to improve the quality of ads. Graepel et al. [2010] represent an ad impression through discrete multi-valued features extracted from the existing statistics of ads and queries. They use a generalized linear model with a cumulative Gaussian link function on the feature weights in order to estimate the click probability via message passing.

Zhang and Jones [2007] examine the correlation of features of query rewrites with the metric of their definition, i.e. ad clicks over expected ad clicks. They compare features which are predictive of relevance and clicks in user query logs. They show that features like the rank position of ads have a high correlation with ad clicks. Regelson and Fain [Regelson and Fain 2006] estimate the clickthrough rate of new ads on a keyword basis by using the clickthrough rates of existing ads with the same bid terms or topic clusters.

These existing efforts study ad clickthrough based on factors related to ad content, landing pages, and bid terms. In a more recent study by Fan and Chang [2010], the authors address the problem of incorporating context into content-based ad placement strategies on blog pages. They argue that ads that conflict with the negative orientation of a blog page are less likely to result in click. Hence, even if the ad’s content matches with the content of a blog page, it should not be displayed on the page that mostly discusses the corresponding commercial product or service from a negative point of view. In the sponsored search domain, Ashkan and Clarke [2013] study the relationship between user click behavior and the context of search engine result pages and corresponding queries. They propose models to predict the aggregated click behavior of ads by considering the influence of such factors as the location of ads and the rank of ads, along with the query intent category. In this paper, we augment context
by adding biases that are based upon known properties of sponsored search, and upon the differences that exist in user browsing behavior over advertising links as opposed to organic links.

All these studies also adopt the common assumption of a trust bias [Joachims et al. 2005] for higher ranked results, which has been confirmed through eye-tracking studies [Joachims et al. 2007] and other methods. The trust bias provides a basis for the examination hypothesis proposed by Richardson et al. [2007]. Under this hypothesis, we assume that the clickthrough rate decreases towards lower positions due to reduced visual attention from the user. Jansen et al. [2007] report that the link examination behavior of users is similar to their click behavior, with users preferring top-listed results for both sponsored and non-sponsored links. Hence, two main factors influencing an ad’s clickthrough rate under this examination model are the relevance of the individual ad to a user’s commercial need and the rank position of the ad on the page.

A more intuitive approach considers the influence of the co-appearance of ads, which has been explored from different perspectives in literature. For instance, [Xiong et al. 2012] show empirically that the clickthrough rate of an ad is influenced by the similarity of the ad and its surrounding ads. The reason may be due to the similarity in the content or topics of these surrounding ads and the given ad, which creates a competitive situation and distracts user’s attention away from the ad. Hence, the authors propose a model for ad click prediction that is based on continuous conditional random fields (CRF) using ad content features and their similarity with the other ads displayed at the same time.

Another perspective, which mainly targets the user’s browsing behavior, considers the quality of ads that appear prior to a given ad on the result page. A relatively strong and compelling ad appearing before a second ad may distract the user from the second ad, regardless of its quality. On the other hand, a weak ad may annoy the user into abandoning the list, regardless of the quality of the rest of the list. This concept is introduced by Ghosh and Mahdian [Ghosh and Mahdian 2008] as the externality effect of ads in advertising. In considering the externality effects, they work from a linear browsing assumption, in which users visually scan the ad list from top to bottom [Aggarwal et al. 2008]. Once an ad is examined by the user, ad-specific factors (e.g., relevance of the ad as perceived by the user) will determine the click decision and continuation probability. In contrast to the examination model, the probability of clicking on an ad depends on the content of the ads shown above it on the page.

In this paper, we adopt a linear browsing assumption that considers the externality effect of ads. This assumption is related to the cascade model of user behavior [Craswell et al. 2008] that is common in organic search. Under this model, an ad is examined only if the user first scans over all the previously displayed ads. In the next section, we briefly review the cascade model, and a few other models based upon it.

### 2.2. Browsing and Click Models

The cascade model of user behavior [Craswell et al. 2008] was originally proposed in the context of organic search. Given that a user issues query $q$, the binary hidden variable $E_i$ indicates whether the user examines the document $d_i$ displayed at rank $i$ in the result list. Similarly, $C_i$ is defined as a binary variable representing whether the user clicks on $d_i$ given they viewed its caption.

According to the cascade model, the probability of examining $d_i$ (i.e., $P(E_i = 1)$) is known as the examination probability, which we assume to be dependent on the quality of the documents shown prior to it (i.e., listed at earlier ranks) on the page:

$$P(E_i = 1) = \prod_{j=1}^{i-1} (1 - \omega_{d_j}^q)$$
where $\omega_d^q$ represents the attraction probability of the document $d_j$ with respect to the query $q$. This probability represents the chance that the user perceives $d_j$ to be relevant to their information need and clicks on it, given they examined it. The attraction probability is sometimes known as perceived relevance [Chapelle and Zhang 2009].

The cascade model makes the following assumptions about user browsing and click behavior: i) the user performs a linear scan of the result list starting from the top ii) there is at most one click per search; hence, the model cannot explain multiple clicks, and iii) if the user does not click on a viewed link, they continue examining links, i.e., the user is infinitely persistent:

$$P(E_1 = 1) = 1$$
$$P(E_{i+1} = 1|E_i = 0) = 0$$
$$P(C_i = 1|E_i = 1) = \omega_d^q$$
$$P(E_{i+1} = 1|E_i = 1, C_i = c_i) = 1 - c_i$$

(1)

where $c_i$ presents the click event value (0 or 1) for the variable $C_i$.

There are other biases and factors addressed in related research work. The user browsing model (UBM) [Dupret and Piwowarski 2008] and Bayesian browsing model (BBM) [Liu et al. 2009] are among the click models that do not employ the cascade assumption. They extend the examination hypothesis by considering the dependency on the positional distance to the previous click in the query session. The task-centric click model (TCM) [Zhang et al. 2011] considers the sequences of queries and clicks in a session as a task and characterizes user behavior related to a task as a collective whole. They formalize user behavior with respect to two biases; one is query reformulation and the other is the user’s desire for unseen documents in a session. Hu et al. [2011] study the impact of query intent diversity on the existing click models. They argue that user click can not be explained only by the relevance and position of the document, but also by the diversity of the user’s queries. The whole page click model [Chen et al. 2011] differs from the previous approaches as the authors explore the whole search result page including all the click blocks (e.g. organic results, sponsored results, etc) on the page as an integrated entity. Their findings include that if there is a click in a given block, a user is less likely to examine the next block.

There have been further efforts to extend the cascade model through various hypotheses and assumptions, all aiming at modeling user browsing/click behavior in a more realistic fashion, in which multiple clicks are permitted. All of these models share a notion of user patience and persistence as they move from document to document. Zhu et al. [2010] define a group of user and URL specific attributes, such as query, browser type, local hour, and the position to model the relevance and examination transitions effects as random variables. The click chain model (CCM), proposed by Guo et al. [2009b], defines the transition probability from document $i$ to $i + 1$ in the cascade model through three global parameters. These parameters are fixed and independent of the users and URLs.

The dynamic Bayesian network (DBN) model, proposed by Chapelle and Zhang [2009], defines a persistence factor that is assumed to be fixed and shared across query sessions. According to DBN, a user starts from the first document and keeps on examining $d_{i+1}$, given they already examined $d_i$ in two cases: i) either they do not click on $d_i$ and skip it with a probability of $\lambda$, or ii) they click on $d_i$ and find it un-satisfying (non-relevant), so they move on to the next document with a probability of $\lambda$. In both cases, a binary variable $S_i$, which indicates the satisfaction status, becomes 0. In case of a click on $d_i$, this variable will be set to 1 if and only if $d_i$ satisfies the user. Here, $\nu_d^q$
represents the satisfaction probability, also known as the post-click relevance:

\[
P(E_1) = 1 \\
P(E_{i+1} = 1|E_i = 0) = 0 \\
P(E_{i+1} = 1|E_i = 1, S_i = 0) = \lambda \\
P(S_i = 1|C_i = 1) = \nu^d_i
\] (2)

where \(\lambda\) represents the persistence of the user in browsing, and it is considered to be a fixed parameter that is constant across all query sessions.

The dependent click model (DCM) [Guo et al. 2009a] is also based on the cascade assumption but it models the user persistence in a different fashion. Here, a position-dependent form of the \(\lambda\) parameter, denoted by \(\lambda_i\), is defined as the chance that the user would be willing to see more results after a click at position \(i\). It is assumed that the user starts from the first document and continues examining the next document with a probability that depends on their click action at rank \(i\). The next document is examined with a probability of one or \(\lambda_i\) given that the user skips or clicks the document at rank \(i\), respectively:

\[
P(E_1 = 1) = 1 \\
P(E_{i+1} = 1|E_i = 0) = 0 \\
P(E_{i+1} = 1|E_i = 1, C_i = c_i) = \lambda_i^c
\] (3)

where the maximum likelihood estimate of the \(\lambda_i\) values is empirically computed for the various positions on the result pages.

Following from the related work in click and browsing modeling in Web search, Ashkan and Clarke [2012] propose query biases in the domain of sponsored search in order to better cope with the actual user behavior in this domain. We extend the previous effort by introducing and formulating the location bias in sponsored search as well as the query biases. The proposed query- and location-aware browsing/click model and the inference of the parameters of the model are detailed in this paper. We also extend the experimental studies in order to further explore the impact of various combinations of the biases on a group of cascade-based click models, and to explore whether the introduced biases reflect varying behavioral patterns for different users.

3. MODELING BROWSING BEHAVIOR IN TERMS OF LOCATION AND QUERY BIASES

As the primary goal of our work, we aim to exploit characteristics of user behavior to improve user browsing models and click prediction. One of these characteristics is a user bias against sponsored search [Jansen and Resnick 2006], leading us to study the impact of variability in user persistence in the sponsored search domain. Another characteristic is the user’s response to the page structure, specifically the locations on the page where ads appear, such as the top and side, and the ordering or ranking of ads at each location.

Building on related work, we develop models that incorporate location and query biases. We re-visit a group of well known click models from the organic search domain, adapting and extending them to the sponsored search domain with respect to these biases. As a particular example, we adapt and extend the DBN [Chapelle and Zhang 2009] click model. Under our approach, we model the set of ads displayed at each location through a separate dynamic Bayesian network, with an extended version of DBN [Chapelle and Zhang 2009] dedicated to each location. Other models could be extended in a similar fashion, and we generally refer to models that accommodate query and location biases as location- and query-aware browsing/click models.
3.1. Location Bias
In contrast to organic links, ad links often do not strictly appear one after the other on a result page. In addition to a rank position, ads can be characterized according to the location on the page on which they appear, typically the top (north) or the right (east) side. Various user studies [Jansen 2007; Jansen and Resnick 2006] suggest that users expect top-listed ads to be more relevant than side-listed ads, and therefore they are examined by the user more frequently.

In order to provide additional evidence for this suggestion, in Figure 1 we plot the relative click rate for different locations and positions of ads on result pages. These statistics have been collected from the logs of a commercial search engine, and full details about these logs is provided in Section 4.1. Relative click rates are computed from the clicks recorded for a set of SERPs with eight ads such that three of them appear at the top of each page (denoted by the first three positions on the X axis) and five of them appear at the side of the page (denoted by the last five positions on the X axis). For the first position of the top-listed ads, a click rate of one is assigned in the plot, while for the rest of positions and locations the relative click rate to that of the first position at the top are reported.

At first glance, the impact of the position bias on click rates is obvious. The frequency of clicks decreases towards lower rank positions, possibly due to reduced visual attention from the user. However, our main intention here is to show the significant drop of click rate starting from position 4 which represents the first rank of the side-listed ads. This evidence confirms the impact of ad location on user behavior.

Let $N_t$ and $N_s$ be the number of displayed ads on the top and at the side of a result page, and define $N = N_t + N_s$. An ad displayed at the rank position $i$ placed on the location $l$ of the page is denoted by $a_{l,i}$. Ads displayed at the top are denoted according to their ranking: $a_{t,1},..., a_{t,N_t}$. Similarly, the side-listed ads are denoted as $a_{s,1},..., a_{s,N_s}$. For a result page with a total of $N$ ads displayed for the query $q$, the following variables also are defined to model various characteristics of the user, the query, and the displayed ads:

- $U_l$ is a binary hidden variable representing whether or not the user initiates browsing at location $l$.
- $E_{l,i}$ is a binary hidden variable indicating whether the user examines the ad at rank $i$ of location $l$. 

Fig. 1: Relative click rate for different locations/positions of the result pages.
- $C_{l,i}$ is a binary variable representing the click observation at rank $i$ of location $l$.
- $A_{l,i}$ is a binary hidden variable reflecting the user’s perceived relevance of the ad displayed at rank $i$ of location $l$.
- $S_{l,i}$ is a binary hidden variable reflecting the user’s satisfaction (post-click relevance) once they click on the ad displayed at rank $i$ of location $l$.

For all these variables, $l$ represents the location of the ad displayed on the page which may be $t$ (for the top-listed ads) and $s$ (for the side-listed ads). Also note that $U_l$ and $E_{l,i}$ are properties of the query (thus, properties of the user who issues the query) and of the location, which means they are shared across result pages with the same query. The sequence of click observations obtained from the page with respect to its ads represent $C_{l,i}$ values. Finally, $A_{l,i}$ and $S_{l,i}$ are considered to be properties of ads with respect to the query for which they are displayed, so they are defined over ad-query pairs.

### 3.2. Query Bias

The query itself represents an important aspect of ad context, which can significantly impact the expected click rate. Even though terms in the query act as triggers for ad selection, the nature of the query, and the user intent underlying the query, still plays a major role.

To some extent, information about user intent can be inferred from the organic results appearing on the search result page [Ashkan 2013]. If the query is commercially oriented (e.g., the user is intending to purchase a product or service) they may be more likely to click on an ad. A weakly related ad appearing with the results of a commercially oriented query may receive more clicks than a strongly related ad appearing with the results of a less commercially oriented query. On the other hand, if there is no commercial intent underlying the query, the user may not consider the ads at all.

To provide an insight into user intent, we asked a group of annotators to judge the presumed intent of a sample of 4000 queries from the perspective of a general user as either mostly commercial or mostly non-commercial. For the purpose of this judging exercise, we define a commercial query as a query with the underlying intention to make an immediate or future purchase of a product or service, while anything else falls into the noncommercial category. The annotation process is further explained in Section 5.1, while more details can be found in [Ashkan 2013]. We matched the annotated queries against a search engine logs to collect SERPs that belong to these queries and contain at least one ad. It is interesting to note that out of the SERPs that belong to the commercial queries, 71% did not end up with a click on any ad, whereas for the non-commercial queries 90% did not end up with a click.

Within the category of commercial queries, we may also consider finer levels of query intent. For instance, consider the query “cheap airline ticket” versus the query “American airline ticket”. Both are commercial queries, making their result pages appropriate for sponsored links. The first query targets a specific commercial product, i.e., an air ticket, without regard for the brand, while the second targets a special brand of this product, i.e., an American Airlines ticket. One could argue the user who enters the first query would be more engaged in the browsing process, since any airline might do. On the other hand, the user who issues the second query may be a relatively loyal user who is looking for their favorite brand among the top results. If they do not find such results among the top-listed sponsored links, they may either abandon the search or move on to the organic results [Jansen and Resnick 2006]. In other words, a user who issues a query for a specific product, with no specific brand in mind, may have a greater tendency to scan through the entire sponsored links as opposed to a user who is looking for a specific brand.
As the query intent tends to vary across multiple dimensions, we must formulate parametric biases that depend on user queries. Two types of query bias are introduced in this work, as factors involved in browsing behavior. The first query bias deals with the initiation of browsing over advertising links, while the second one reflects user persistence in browsing advertising links.

**Initiation:** The trigger point to start examining an ad list should matter more for the sponsored results as opposed to the organic results. The reason goes back to the user’s bias against sponsored links, as compared to organic links [Jansen and Resnick 2006], where users appear to have a greater tendency to examine organic links rather than the sponsored links. Hence, we formulate the following hypothesis for the way a user targets a list of ads at particular location of a result page:

$$U_l = 1 \iff E_{l,1} = 1$$  (4)

Therefore, the initiation probability $u_q^l$ is defined at the query-level and the location-level, where $q$ represents the query and $l$ represents the location. The initiation probability $u_q^l$ represents the chance that the user will eventually initiate browsing ads listed at location $l$ of the page, i.e., $P(U_l = 1) = u_q^l$. Using this definition and Equation 4, given query $q$, the probability of examining the first ad at location $l$ can be calculated as follows:

$$P(E_{l,1} = 1) = P(U_l = 1) = u_q^l$$  (5)

Once a user starts browsing ads listed at a particular location, their browsing persistence can be addressed using a variation of the cascade model, as explained next.

**Persistence:** Newer versions of the cascade model [Chapelle and Zhang 2009; Guo et al. 2009b,a; Zhu et al. 2010] share a notion of the user’s patience or persistence ($\lambda$) as they move from link to link. In order to model the relationship between user persistence and the associated query, this paper introduces variability into the persistence parameter $\lambda$. We assume that different users have different levels of patience in browsing through an ad list. As a result, the persistence parameter for the DBN model in Equation 2 can be revised to $\lambda_q^l$ in order to take into account the user’s query ($q$) and the ad’s location on the page ($l$). Our extended user browsing and click model for ads is depicted in Figure 2. Note that this model accounts for both the location bias and the query bias with respect to a user's initiation of browsing and also for query-dependent persistence in browsing. Given query $q$, if the user initiates browsing (see Eq. 5), they begin examining the list of ads displayed at location $l$. If the user examines ad $a$ displayed at the rank position $i$ at this location, two scenarios are possible: click or no click. If they do not click on the ad (i.e. $C_{l,i} = 0$), they either move on to the next ad with probability $\lambda_q^l$, or they abandon their search with probability $1 - \lambda_q^l$:

$$P(E_{l,i+1} = 1|E_{l,i} = 1, C_{l,i} = 0) = \lambda_q^l$$  (6)

On the other hand, the user may examine the ad and perceive it to be relevant to their commercial need (i.e. $A_{l,i} = 1$) with the attraction probability of $\omega_q^a$ and click on it (i.e. $C_{l,i} = 1$). The perceived relevance may depend on the presentation quality of the ad, such as the ad’s header or its creative, and of course on the user’s understanding of the relevance of the ad to their query. The click probability at rank $i$ given the examination state can therefore be formulated as:

$$P(C_{l,i} = 1|E_{l,i} = 1) = P(A_{l,i} = 1) = \omega_q^a$$  (7)

$$P(C_{l,i} = 1|E_{l,i} = 0) = 0$$
Fig. 2: The location- and query-aware browsing model.

Clearly, when there is no examination, there is not going to be a click according to the cascade assumption. Three scenarios are likely upon a click:

1. With a probability of \( \nu_a \), the user may be satisfied by the content of the landing page and exit the page:

\[
P(S_{l,i} = 1 | C_{l,i} = 1) = \nu_a
\]
\[
P(E_{l,i+1} = 1 | S_{l,i} = 1) = 0
\]

2. With a probability of \((1 - \nu_a)\lambda_l\), they may be unsatisfied by the landing page and move on to the next ad:

\[
P(S_{l,i} = 0 | C_{l,i} = 1) = 1 - \nu_a
\]
\[
P(E_{l,i+1} = 1 | E_{l,i} = 1, S_{l,i} = 0) = \lambda_l
\]

3. With a probability of \((1 - \nu_a)(1 - \lambda_l)\), they may abandon their search.

\[
P(E_{l,i+1} = 0 | E_{l,i} = 1, S_{l,i} = 0) = 1 - \lambda_l
\]

Here, \( \nu_a \) is the satisfaction probability representing the probability that the user, who clicked on ad \( a \) and viewed its landing page, finds the ad satisfactory with respect to their query \( q \).
3.3. Parameter Inference

In this section, we examine the inference algorithm for our extended version of the DBN model. The algorithm is essentially the same for any other cascade-based click model, once extensions are applied to accommodate the location bias and the two types of query bias. The inference algorithm finds the maximum likelihood estimates of the parameters set \( \theta = (u_l^q, u_t^q, \lambda_l^q, \lambda_t^q, \omega_s^q, \nu_u^q) \) corresponding to the hidden variables of the model. There is no known way to analytically solve for the model which maximizes the expected complete-data log-likelihood function (M-step). We consider the initial estimate of all the parameters to be 0.5 for the first iteration, reflecting a random initial value of 50% for each parameter. In what follows, we explain the inference procedure for the model’s parameters. Note that the superscript \( j \) added to the variables, originally introduced in Section 3.1, indicates the SERP \( j \) to which these variables belong.

**Initiation Probability:** As for \( u_l^q \) and \( u_t^q \), the posterior distribution of their corresponding hidden variables is calculated in the E-step of each iteration. We explain the analysis for \( u_l^q \) in the general form which can be expanded to the top-listed ads and the side-listed ads by substituting \( t \) with \( t \) and \( s \), correspondingly. Define \( \phi(U_l^j) \) as the posterior distribution of the variable \( U_l^j \) given the click sequence \( C_l^j \) observed on the location \( l \) of the SERP \( j \) and the current value of the parameter \( u_l^q \):

\[
\phi(U_l^j = 1) = P(U_l^j = 1|C_l^j, u_l^q) = \begin{cases} u_l^q & \text{if there is no click} \\ 1 & \text{otherwise} \end{cases}
\]

For a given SERP of query \( q \), if no ad click is observed over the ads listed at location \( l \) of the page, there is a probability of \( u_l^q \) that the user started examining these ads, and a probability of \( 1 - u_l^q \) that the user skipped the list in the first place. On the other hand, if any click is recorded for this location, the cascade assumption implies that the user considered the ad list and started examining from the first ad; hence, the posterior probability becomes 1 (See Eq. 4).

Given the posterior distribution \( \phi(U_l^j) \) for the \( j \)th SERP of \( q \) (1 ≤ \( j \) ≤ \( M \)), the expected complete-data log-likelihood function \( Q(u_l^q, u_l^q(k)) \) [Dempster et al. 1977] at the iteration \( k \) of the inference procedure can be computed as follows:

\[
Q(u_l^q, u_l^q(k)) = \sum_{j=1}^{M} P(U_l^j = 0|C_l^j, u_l^q(k)) \log(1 - u_l^q) + P(U_l^j = 1|C_l^j, u_l^q(k)) \log(u_l^q)
\]

\[
= \sum_{j=1}^{M} \phi(U_l^j = 0) \log(1 - u_l^q) + \phi(U_l^j = 1) \log(u_l^q)
\]

where \( u_l^q(k) \) denotes the current value of the parameter \( u_l^q \), and it is used for computing \( \phi(U_l^j) \) according to the Equation 10.

The expected complete-data log-likelihood function can then be locally maximized by solving for the partial derivative of the function with respect to the parameter \( u_l^q \) to
be 0 at this iteration:

$$\frac{\partial Q(u_q^l, u_q^{l(k)})}{\partial u_q^l} = 0$$

(12)

$$\Rightarrow \sum_{j=1}^{M} \phi(U_{j}^l) = 0 \frac{-1}{1 - u_q^l} + \phi(U_{j}^l = 1) \frac{1}{u_q^l} = 0$$

$$\Rightarrow \sum_{j=1}^{M} (1 - \phi(U_{j}^l = 1)) \frac{-1}{1 - u_q^l} + \phi(U_{j}^l = 1) \frac{1}{u_q^l} = 0$$

$$\Rightarrow \sum_{j=1}^{M} \phi(U_{j}^l = 1) - u_q^l \frac{u_q^l(1 - u_q^l)}{u_q^l} = 0$$

$$\Rightarrow u_q^l = \frac{\sum_{j=1}^{M} \phi(U_{j}^l = 1)}{M}$$

Therefore, after the M-step of the algorithm at the iteration $k$, the value of the parameter $u_q^l$ will be updated to the above value that locally maximizes the $Q$ function:

$$u_q^{l(k+1)} = \arg\max_{u_q^l} Q(u_q^l, u_q^{l(k)}) = \frac{\sum_{j=1}^{M} \phi(U_{j}^l = 1)}{M}$$

(13)

**Persistence Probability:** One of the major challenges in the inference procedure relates to the variability of the persistence probability. As stated before, prior models incorporate a fixed parameter that is usually obtained from domain knowledge. However, we assume persistence varies as follows:

The persistence in transitioning from ad to ad depends on the user’s query and on the location of the ad list. Hence, two parameters $\lambda_{q,t}$ and $\lambda_{q,s}$ are defined to reflect the user’s query $q$ and the location $l$ on the page (i.e. either $t$ or $s$ representing the top and the side locations respectively).

The analysis is explained for the general form of parameter $\lambda_{q,t}$, where $\phi(E_{i+1,j}^l)$ is defined as the posterior distribution of variable $E_{i+1,j}^l$ for the SERP $j$, given the observed click sequence $C_{j}^l$ at the location $l$, the previous examination state at rank $i$, and the current value of $\lambda_{q,t}$:

$$\phi(E_{i+1,j}^l = 1) = \frac{P(E_{i+1,j}^l = 1|C_{i}^l, E_{i,j}^l = 1, \lambda_{q,t}^l)}{P(E_{i,j}^l = 1, C_{i,t}^l, \lambda_{q,t}^l)}$$

$$= \frac{P(E_{i+1,j}^l = 1, E_{i,j}^l = 1, C_{i,j}^l, \lambda_{q,t}^l)}{P(E_{i,j}^l = 1, C_{i,j}^l, \lambda_{q,t}^l)}$$

$$= \frac{P(E_{i+1,j}^l = 1, E_{i,j}^l = 1|C_{i}^l, \lambda_{q,t}^l)P(C_{i}^l|\lambda_{q,t}^l)}{P(E_{i,j}^l = 1|C_{i,j}^l, \lambda_{q,t}^l)P(C_{i,j}^l|\lambda_{q,t}^l)}$$

$$= \frac{P(E_{i+1,j}^l = 1, E_{i,j}^l = 1|C_{i}^l, \lambda_{q,t}^l)}{P(E_{i,j}^l = 1|C_{i,j}^l, \lambda_{q,t}^l)}$$

(14)

We note that the posterior distribution for the transition variable $E_{i+1,j}^l$ is computed at the presence of the previous examination state that occurs at rank $i$ (i.e. $E_{i,j}^l$). This
is due to the cascade assumption which implies there is no continuation in transition at the rank position \( i + 1 \) unless \( E_{l,i+1}^j = 1 \).

The numerator and denominator of Equation 14 are computed according to Equation 24 using the forward-backward algorithm described in the Appendix.

In each iteration, the posterior distribution is computed for the possible transitions across the existing rank positions of the location \( l \) of each SERP \( j \). The value of the persistence parameter \( \lambda_l^q \) at the \( k^{th} \) iteration can then be updated by maximizing the expected complete-data log-likelihood function:

\[
\lambda_l^q(k+1) = \arg\max_{\lambda_l^q} Q(\lambda_l^q, \lambda_l^{(k)})
\]

\[
= \arg\max_{\lambda_l^q} \sum_{j=1}^M \sum_{i=1}^{N_l^j-1} [P(E_{l,i+1}^j = 0|C_{l,i}^j, E_{l,i}^j = 1, \lambda_l^{(k)}) \log(1 - \lambda_l^q) + P(E_{l,i+1}^j = 1|C_{l,i}^j, E_{l,i}^j = 1, \lambda_l^{(k)}) \log(\lambda_l^q)]
\]

\[
= \arg\max_{\lambda_l^q} \sum_{j=1}^M \sum_{i=1}^{N_l^j-1} \phi(E_{l,i+1}^j = 0) \log(1 - \lambda_l^q) + \phi(E_{l,i+1}^j = 1) \log(\lambda_l^q)
\]

\[
= \frac{\sum_{j=1}^M \sum_{i=1}^{N_l^j-1} \phi(E_{l,i+1}^j = 1)}{\sum_{j=1}^M N_l^j - 1}
\]

where \( N_l^j \) is the number of ads appearing at the location \( l \) of the SERP \( j \), resulting in \( N_l^j - 1 \) possible transitions between ads at this location. The \( \arg\max \) function, similar to the case for the initiation parameters, finds the local maximum by taking the partial derivative of the expected complete-data log-likelihood function with respect to the \( \lambda_l^q \) parameter.

**Relevance parameters:** Parameters \( \omega_a^q \) and \( \nu_a^q \) can be estimated in the same fashion as Chapelle and Zhang [2009], but taking different locations of the page into account. We skip the details of the analysis here and provide only the final answer for each. The update formulas for these parameters in the \( k^{th} \) iteration are as follows:

\[
\omega_a^q(k+1) = \frac{\sum_{j=1}^M \sum_{l \in \{t,s\}} \sum_{i=1}^{N_l^j} I(a_{l,i}^j = a) \phi(A_{l,i}^j = 1)}{\sum_{j=1}^M \sum_{l \in \{t,s\}} \sum_{i=1}^{N_l^j} I(a_{l,i}^j = a)}
\]

\[
\nu_a^q(k+1) = \frac{\sum_{j=1}^M \sum_{l \in \{t,s\}} \sum_{i=1}^{N_l^j} I(a_{l,i}^j = a, C_{l,i}^j = 1) \phi(S_{l,i}^j = 1)}{\sum_{j=1}^M \sum_{l \in \{t,s\}} \sum_{i=1}^{N_l^j} I(a_{l,i}^j = a, C_{l,i}^j = 1)}
\]

where \( I(.) \) is a binary indicator function, such that \( I(a_{l,i}^j = a) = 1 \) if the \( i^{th} \) ad placed on the location \( l \) of the \( j^{th} \) SERP is \( a \), and is 0 otherwise. Similarly, \( I(a_{l,i}^j = a, C_{l,i}^j = 1) = 1 \) if there was also a click on this ad, and 0 otherwise. Finally, \( \phi(A_{l,i}^j) \) and \( \phi(S_{l,i}^j) \) are the posterior distributions of variables \( A_{l,i}^j \) and \( S_{l,i}^j \), respectively, which are computed in a similar fashion as in the original DBN [Chapelle and Zhang 2009] but adopted to account for separate locations on the page.
4. EXPERIMENTAL STUDY

Our experimental study is based on sponsored search data [MSR Data 2007] from Microsoft adCenter sampled over a few months in 2007. We selected two subsets of the data for evaluation purposes, which proceeds in two steps, as follows:

1. We evaluate the impact of the query biases, user initiation and persistence, on user browsing and click behavior.
2. We evaluate the impact of the location bias and the result page structure, on user browsing and click behavior, along with the query biases.

4.1. Data Set

A subset of the data is selected for the first step of the evaluation. We refer to this subset as set \( A \) in the remainder of the paper. There are approximately 41M SERPs in this set which were sampled over three months. These SERPs correspond to a sample of about 2.8M queries, all of which were required to have at least one ad click recorded in the log. As also stated in Zhu et al. [2010], for the SERPs with multiple clicks, it is assumed that the clicks order is based on the cascade assumption, i.e. from top to bottom.

The location of the ads is not recorded in the data set which could create ambiguity for the location-based analysis. For this reason, in the second round of experiments (concerning the location bias experiments), a sample of search result pages needs to be used such that the location of ads can be certainly identified. This is possible due to the ad placement strategy of the search engine back in 2007, i.e. at most three ads were displayed on the top and a maximum of five ads were displayed at the side. Given this constraint, and assuming that the ads displayed on the top are ranked higher than the ones displayed at the side, we can be assured that the SERPs with eight ads belong to result pages with three top-listed ads and five side-listed ads. All the SERPs with eight ads in the set \( A \) have therefore been placed into a second set, called \( B \), for evaluation with respect to the second step. Approximately 7M SERPs appear in the set \( B \), which corresponds to a sample of about 712K queries.

In each round of experiments, using either set \( A \) or set \( B \), one pass is made over the SERPs and the corresponding ad impressions in order to do online learning and testing. In each case, before taking the current impression into account for training purposes, we make a prediction of the click probability of its ads using the values of the parameters obtained from the previously observed impressions. Once the prediction is finished for the current SERP, its actual click signals are added to the rest of the training samples to perform another round of inference. One of the benefits of this process is its ability to perform online learning and testing. Therefore, the contextual factors for any pair of query and ad keep getting updated as more instances of the query and ad appears in the log.

4.2. Click Models

A group of cascade-based click models from the organic search domain are borrowed as the baselines for the experiments conducted in this section. The reason is due to the nature of browsing behavior studied in this paper which targets the linear browsing of users through the result list. We also consider one of the traditional click models from both organic and sponsored search domains, the logistic regression (LR) model [Craswell et al. 2008; Richardson 2008]. The reason to choose this model is that, apart from being among the primary click models, it can be implemented with respect to the factors considered in this work that do not relate to the features extracted directly from the content of ads, bid terms, or the client side information. We adopt the LR implementation of Chapelle and Zhang [2009] to our domain, such that the
click probability is considered as a function of the ad and its position. It is noted that the LR model is among the click models that are based on the examination hypothesis [Richardson et al. 2007] of user behavior.

The cascade (CAS) model [Craswell et al. 2008] has been selected as another baseline of the analysis. Under this model, the user is assumed to be infinitely persistent, continuing their examination unless they make a click (see Eq. 1). In addition, the dynamic Bayesian network (DBN) model [Chapelle and Zhang 2009] has been selected as it is one of the state-of-the-art click models in which both the perceived relevance and the post-click relevance, and also the user’s persistence, are modeled. However, the persistence probability is assumed to be constant across all sessions in DBN (see Eq. 2).

Finally, the dependent click model (DCM) [Guo et al. 2009a] has been selected as one of the variations of the cascade model in which a notion of varying persistence probability has been considered. However, the assumption of an infinitely persistent user is still partially included in this model, as the user is assumed to examine a document with a probability of one given they have not clicked on the previously examined document. There is a position-dependent parameter defined in this model, representing the probability that the user would be willing to see more results after a click (See Eq. 3). However, this transition probability is assumed to be the same across all queries.

None of these click models require extra information about the client side, ad content, or bid terms, which enables us to reproduce them on our log data. For each cascade-based click model, we evaluate the performance of the model in the presence of the biases that were introduced in this work. Our goal is to understand whether any of these models and settings are able to better predict user behavior on advertisement links. The first group of settings reflect the query biases (initiation probability and persistence). The second group of settings reflects the query biases and the location biases together.

4.3. Click Prediction

As discussed before, an ad is assumed to be clicked if and only if the user gets the chance to view the ad within the context that it is shown and finds it relevant to their commercial need. Therefore, given a new search result page $j$ for the query $q$, we predict the probability of click for a given ad $a$ displayed at rank $i$ on the location $l$ of the SERP as follows:

$$P(C^j_{il} = 1) = P(C^j_{il} = 1|E^j_{il} = 1)P(E^j_{il} = 1)$$

$$= \omega^q_a P(E^j_{il} = 1)$$

(16)

where the attraction probability $\omega^q_a$ is substituted according to Equation 7, and $P(E^j_{il} = 1)$ is estimated according to Equation 22 from Appendix using the forward-backward algorithm [Rabiner 1989] and with respect to the parameters of the model that has been learned throughout the inference procedure. Incorporating $P(E^j_{il} = 1)$ into the click probability model assures that the quality of the ad is evaluated in the context of preceding ads; satisfying one of the major goals of our work.

The above equation belongs to the model augmented with the location and query biases in the previous section. For the original models (i.e. CAS, DCM, and DBN) and the models within the other settings, a simpler form of this equation can be created in the same way (for instance, the parameter $l$ should be removed from the equation for the models that do not implement the location bias). For the LR model, the probability of click is a function of the ad and of the position. Using the logistic function ensures this probability is always between 0 and 1.
4.4. Evaluation Metrics

We evaluate the performance of the models within different settings based on two standard metrics: average log-likelihood and average perplexity. For each model, the average log-likelihood (LL) is calculated with respect to the predicted click probability for each ad and the actual click signals observed in the log. For instance, the average log-likelihood for a click model augmented with the location bias is formulated as:

$$\text{LL} = \frac{\sum_{j=1}^{M} \sum_{l \in \{t,s\}} \sum_{i=1}^{N_j^l} c_{j,l,i} \log P(C_{j,l,i} = 1) + (1 - c_{j,l,i}) \log(1 - P(C_{j,l,i} = 1))}{\sum_{j=1}^{M} N_j^t + N_j^s}$$  \hspace{1cm} (17)

where $c_{j,l,i}$ represents the actual click signal recorded for the ad displayed at the rank position $i$ of the location $l$ of the SERP $j$, and $P(C_{j,l,i} = 1)$ is the probability of click that the model predicts for this ad (Eq. 16). $N_j^l$ indicates the number of ads displayed at location $l$ of the SERP $j$. A larger value of LL indicates a better model fit, where the ideal value is zero.

To further study the impact of the biases at different rank positions, the perplexity measure is used as the log-likelihood powers computed independently at each rank position [Zhang et al. 2011]. For instance, the average perplexity for the rank position $i$ of location $l$ can be computed as follows:

$$\rho_{l,i} = \exp\left(-\frac{1}{M_{l,i}} \sum_{j=1}^{M_{l,i}} c_{j,l,i} \log P(C_{j,l,i} = 1) + (1 - c_{j,l,i}) \log(1 - P(C_{j,l,i} = 1))\right)$$  \hspace{1cm} (18)

where $M_{l,i}$ denotes the number of pages in the log that displayed an ad at the rank position $i$ of the location $l$. A lower value of the perplexity indicates a better fit between the model and the actual data. The details of the experiments based on each group of settings, along with their results, are presented next.

4.5. Evaluating the Impact of Query Biases

Settings that reflect different combinations of the two query biases are categorized and labeled in order to be referenced easily throughout the experiments. Their description can be found in Table I. The formulation and inference of all the extended models are similar to those explained in the previous section. These settings do not reflect the location bias since they are evaluated on the larger sample of the data, set $A$, which includes various number of ads on the search result pages.

Table I: The settings related to the query biases for the cascade-based models under the experiment.

<table>
<thead>
<tr>
<th>settings</th>
<th>click models</th>
</tr>
</thead>
<tbody>
<tr>
<td>original setting of the model</td>
<td>CAS, DCM, DBN (0.9, 0.5, 0.1, 0.001)</td>
</tr>
<tr>
<td>varying persistence</td>
<td>CAS+VP, DCM+VP, DBN+VP</td>
</tr>
<tr>
<td>varying persistence and initiation</td>
<td>CAS+VPI, DCM+VI, DBN+VPI</td>
</tr>
</tbody>
</table>

Note that only for DCM, the variability of persistence factor is not accommodated in the model since DCM itself has a notion of variability in user persistence (as described before). Hence, the browsing initiation probability is the only factor added to DCM in this round of experiments.
In addition, as the persistence factor (\(\lambda\)) is fixed in the original DBN model (see Eq. 2), we train DBN using various values of this parameter. Thus, under the first setting of DBN in Table I, four values for \(\lambda\) are provided, each resulting in a separate run for DBN. Among these, the run with \(\lambda = 0.9\) represents the original DBN model with a relatively patient user while the rest, particularly the one with \(\lambda = 10^{-3}\), represents a relatively impatient user.

Table II: Runs for the LR model and across different settings of the query biases for the cascade model, DCM model, and DBN model.

<table>
<thead>
<tr>
<th>model</th>
<th>LR</th>
<th>Cascade</th>
<th>DCM</th>
<th>DBN</th>
</tr>
</thead>
<tbody>
<tr>
<td>setting</td>
<td>-</td>
<td>orig.</td>
<td>VP</td>
<td>VPI</td>
</tr>
<tr>
<td>run#</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

The LR model is also included in this round of experiments in order to compare its performance against the cascade-based models. Consequently, a total of 12 different runs are generated for evaluation purposes. Table II depicts these runs across different settings that are numbered from 1 to 12. Note that the LR model is based on the examination hypothesis of user behavior (as opposed to the cascade-based models) and therefore it can not be extended with respect to the proposed query biases.

![Fig. 3: The average log-likelihood for the 12 runs.](image)

The average log-likelihood for the runs are computed according to Equation 17 and plotted in Figure 3, where the run ids match those shown in Table II. Each group of models are plotted in a different pattern. It can be seen that all the cascade-based models and their variations outperform the LR model. This is in agreement with previous studies, such as [Craswell et al. 2008; Chapelle and Zhang 2009], confirming that the cascade assumption can better model user behavior as opposed to the examination hypothesis used in traditional models such as LR. It can be also seen that the overall performance of the DBN runs is relatively better than the other three groups of models, as expected.

A variation of user persistence is already included in the original DCM model, whereas the inclusion of varying user persistence provides improvements over the original settings of the other two models. As for the DBN model, the inclusion of a varying persistence probability (i.e. DBN+VP in run 11) shows improvement over the original DBN (i.e. DBN with \(\lambda = 0.9\) in run 7) and over DBN with \(\lambda = 0.5\) in run 8.
However, DBN+VP shows similar performance to that of the two settings of DBN with impatient users (i.e. DBN with $\lambda = 0.1$ in run 9 and DBN with $\lambda = 0.001$ in run 10). This may be due to the fact that the inclusion of the varying persistence probability is intended to reflect the lower persistence of the users in browsing through the advertisement links which is essentially the same as assuming the user to be impatient. Thus, runs 9, 10, and 11 are relatively comparable. One could argue that DBN+VP is able to reflect the impatience of the users in browsing through the advertisement links without assuming a fixed persistence probability for all users.

Runs 4, 6, and 12 substantially outperform their peers in cascade group, DCM group, and DBN group respectively, suggesting that modeling user initiation and persistence better reflects user browsing and click behavior in sponsored search. In other words, the inclusion of both user initiation probability and varying persistence in examination improves the performance of all the original cascade-based models.

In order to compare the performance of the models at different stages of online training, we considered the queries in the sequence of their SERPs as they appeared in the log and therefore in the learning process. Five intervals of frequency are depicted across the $x$–axis in Figure 4, each indicating the number of times that a query has been seen before the current prediction performance is calculated and averaged across

![Fig. 4: The average log-likelihood of the cascade-based click models with various settings of query biases and across different query frequencies.](image)
all the queries that fall in this query interval. Note that the last interval contains all
the queries in set $A$. Except for the early stage of the learning process (i.e. the first
interval), the performance of the models across the rest is pretty consistent. As can be
seen in these plots, for all the cascade-based models, the inclusion of both query biases
(i.e. CAS+VPI, DCM+VI, and DBN+VPI runs) results in substantial improvement in
the performance of the corresponding models across all the frequency groups.

4.6. Evaluating the Impact of the Location Bias in Addition to the Query Biases

In this section, we evaluate the performance of the cascade-based models under the
location bias, along with the query biases. For this purpose, set $B$ is used as it contains
only those SERPs from set $A$ for which eight ads are displayed. The location of these
ads is known: the top three appear on the top and the bottom five appear at the side of
the corresponding search result page.

Table III: The additional setting related to the location bias for the cascade-based mod-
els under the experiment.

<table>
<thead>
<tr>
<th>setting</th>
<th>click models</th>
</tr>
</thead>
<tbody>
<tr>
<td>varying persistence and initial motivation for different locations</td>
<td>CAS+VPIL</td>
</tr>
</tbody>
</table>

In addition to the settings listed under Table I, a new setting is considered for this
round of experiments, in order to study the impact of the location bias on the click
models. Table III depicts and labels this setting in order to be referenced along with
the ones introduced in Table I. We note again that the variability of the persistence
probability is not incorporated into DCM under this setting, since DCM itself has a
notion of variability in user persistence.

By adding the 3 runs labeled in Table III to the 11 cascade-based runs labeled in
Table I, a total of 14 runs are generated for this round of the experiments. These runs
are numbered from 1 to 14 as depicted in Table IV.

Table IV: Runs across different settings of the query biases and the location bias for
the cascade model, DCM Model, and DBN model.

<table>
<thead>
<tr>
<th>model</th>
<th>Cascade</th>
<th>DCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>setting</td>
<td>orig.</td>
<td>VP</td>
</tr>
<tr>
<td>run#</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>model</td>
<td>DBN</td>
<td></td>
</tr>
<tr>
<td>setting</td>
<td>$\lambda = 0.9$ (orig.)</td>
<td>$\lambda = 0.5$</td>
</tr>
<tr>
<td>run#</td>
<td>8</td>
<td>9</td>
</tr>
</tbody>
</table>

To summarize the performance of the 14 runs reported in Table IV, the average
log-likelihood (LL) across all the rank positions and locations are plotted in Figure 5.
The run ids in the figure match those shown in the table. Each group of models are
presented in a different pattern, highlighting the superiority of the performance of the
models extended with query and location biases comparing to their original setting in
each group.

Similar to the experiments performed on set $A$ and depicted in Figure 3, the overall
performance of the runs over set $B$ report relative superiority of the DBN runs over
the other two models in Figure 5. This result could be due to the more realistic way
of interpreting user behavior under the DBN model, e.g. the consideration of the per-
ceived and post-click relevance, which appears to reflect user behavior in sponsored
search as well. The numbers reported in Figure 5 are, generally speaking, lower than the overall performance values reported in Figure 3 due to the limitations of the set $B$. Remember that set $B$ has been chosen as a subset of set $A$ in order to be able to identify the precise location of ads on the SERPs.

To further study the impact of the introduced biases, especially the location bias, the perplexity measure is used as the log-likelihood powers computed independently at each rank position. A lower value of perplexity indicates a better fit between the model and the actual data. The plots in Figure 6 depict these results for each click model under the various settings of the biases. The performance of these models in their original settings is also plotted. We note that the first three rank positions on the X-axis represent the possible positions for the top-listed ads, while the last five (from 4 to 8) represent the side-listed ads.

Paired t-tests were carried out between the probability values obtained for the ads in each pair of runs at a significant level of 95%, which all resulted in p-values under 0.001, indicating the significance of the differences in the results.

As we can observe in Figure 6, for all models, the inclusion of various combinations of the biases results in better performance comparing to the original setting of the models. Among the various settings, the ones that accommodate the location bias as well as the query biases have superior performance in all plots, suggesting that the introduction of the location bias brings further improvement. The superiority of the performance appears stronger for the top-listed ads. This could be due to the fact that the top-listed ads receive the major attention and clicks from the users whereas the click chance over the side ads is much lower resulting in a more sparse samples for the side location.

5. PATTERNS FOUND IN USER BEHAVIOR

A major purpose of this paper is to determine if query and location biases can better reflect user browsing and click behavior. To a large extent, the findings of our experimental studies confirm the superior performance of our click models when these biases are incorporated. The current section aims to exploring whether these biases reflect clear distinctions in user behavior, such that they can be used as helpful signals in future click prediction models for sponsored search.
5.1. User Intent and Query Biases

In order to study the relation between the query bias parameters (initiation probability and persistence probability) and search intent, we used a semi-automatic approach for labeling a batch of queries in various dimensions of query intent by using crowdsourcing. Details of the annotation process can be found in [Ashkan 2013]. A set of 1000 queries was first labeled by three local annotators. These queries were then used as seed queries submitted along with an additional set of 3000 queries to the Amazon Mechanical Turk [2009] in order to be labeled by five different annotators. The seed queries were used to validate and approve the results obtained from Mechanical Turk. The results have also been validated through inter-annotator agreement analysis and through binary query intent classifiers. The agreement obtained among the annotators along different dimensions of query categories and the accuracies obtained from the trained query classifiers confirm that the studied query categories are reasonably distinguishable. It is worth noting that one could also use existing commercial query classifiers [Ashkan and Clarke 2009; Dai et al. 2006; Hu et al. 2009] for this analysis.
As a result of the annotation process, a set of 4000 queries were obtained to represent the commercial/ non-commercial intent of queries. If the assumed purpose of submitting a query is to make an immediate or future purchase of a product or service, the query is considered as “commercial”. Otherwise, if the purpose of the query is assumed to have little to do with commercial activity, it is considered as “noncommercial”. Moreover, a set of 2010 commercial queries was obtained along three dimensions of commercial intent:

- If the query is related to a specific retailer, it is labeled as a specific retailer, otherwise as a generic (unknown) retailer.
- If the query is related to a specific product, it is labeled as a specific product, otherwise as a generic (broad) category of product.
- If the query is related to a specific brand, it is labeled as a specific brand, otherwise as a generic (unknown) brand.

For instance, the commercial query “United Airlines ticket” is assumed to be product-specific and brand-specific, but with an unknown retailer since a United Airlines ticket can be purchased from different travel services.

From this process, we obtain intent labels for each query along each dimension (i.e., commercial/non-commercial, brand, product, and retailer) based upon majority agreements among annotators. These labels are used to determine whether those parameters of the models that were defined over user queries and learned through expectation maximization reflect varying behavioral patterns across different query intent categories.

We specifically focus on the result of DBN+VPI in order to study user intent with respect to initiation probability ($u^q$) and persistence probability ($\lambda^q$). For the commercial/non-commercial dimension, the cumulative density function (CDF) is calculated with respect to the values of $u^q$ separately for commercial queries and non-commercial queries. The two CDF curves are depicted in Figure 7. We note that each point ($u^q, \rho$) on a CDF curve indicates that the probability of having an initiation probability value less than $u^q$ is $\rho$ for the corresponding query category. We also note that

![Fig. 7: Difference in the initiation probability for the commercial/non-commercial dimension.](image-url)
for each pair of CDFs reported in this section we performed a two-sample KS-test (Kolmogorov-Smirnov test) at a significance level of 95%, where the difference between the CDFs for each pair has reported to be significant.

By comparing the trend of the curves in Figure 7, we can observe that the commercial category on average has higher probabilities for larger values of $u^q$ compared to non-commercial queries. This observation can be intuitively justified, since a user with commercial intent is more likely to initiate browsing an ad list compared to a user with non-commercial intent.

In addition to the initiation, the persistence probability could vary across different query intents. If the commercial intent behind a user’s query provides a higher probability to initiate browsing, different aspects of their commercial intent could also result in varying browsing behavior. To illustrate this variability, we calculated CDF with respect to the $\lambda^q$ parameter for our subcategories of commercial intent (i.e., retailer, brand, and product).

Fig. 8: Distinctions found in the user behavioral parameters across the commercial intent subcategories.

As we can observe in Figure 8a, retailer-specific queries exhibit lower persistence compared to non-retailer-specific queries. This observation is consistent with that of Ghose and Yang [Ghose and Yang 2008] for the impact of keyword attributes on consumer search and purchase behavior. Retailer-specific queries are usually navigational queries, for which the user may be a loyal customer, perhaps looking for information about a particular retailer, and expecting to find this information towards the top of the list.

Figure 8b, on the other hand, depicts CDF curves for product-specific queries against the generic ones, such that none of the categories include any particular retailer name in their queries, in order to avoid the impact of the retailer name. We can observe that users are relatively more persistent in browsing through the ad list if their query names a specific product. This observation could indicate that the user may need a commercial product, but does not yet know where to buy it, providing competitive search situations where the user is more persistent in browsing.
We further study the persistence of the user in the product dimension with respect to the presence and absence of the brand information. Figure 9 shows CDF curves for the four combinations of product and brand categories. Consistent with our earlier observation, queries that reflect specific product names have a relatively higher persistence probability. This observation can be confirmed by comparing the CDF for specific products against that of generic products in the presence of a brand name (specific brand), and also by comparing the CDF for specific products against that of generic products in the absence of a brand name (generic brand).

We draw attention to the effect of brand name on product-specific queries. By comparing the CDFs of the product-specific queries in absence of a brand name (generic brand) against the product-specific queries with a specific brand name, we can observe that once the brand name is included in the query users are less persistent in browsing. The former query type targets a specific commercial product (e.g. an air ticket), while the latter targets a specific brand of a specific product (e.g. an American Airlines ticket). One could argue the user who enters the first query would be more engaged in the browsing process, since any airline might do. On the other hand, the user who issues the second query may be a relatively loyal user who is looking for their favorite brand among the top results. If they do not find the brand among the top results, they may either abandon the search or move on to the organic results. This behavior could suggest that showing few but highly targeted ads for a specific brand of a product query is better than showing many ads that reflect various competing brands. Whereas, presenting various ads for competing brands of a product could be more effective for brand-generic product queries.

These sorts of observations are interesting and at the same time may be helpful in the sense that they are obtained from independent experiments. One group of experiments (from the previous section) empirically calculates the bias parameters with respect to the click signals recorded in the search engine log. The second round of experiments (in the current section) matches these values against query types that have been obtained independently, and finds distinctive patterns of user behavior based on them. These patterns not only shed light on user behavior, they may also suggest the
development of user dependent properties to be used as signals for ad click analysis in sponsored search.

5.2. User Behavior and Location Bias

With respect to the location bias, we study whether the behavioral parameters that are defined for different locations and are learned through the expectation maximization can reflect distinctions in user behavior at different locations on result pages. We focus on the result of DBN+VPIL on set $B$ to address the location bias parameters: $u_q^t$, $u_q^s$, $\lambda_q^t$, and $\lambda_q^s$.

Figure 10 depicts two types of comparison results for our purposes: one depicts the sorted values of $u_q^t - u_q^s$ for the corresponding queries, and the other contains the sorted values of $\lambda_q^t - \lambda_q^s$ across the queries. The idea is to examine whether users behave differently for top-listed ads and side-listed ads in terms of: i) their initiation probability to begin examining ads listed at different locations, and ii) their persistence probability in continuing to examine ads at different locations.

As we see in both plots of Figure 10, the parameters reflecting user behavior over the top ads appear to be larger than the corresponding values for the side ads for most of the queries. In more detail, the initiation probability for the top location is found to be higher than or equal to the initiation probability at the side for about 88% of the queries (note the vertical line that shows the threshold of 12%). This number appears to be about 78% with respect to the persistence probability at different locations. In other words, users are found to be more likely to initiate browsing of the top ads compared to the side ads. They are also found to be more persistent in browsing through the top ads as opposed to the side ads.

The hidden parameters $u_q^t$, $u_q^s$, $\lambda_q^t$, and $\lambda_q^s$ are learned through the expectation maximization technique, as explained previously, independent of any assumption about the superiority of the top ads over the side ads. However, the results of learning indicate such a superiority in Figure 10 for most cases. These observations confirm that users generally pay more attention to the top ads as opposed to the side ads; a signal that can be accommodated by the click models for better understanding of ad click prediction in sponsored search.
6. CONCLUSIONS
This paper studies the impact of a group of contextual factors on modeling user behavior in sponsored search, allowing us to improve the accuracy of click predictions. These factors include the probability that the user will initiate browsing the advertising links at different locations on the page and their persistence in continuing to browse these links. User initiation and persistence are modeled as query biases, while ad placement is modeled as a location bias. A group of existing probabilistic click models are adapted and extended to incorporate these contextual factors. The newly introduced parameters can be learned from click signals recorded in the logs of a commercial search engine.

Our evaluation results indicate that significant improvements can be achieved in click prediction once the overall quality of ads shown on a result page, along with location bias and query biases, are taken into consideration. Further investigation in this direction over other well-known click models are among the future directions for this work. Comparing the effectiveness of these factors in sponsored search versus organic search is also among the directions for future work. Our findings confirm that user behavior in sponsored search depends on the nature of the commercial intent underlying the query. This information may provide a better understanding of user behavior with respect to their query which may be used as helpful signal to better target context-based ad click prediction in the sponsored search domain.

We are aware that our simplifying assumptions regarding a user’s approach to browsing an ad list introduces limitations into our work. Instead of linearly browsing through the list, a user may randomly view an ad at a particular rank position or location, or they may move up and down in the list during their browsing session. However, the cascade assumption of linearly browsing enables us to represent the behavior of the many of users, which may be considered as a reasonable starting point to better understand user behavior in sponsored search. It also enables us to model ads in the context of the preceding ads. More complex models are required in the future in order to address random viewing and skipping over different positions in the ad list. We also model variability of user behavior through the parameters defined over queries by assuming that users issuing the same query have generally similar behavior. While a query-based representation of users appears to be effective in the domain of sponsored search [Yan et al. 2009], a more realistic assumption would define the parameters over user/query pairs, suggesting a possible direction for future work.

Another point that can be addressed in future work is to explore the extent to which incorporating the proposed biases would be helpful in click analysis. An effective way of evaluating the performance of these factors would be to study the performance of a general ad click model that incorporates both the content-based features and the proposed biases. This could not be studied in this paper due to the lack of information about the client side, ad content, and bid terms in the data.

ACKNOWLEDGMENTS
The authors would like to thank Microsoft Research and Microsoft adCenter for providing the Beyond Search data.

REFERENCES


APPENDIX

Details on the inference algorithm used throughout the paper are provided in this appendix. Note that the variables across this section are mostly used in their general form. By adding a superscript \( j \) to the variables, the same formulations can be used for a particular SERP \( j \) from the search log.

A. FORWARD-BACKWARD VARIABLES

Following the forward-backward algorithm [Rabiner 1989] and the details provided in Chapelle and Zhang [2009], the forward variable \( \alpha_{l,i}(e) \) and the backward variable \( \beta_{l,i}(e) \) can be defined as follows:

\[
\alpha_{l,i}(e) = P(C_{l,1}, \ldots, C_{l,i-1}, E_{l,i} = e | \theta) \\
\beta_{l,i}(e) = P(C_{l,i}, \ldots, C_{l,N_l} | E_{l,i} = e, \theta)
\]

where \( \alpha_{l,i}(e) \) is the probability of the partial click observations sequence \( C_{l,1}, \ldots, C_{l,i-1} \) and the examination state \( e \) for the ad \( a \) listed at rank position \( i \) and location \( l \), given the parameters of the model, \( \theta = (u_q^l, \eta_q^l, \lambda_q^l, \omega_q^l, \nu_q^l) \). Similarly, \( \beta_{l,i}(e) \) is the probability of the partial click observations sequence \( C_{l,i}, \ldots, C_{l,N_l} \) given the user's examination state \( e \) for ad \( a \) and \( \theta \). The recursion formula for these variables in our problem setting can be stated, as follows:

\[
\alpha_{l,i+1}(e) = \sum_{e' \in \{0, 1\}} \alpha_{l,i}(e') P(E_{l,i+1} = e, C_{l,i} | E_{l,i} = e') \\
\beta_{l,i-1}(e) = \sum_{e' \in \{0, 1\}} \beta_{l,i}(e') P(E_{l,i} = e', C_{l,i-1} | E_{l,i-1} = e)
\]

where the conditional probability \( P(E_{l,i+1}, C_{l,i} | E_{l,i}) \) can be computed based on the DBN model [Chapelle and Zhang 2009] and adopted to our setting by using Equations 7, 8, and 10, as follows:

\[
P(E_{l,i+1}, C_{l,i} | E_{l,i}) = \sum_{s \in \{0, 1\}} P(E_{l,i+1} | S_{l,i} = s, E_{l,i}) P(S_{l,i} = s | C_{l,i}) P(C_{l,i} | E_{l,i})
\]

Table V depicts the values of this conditional probability computed according to the above equation and based on the values of the variables.

Table V: The probability distribution for \( P(E_{l,i+1}, C_{l,i} | E_{l,i}) \).

| \( E_{l,i+1} \) | \( C_{l,i} \) | \( E_{l,i} \) | \( P(E_{l,i+1}, C_{l,i} | E_{l,i}) \) |
|-----------|----------|----------|-----------------|
| 0         | 0        | 0        | 1               |
| 0         | 0        | 1        | \((1 - \lambda_q^l)(1 - \omega_q^l)\) |
| 0         | 1        | 0        | 0               |
| 0         | 1        | 1        | \(\omega_q^l(1 - \lambda_q^l + \nu_q^l \lambda_q^l)\) |
| 1         | 0        | 0        | 0               |
| 1         | 0        | 1        | \(\lambda_q^l(1 - \omega_q^l)\) |
| 1         | 1        | 0        | 0               |
| 1         | 1        | 1        | \(\lambda_q^l \omega_q^l(1 - \nu_q^l)\) |
Finally, the base cases for the forward and backward variables can be obtained with respect to user initiation probability stated in Equation 5, as follows:

\begin{align*}
\alpha_{l,1}(0) &= 1 - u_l^q, & \alpha_{l,1}(1) &= u_l^q \\
\beta_{l,N_l+1}(0) &= 1, & \beta_{l,N_l+1}(1) &= 1
\end{align*}

where \(N_l\) is the number of ads appearing at the location \(l\) of the page.

**B. ESTIMATING POSTERIOR PROBABILITIES**

We use the forward and backward variables to compute the posterior probability of the transition state (i.e. \(E_{l,i}\)) that is a hidden variable in our model. Given the click sequence \(C_l\) observed for the ads listed on the location \(l\) of a result page, the posterior probability of the user examining or not examining an ad listed at the rank position \(i\) can be formulated using Bayes’ rule as:

\[
P(E_{l,i} = e|C_l, \theta) = \frac{P(C_l|E_{l,i} = e, \theta)P(E_{l,i} = e|\theta)}{P(C_l|\theta)}
\]

According to the Markov property, given a state, past observations are independent of the future observations. In our case, given the user’s decision about examining the ad at rank \(i\), the previous click events are independent of the future ones. This allows us to express the above equation as follows:

\[
P(E_{l,i} = e|C_l, \theta) = \frac{P(C_{l,1}, \ldots, C_{l,i-1}, E_{l,i} = e, \theta)P(C_{l,i}, \ldots, C_{l,N_l}|E_{l,i} = e, \theta)P(E_{l,i} = e|\theta)}{P(C_l|\theta)}
\]

where the numerator consists of the forward and backward variables, and the denominator is a normalization factor to make \(P(E_{l,i} = e|C_l, \theta)\) a probability measure such that \(\sum_{e \in \{0,1\}} P(E_{l,i} = e|C_l, \theta) = 1\). As a result, the conditional probability can be estimated using the forward-backward algorithm.

\[
P(E_{l,i} = e|C_l, \theta) = \frac{\alpha_{l,i}(e)\beta_{l,i}(e)}{\sum_{e' \in \{0,1\}} \alpha_{l,i}(e')\beta_{l,i}(e')}
\]  \hspace{1cm} (21)

The probability of examination given the model can be similarly estimated, as follows:

\[
P(E_{l,i} = e|\theta) = \sum_{C_l} P(C_l, E_{l,i} = e|\theta)
\]

\[
= \sum_{C_l} P(C_l|E_{l,i} = e, \theta)P(E_{l,i} = e|\theta)
\]

\[
= \sum_{C_l} P(C_{l,1}, \ldots, C_{l,i-1}, E_{l,i} = e, \theta)P(C_{l,i}, \ldots, C_{l,N_l}|E_{l,i} = e, \theta)P(E_{l,i} = e|\theta)
\]

\[
= \sum_{C_l} P(C_{l,1}, \ldots, C_{l,i-1}, E_{l,i} = e|\theta)P(C_{l,i}, \ldots, C_{l,N_l}|E_{l,i} = e, \theta)
\]

\[
= \sum_{C_l} \alpha_{l,i}(e)\beta_{l,i}(e)
\]

Finally, the probability of user being in the examination state $e$ at rank $i$ and in state $e'$ at rank $i + 1$, given the model and the click observations sequence $C_l$, can be expressed as follows:

$$P(E_{l,i} = e, E_{l,i+1} = e'|C_l, \theta)$$

$$(23)$$

$$= \frac{P(E_{l,i} = e, E_{l,i+1} = e', C_l|\theta)}{P(C_l|\theta)}$$

$$(24)$$

$$= \frac{1}{P(C_l|\theta)} \times P(C_{l,1}, ..., C_{l,i-1}, E_{l,i} = e|\theta)$$

$$\times P(E_{l,i+1} = e', C_{l,i}|E_{l,i} = e, \theta)$$

$$\times P(C_{l,i+1}, ..., C_{l,N_l}|E_{l,i+1} = e'|\theta)$$

$$= \frac{\alpha_{l,i}(e) P(E_{l,i+1} = e', C_{l,i}|E_{l,i} = e) \beta_{l,i+1}(e')}{\sum_{e'' \in \{0, 1\}} \alpha_{l,i}(e'') \beta_{l,i}(e'')}$$

As a result, $P(E_{l,i+1} = 1, E_{l,i} = 1|C_l)/P(E_{l,i} = 1|C_l)$ can be computed by dividing Eq. 23 by Eq. 21:

$$\frac{P(E_{l,i+1} = 1, E_{l,i} = 1|C_l)}{P(E_{l,i} = 1|C_l)}$$

$$(24)$$

$$= \frac{1}{P(C_l|\theta)} \frac{\alpha_{l,i}(1) P(E_{l,i+1} = 1, C_{l,i}|E_{l,i} = 1) \beta_{l,i+1}(1)}{\sum_{e'' \in \{0, 1\}} \alpha_{l,i}(e'') \beta_{l,i}(e'')}$$

$$= \frac{\beta_{l,i+1}(1)}{\beta_{l,i}(1)} P(E_{l,i+1} = 1, C_{l,i}|E_{l,i} = 1)$$

where $e$ and $e'$ from Equations 21 and 23 are substituted by 1, and $P(E_{l,i+1} = 1, C_{l,i}|E_{l,i} = 1)$ can be estimated according to Equation 20.