Search Costs vs User Satisfaction on Mobile

Manisha Verma, Emine Yilmaz

University College London
mverma@cs.ucl.ac.uk, emine.yilmaz@ucl.ac.uk

Abstract. Information seeking is an interactive process where users submit search queries, read snippets or click on documents until their information need is satisfied. User cost-benefit models have recently gained popularity to study search behaviour. These models assume that a user gains information at expense of some cost. Primary assumption is that an adept user would maximize gain while minimizing search costs. However, existing work only provides an *estimate* of user cost or benefit per action, it does not explore how these costs are correlated with *user satisfaction*. Moreover, parameters of these models are determined by desktop based observational studies. Whether these parameters vary with device is unknown. In this paper we address both problems by studying how these models correlate with user satisfaction and determine parameters on data collected via mobile based search study. Our experiments indicate that several parameters indeed differ in mobile setting and that existing cost functions, when applied to mobile search, do not highly correlate with user satisfaction.

Keywords: Cost-Benefit analysis, Effort, User Satisfaction

1 Introduction

Search is an extremely popular means of finding information online. Users repeatedly interact with a search engine to satisfy their information need which makes interactive information retrieval (IIR) an active area of research. Recently, large body of formal models [4, 5] have been proposed that capture user cost (or effort) and benefit by incorporating several user actions. Users incur some cost for each of these actions: input a search query, read snippets, click results or scroll up/down search engine result page (SERP). At present, cost of each action is measured in time, keystrokes or number of documents. For instance, query cost can be estimated via $W * c_w$ [4] where W is number of words in query and c_w is the average time it takes a user to type each word. Several models have been proposed [5], simulated [2] or empirically evaluated [3] on real datasets.

However, existing work only provides an *estimate* of user cost or benefit per action, it does not explore how these costs are correlated with user satisfaction. It remains to be seen what cost functions correlate best with user satisfaction. Existing research in IIR is also limited to a desktop setting. User models of search and interaction have been developed for desktop environments and lab studies have been conducted to empirically evaluate and learn these models. However, today users have quick access to information on several devices such as desktops, mobiles and tablets. Whether these models highly correlate with user satisfaction needs to be evaluated on different devices.

In this work we address above mentioned limitations of existing work. We begin by introducing a mobile specific dataset collected during a lab study. We explore different actions and their costs across 25 users and 193 sessions. We also investigate how these cost functions correlate with user satisfaction. Our

experiments show that once trained, cost-benefit model parameters are different for mobile search. We also found varied correlation between satisfaction and cost functions proposed in the literature. In following sections, we briefly explain cost functions proposed in literature, followed by examining correlation between user satisfaction and search costs (or effort), benefit (or gain) and profit respectively as proposed in previous work.

2 User Study and Data Statistics

We conduct a small scale search study to collect fine grained user interaction data and explicit labels for satisfaction from some users. We collected data for 10 topics chosen from publicly available dataset [6]. We tailored topic descriptions for mobile search and did not impose any time restrictions for completing these search tasks. We built an Android app[1] for our experimental study. Participants were free to issue as many queries as they liked. Search results were retrieved using Bing Search API[2] with fixed parameters. We ensured that if two participants issued the same query, they would see the same results by caching results for each query. We customized search interface for image, video and wiki results respectively to reflect existing commercial search engine result pages (SERPs). We logged several interaction signals such as clicks, taps and swipes on SERP. Participants were asked to provide feedback for SERP relevance and satisfaction on Likert scale of 1 (non-relevant/dissatisfied) to 5 (highly-relevant/satisfied). They could begin with any task and perform as many search tasks as they liked.

Participants were recruited via university mailing lists and social media websites. We collected data from 25 participants (7 females and 18 males) for this study whose age lies between 22-55. We asked participants that were familiar with search in mobile browser, to complete the study on their *personal* android phones.

Our data consists of 193 search sessions, 104 unique queries, 161 unique SERP result (URL) clicks and 192 relevance/satisfaction labels for SERPs. The distribution of SERP satisfaction labels is 1=13, 2=12, 3=32, 4=54, 5=81 respectively.

3 Cost/Benefit vs Satisfaction Analysis

Cost (or effort) and benefit can be analysed in multiple ways. Existing work [4] investigates user costs on a per-action basis. In this paper, we limit our investigation to two types of costs: query cost and click/scroll cost. Cost of querying solely depends on user's input query i.e. it is directly proportional to query length. However, click/scroll costs are relatively more complex as they depend on factors such as number of snippets read, clicked and number of SERPs examined by the user. We explain different cost/benefit functions, discuss their correlation with SERP satisfaction labels from our study and finally estimate their parameters by optimizing different cost functions in following subsections.

3.1 Query Cost-Benefit and User Satisfaction

Users rely on keywords to formulate their information needs. They may incur different costs for issuing query on different mediums. For instance, users can

[1] Topics, results and app at http://www0.cs.ucl.ac.uk/staff/M.Verma/app.html
[2] http://datamarket.azure.com/dataset/bing/search

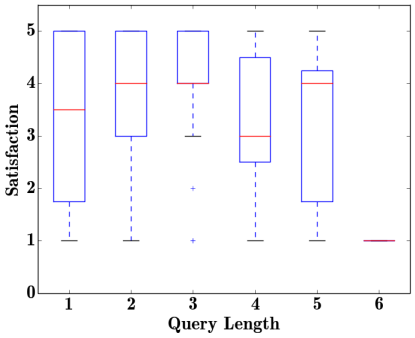

Fig. 1. sat vs query length

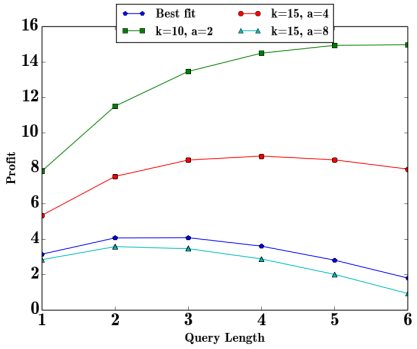

Fig. 2. profit curves

issue a query via keyboard or touch screens on desktop and mobile respectively. Users of our app were required to touch type their queries and we did not provide query auto completion, to ensure that users type all queries explicitly.

Given that a user enters a query with W words and c_w captures the effort required to input each word, we use the model from [4], in Equation 1, to compute net profit (π), benefit $b(W)$ and cost $c(W)$ for each query:

$$\begin{aligned} b(W) &= k.log_\alpha(W+1) \\ c(W) &= W.c_w \\ \pi(W) &= b(W) - c(W) \end{aligned} \tag{1}$$

Here, k represents a scaling factor and α captures diminishing returns of typing subsequent words. Distribution of satisfaction labels for queries of varying length is shown in Figure 1. We use same values for $k \in \{10, 15\}$ and $\alpha \in \{2, 4, 8\}$ as in [4] to compute Pearson correlation (ρ) between query profit and satisfaction. Correlation between satisfaction and profit for each combination of k and α is given in Table 1.

We obtain values of c_w, k, α by optimizing objective function in Equation 2 which minimizes the difference between user satisfaction ($\hat{\pi}$) and net user profit.

$$\min_{c_w,k,\alpha} \sum_{i=1}^{n} (\hat{\pi} - \pi(W))^2 \tag{2}$$

We can estimate parameters c_w, k and α by minimizing squared loss on satisfaction labels from our study. Parameter values $c_w = 2.18$, $k = 8.5$ and $\alpha = 3.0$ yield best fit on our data. When substituted, net profit has Pearson's ρ of 0.314 (p-val < 0.001) with satisfaction. Profit curves for different parameter settings are shown in Figure 2. We observe that as the length of query increases, overall profit of user decreases which was also reported in [4]. We also observe a similar trend in our data where profit is highest for three word queries and rapidly drops thereafter. Table 1 shows that higher α yields stronger correlation between satisfaction and user profit which indicates rapid diminishing returns of typing subsequent words. While query cost does not model entire search process, experiments on our data suggest that query costs (in Equation 2) can affect overall user satisfaction.

	k		
α	10	15	20
2	-0.10	-0.14	-0.15
4	**0.312***	-0.009	-0.10
6	0.271*	0.27*	-0.02
8	0.256*	**0.312***	-0.09
10	0.248*	0.295*	0.23

Table 1. Pearson's ρ

	k			
β	2.0	5.0	10.0	16.0
0.03	0.16*	0.14*	0.10*	0.09
0.3	**0.17***	0.13*	0.09*	0.08
0.43	0.16*	0.12	0.08	0.08
1.0	0.11	0.08	0.07	0.06

Table 2. Pearson's ρ

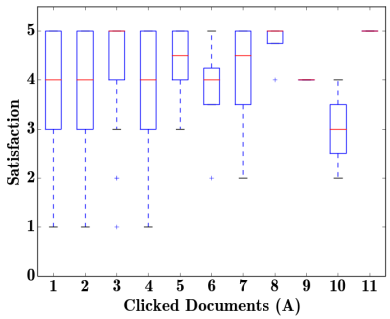

Fig. 3. # clicked documents

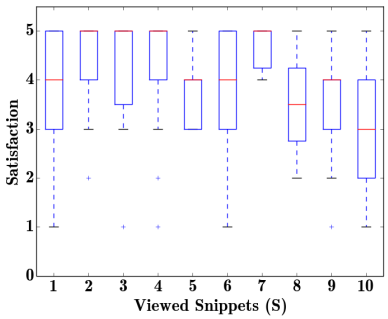

Fig. 4. # viewed snippets

3.2 Search Cost-Benefit and User Satisfaction

A user has choice of several actions on submitting any query to the search engine. They can either choose to examine a snippet, click a result, go to the next page or issue a new query. We assume that user submits Q queries, reads S snippets, views V SERP pages per query and reads A clicked documents. If the cost of querying is c_w, the cost of viewing a SERP page is c_v, the cost of reading a snippet is c_s and the cost of reading a clicked document is c_a respectively, we can use cost $c(Q, V, S, A)$ and gain/benefit $b(Q, A)$ function from [2] to compute the net profit π given in Equation 3. Here, α and β capture user's frequency of issuing multiple queries and reading documents respectively.

$$\begin{aligned} c(Q, V, S, A) &= (c_w + c_v.V + c_s.S + c_a.A).Q \\ b(Q, A) &= k.Q^{\alpha}.A^{\beta} \\ \pi &= b(Q, A) - c(Q, V, S, A) \end{aligned} \tag{3}$$

Distribution of satisfaction with respect to time spent on reading (or examining) A clicked documents, viewing S snippets, cost of reading each snippet (c_s) and clicked document (c_a) is shown in Figure 3, 4, 5 and 6 respectively. Some users in our study, despite clicking on more than 10 documents for a query, have assigned higher satisfaction grade to SERP. It is worth noting that the median cost of reading a snippet (in milliseconds) is higher on low satisfaction SERPs than on high satisfaction SERPs. However, the trend reverses in the curve depicting examination cost of clicked documents i.e. Figure 3 where users spend less time reading a document clicked on low satisfaction SERP than on high satisfaction SERP.

We optimize the function in Equation 2 with satisfaction labels and net profit for each SERP. Since our satisfaction labels are per SERP basis, we set Q=1 to compute per SERP cost and benefit function.

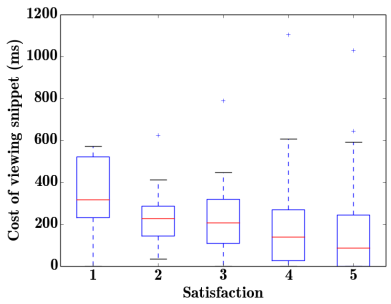

Fig. 5. cost of reading snippet (c_s)

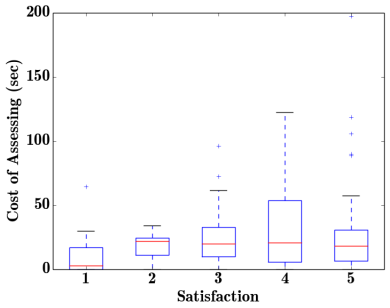

Fig. 6. cost of reading clicked doc (c_a)

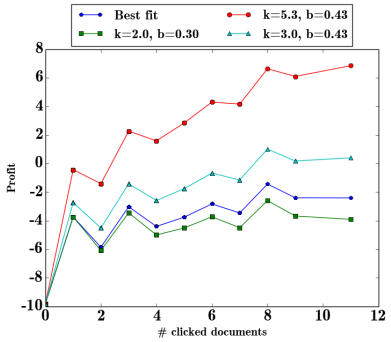

Fig. 7. clicked doc profit

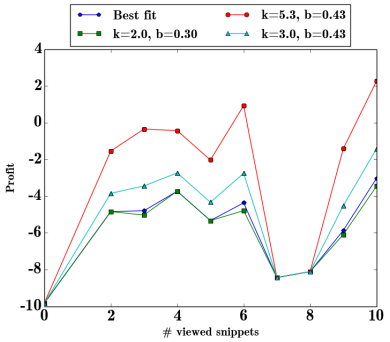

Fig. 8. viewed snippet profit

We perform optimization similar to Equation 2 where we minimize the difference between satisfaction labels and benefit obtained from total SERP interaction. We obtained lower value of $k = 2.0$ and $\beta = 0.30$ than previously reported values $k = 5.3$ and $\beta = 0.43$ as given in [1]. Variation in profit curves for different combinations of k and β for clicked documents and viewed snippets is given in Figure 7 and Figure 8, respectively. Pearson correlation ρ between net profit and satisfaction for different values of k and β is shown in Table 2[3].

Best fit ($k = 2.0$ and $\beta = 0.30$) net profit curve in Figure 7 shows that *change* in net user gain is highest when only one document is clicked. Net profit gradually increases as more documents are clicked. The kink in curve for two clicked documents suggest that other costs dominate cost function, thereby lowering net profit. We did not observe a significant drop in the profit with increase in number of clicked documents. However, net profit when $k = 5.3$ and $\beta = 0.43$ (from [1]) rapidly increases as more documents are clicked. Our data suggests that lower number of clicked documents yield higher user satisfaction on mobile. Profit curves for number of viewed snippets in Figure 8 shows a different trend. Net gain rapidly increases as users view more snippets but drops significantly when they read between six to eight snippets. Best fit curve shows highest profit when user views four snippets and declines thereafter. Best fit profit curve is similar to curve with $k = 5.3$ and $\beta = 0.43$ (from [1]) when plotted against viewed snippets. Table 2 shows that correlation between satisfaction and net benefit weakens as k and β increase.

[3] * indicates p-val< 0.05

Pearson correlation ρ between satisfaction and net search benefit on our data, for parameters obtained by optimizing objective function in Equation 3 ($k = 2.0$ and $\beta = 0.30$) was significantly low, only 0.17 (p-val<0.05) which indicates that linear combination of query, snippet examination and clicked document examination costs may not be optimal for mobile search. Pearson correlation ρ of each variable with satisfaction is as follows:

- $c_w * w$ = -0.33*
- c_v = 0.03, v = -0.02, $c_v.v$ = 0.03
- c_a = 0.07, A = 0.06, $c_a.A$ = 0.09
- c_s = -0.13*, S = -0.17*, $c_s.S$ = -0.16*

It is worth noting that each variable is correlated differently with satisfaction which is expected. While snippet (c_s) and query (c_w) costs are negatively correlated with satisfaction, cost of examining clicked document (c_a) and search result pages (c_v) are positively (but not significantly) correlated with user satisfaction.

Overall, for both query and search cost-benefit functions, we observed a different optimal value for each parameter on mobile. We observed higher correlation between net query benefit and satisfaction on mobile search data. However, satisfaction correlation with net search benefit was relatively low, which suggests that linear combination of search costs may not be suitable for a mobile setting.

4 Conclusion

Existing models of cost-benefit analysis models in IIR estimate how users maximize their net gain while minimizing search costs. These models do not provide any insight into how these strategies correlate with user satisfaction. Empirical study of these models is also limited to desktop setting. This paper was an investigation of correlation between cost-benefit of querying/searching and user satisfaction in mobile search. We found that optimal parameters of these models differ from desktops. We also found satisfaction to be highly correlated with net query profit but weakly correlated with net search profit. Our study motivates further investigation of non-linear cost models to better capture user behaviour on mobile devices.

References

1. L. Azzopardi. Economic models of search. In *Proceedings of the 18th Australasian Document Computing Symposium*. ACM, 2013.
2. L. Azzopardi. Modelling interaction with economic models of search. In *Proc. SIGIR*. ACM, 2014.
3. L. Azzopardi, D. Kelly, and K. Brennan. How query cost affects search behavior. In *Proc. SIGIR*. ACM, 2013.
4. L. Azzopardi and G. Zuccon. An analysis of the cost and benefit of search interactions. In *Proc. ICTIR*, pages 59–68. ACM, 2016.
5. L. Azzopardi and G. Zuccon. Two scrolls or one click: A cost model for browsing search results. In *Proc. ECIR*. Springer International Publishing, 2016.
6. T. Demeester, D. Trieschnigg, D. Nguyen, D. Hiemstra, and K. Zhou. Fedweb greatest hits: Presenting the new test collection for federated web search. In *Proc. WWW*. ACM, 2015.