Supporting Cross-Device Web Search with Social Navigation-Based Mobile Touch Interactions

Shuguang Han[1], Daqing He[1], Zhen Yue[2], and Peter Brusilovsky[1]

[1] University of Pittsburgh, PA, 15213, USA
[2] Yahoo! Labs, 701 First Ave., Sunnyvale, CA, 94089, USA
{shh69, dah44, peterb}@pitt.edu, zhenyue@yahoo-inc.com

Abstract. The wide adoption of smartphones eliminates the time and location barriers for people's daily information access, but also limits users' information exploration activities due to the small mobile screen size. Thus, cross-device web search, where people initialize information needs on one device but complete them on another device, is frequently observed in modern search engines, especially for exploratory information needs. This paper aims to support the cross-device web search, on top of the commonly used context-sensitive retrieval framework, for exploratory tasks. To better model users' search context, our method not only utilizes the search history (query history and click-through) but also employs the mobile touch interactions (MTI) on mobile devices. To be more specific, we combine MTI's ability of locating relevant subdocument content [10] with the idea of social navigation that aggregates MTIs from other users who visit the same page. To demonstrate the effectiveness of our proposed approach, we designed a user study to collect cross-device web search logs on three different types of tasks from 24 participants and then compared our approach with two baselines: a traditional full text based relevance feedback approach and a self-MTI based subdocument relevance feedback approach. Our results show that the social navigation-based MTIs outperformed both baselines. A further analysis shows that the performance improvements are related to several factors, including the quality and quantity of click-through documents, task types and users' search conditions.

Keywords: Mobile touch interaction; Cross-device web search; Social navigation

1 Introduction

The wide adoption of mobile phones facilitates people's information seeking process so that they can search for information at any time/place when their information needs are triggered. Although bringing the convenience, it is also observed that mobile users sometimes cannot complete a whole search task at one time and on one device. This is particularly common when users were handling complex search tasks [16]. Therefore, users may sometimes halt a search task on mobile and resume it on desktop. This is called the cross-device web search in

literature [16, 22]. Existing studies of cross-device search mainly focused on providing simple descriptive statistics for cross-device web searches and predicting device transition. In contrast, this paper explores automatic support of cross-device search queries (e.g., better ranking of relevant documents). Since simple tasks that can be fulfilled by one or two queries are less common in cross-device web search, we focus specifically on the exploratory search tasks. Furthermore, considering the difficulty of inferring search contexts for the search queries in late stage of a search process [7, 15], we set our target to support these queries.

A common approach for modeling search context is to utilize search history [12, 18]. The history from mobile devices may be sparse – given the same time, users may produce fewer histories because of the input difficulty on mobile [13]. One solution is to use search and interaction histories from other users who undertook similar tasks. This method, known as social search or social navigation [5, 6], has been explored by Farzan [8], Smyth et al. [19] and White et al. [23]. For example, Farzan [8] found that providing social cues (e.g. highlighted content) from other users helped the current user quickly identify relevant information. Our study is closer to [23] because our goal is to provide better document ranking for search queries rather than visualizing social cues to end users. However, we still name our method as social navigation because it is based on user exploration data that are traditionally used by social navigation. Including additional information increases the risk of having noise. Previous studies found that applying subdocument relevant information can help to reduce noise [4, 10]. Such information can be inferred either from gaze attention [4] or from mobile touch interactions (MTI) [10]. The latter does not require additional resources to obtain eye-gaze information, and thus is easier to scale up in live systems.
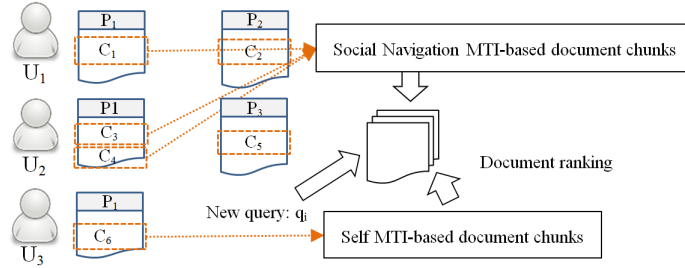
Fig. 1. Self MTI-based and social navigation MTI-based document chunks

Our method combines MTI's ability of pinpointing relevant subdocument content with social navigation's ability of aggregating relevant information. The aggregated content is then used as relevance feedback to support queries in cross-device web searches. To be more specific, a user can have two types of MTI: self MTI and social navigation MTI. The former refers to the MTIs performed by the user herself while the latter refers to the MTIs performed by other users on the web pages visited by the given user. This can be illustrated by Figure 1. Suppose that we have three users (U_1, U_2 and U_3) working on the same task. They visited three web pages (P_1, P_2 and P_3). Based on the logged MTIs, we can

infer six relevant document chunks (C_1, C_2, ..., C_6) using the method described in Section 3.2. To support a new query q_i for U_3, besides using q_i itself, we can also incorporate U_3's search history (i.e., P_1). Instead of using the full text of P_1, we may only use C_6 because we have identified that U_3 is only interested in C_6. We name it as the self MTI based method. When applying social navigation MTIs, we use C_1, C_3 and C_4 (C_6 is not included because it is self MTI) as relevant document chunks because other users (i.e., U_1 and U_2) who conducted the same task have shown interests in those document chunks. Note that, we do not include C_2 and C_5 because U_3 didn't click on P_2 or P_3.

Social navigation information can only be applied when users performed the same or similar tasks. White et al. [23] computed the similarity of two tasks by their query similarity. This approach requires a large-scale search log, which is difficult to acquire. Thus, we followed the approach used in Farzan [8] that performed lab-based user studies and preset the same task goals for all participants. To make our search tasks resemble real-world cross-device web searches, our tasks were directly adopted from Han et al. [10], where tasks were designed based on the results of an online survey. Among the collected information needs, we selected the ones that seek for multi-facet answers to simulate the intrinsic diversity (ID) tasks, which was identified as a popular type of information need in modern search engines [17].

Overall, we are interested in the following two research questions. First, can social navigation MTIs be used to improve the cross-device web search (we only study mobile-to-desktop web search in this paper) performance? Second, can social navigation-based MTIs also be used to improve the same-device cross-session web search (we only study the desktop-to-desktop web search in this paper) performance when MTIs are unavailable?

2   Related Work

Despite the popularity of cross-device web search in modern search engines [16], related studies on this topic, particularly the automatic support of such search scenario, are still rare. Existing work mainly focused on providing simple descriptive statistics and predicting device continuation [16, 22]. For example, Wang et al. [22] studied the task resumption in cross-device web search tasks and identified several interesting patterns that are associated with device transition. The support of mobile web search was explored in Song et al. [20]. They studied the search process on different devices (desktop, tablet and smartphones) and found that mobile search performance could be improved by transferring user behaviors from desktop. However, their focus is not on cross-device web search.

To support cross-device web searches, the techniques from context-sensitive information retrieval [18] and search personalization [1] can be used. A common approach is to treat users' search history as relevance feedback to re-rank relevant documents. Shen et al. [18] found that both query history and click-through are useful feedback resources but the latter is more effective. When applying the click-through, previous studies either incorporated the full text [12]

or subdocument chunks within a document [4]. It is reported that using subdocument content is more effective [4, 14]. However, obtaining subdocument chunks is a non-trivial task. Liu and Croft [14] treated the best-matched passages in a document for the given query as relevant subdocument chunks. Buscher et al. [4] obtained subdocument chunks through eye-tracking, where the content with more gaze attentions are chosen as relevant subdocument chunks. Han et al. [11] located subdocument chunks through touch-based interactions. The study presented in this paper takes a similar approach as Han et al. [11], but also differs from their study. In addition to the touch position and touch speed used in Han et al.[11], the inactive time is also adopted in our work based on a recent study [9], finding that the inactive time is a strong indicator for content relevance. Besides, we study both self MTIs and social navigation-based MTIs, which were not considered in Han et al. [11].

One challenge of utilizing search history is to handle the data sparseness. White et al. [23] found that it is possible to include the feedback information from other users who conducted similar search tasks. Similar idea were also explored as social navigation in literature [6], where social cues generated from other users are used to assist the current user on navigating the complex information space. Previous studies showed that applying social navigation in web search [8] and e-learning systems [3] enables people to access and utilize relevant information more effectively. However, social navigation has the risk of introducing noise because different users may have different search intentions, even when visiting the same page. A possible remedy is to differentiate users intentions and utilize the social cues only from people within the same user cohorts [23] or search community [19]. These approaches require acquiring user information which may be hard to obtain. Other information such as users' MTI behaviors can also help to identify a fine-grained user interest, and thus can be used to filter out the noise [10]. In this paper, we try to combine the advantages of social navigation and MTIs, and expect further improvement on search performance.

3   Experiment Setup

We built a cross-device web search dataset through a controlled lab study. In this paper, the cross-device search referred to mobile-to-desktop (M-D) web search. A desktop-to-desktop (D-D) search was also included for two reasons. First, one of our research questions is to apply social navigation based MTIs in D-D. Second, users can explore information more conveniently and thoroughly on desktop, which help us build a more reliable ground truth. After obtaining such dataset, we applied the context-sensitive retrieval model [18] to re-rank search results for the queries from the second session, as stated in the Introduction.

3.1   Data Collection

Obtaining our experimental data requires: (1) a cross-device web search system to record user behaviors; (2) properly designed cross-device web search tasks; (3) a well-designed experiment procedure; and (4) a ground-truth.

CrossSearch. Recording user behaviors, particularly the MTIs, is the most important component in our study. Such information was logged in CrossSearch, our self-developed system. When receiving a query, CrossSearch triggers Google API and displays search results returned by Google. Besides, CrossSearch also logs users queries and clicked documents.

Search tasks. As stated, the cross-device search tasks we used are directly adopted from Han et al. [10], which included six tasks from three categories: news search (NE), product search (PD) and people search (PE). Each category had two tasks, each task was designed on top of the results from an online survey.

User study design. Our study included both M-D and D-D. We employed a within subject design so that each participant searched for all six tasks with three under M-D and the other three under D-D. The combination of tasks, search conditions (M-D or D-D) and task sequence were rotated based on Latin square to minimize fatigue and learning effects. Each task was divided into two sessions, 7 minutes for each. In the first session, the participants performed three tasks on mobile and the other three on desktop and all the six tasks were resumed on desktop in the second session. Whenever they found relevant documents, the participants can save the documents. At the end of each task, the participants were asked to rate the relevance of each saved webpage on a 5-point Likert scale with 1 denoting not relevant and 5 being highly relevant.

Data collection. We recruited 24 participants (15 females and 9 males, 16 undergraduates and 8 graduates) from University of Pittsburgh and Carnegie Mellon University during October to November in 2013. All participants were self-rated experienced searchers. The participants issued 961 unique search queries, visited 1,790 unique webpages and saved 1,125 of them. In total, we logged 3,286 MTIs on the clicked web pages, where the dragdown (53.1%), dragup (17.5%) and tap (12.7%) were three dominating behavior types. Pinch-in and pinch-out were ignored because of the number was too few (less than 1%).

Ground truth. The ground-truth relevance of each document was computed based on users' post-task self-rated document relevance. Since a document can be saved by different users with different relevance scores, the simplest way for aggregation is to compute average. However, it may be biased for the documents saved only by few users. Thus, a Bayesian smoothing method was adopted to remove this bias [21, 10]. When generating ground truth for the search queries on M-D, we aggregated the saved documents from D-D because the information from M-D was used for producing the relevance feedback. When generating ground truth for D-D, we excluded information from a given participant but included that from all other participants.

3.2   Applying Context-Sensitive Retrieval Model

Our next step is to re-rank search results based on the context-sensitive retrieval model [18]. In this paper, we only used the click-through information. First, previous studies [18, 10, 12] identified that the click-through is superior to query history. Second, the focus of our later studies is to recognize relevant subdocument content chunks from click-through documents. Using click-through

makes our results more consistent in comparing to different click-through models. Following the majority of previous studies [18, 10, 12], we employed language modeling approach to represent different click-through models.

Unigram click-through language model was inferred by the method proposed in Shen et al. [18]. Their approach estimated the click-through language model through averaging all unigram word distributions inferred from the full text of click-through documents.

MTI-based click-through language model. Biedert et al. [2] found that users' touch positions usually lie within their reading zones. Thus, we assumed that if a user touched on position P (string index of the touched content in the whole document), she can read at most M characters before and after P, which means that the reading zone is [P-M, P+M]. We further measured the relevance of a reading zone based on inactive time and gesture speed [9, 11] because they are two strongest indicators for content relevance. We used two rules to define a relevant reading zone: (1) the reading speed that is slower than S and had an inactive time longer than I; (2) top M characters of a clicked document are used as relevant content if there are no MTIs. In this paper, we set S=500(pixels/second), I=1(second) and M=85(characters, ∼14-15 words) because of their best performance in training datasets. The employed MTIs were dragup and dragdown because they took high proportion in MTIs. Tap was not used because it is the same as click in desktop. Finally, depending on the MTIs we used, we can either build a self-MTI or social navigation-MTI based click-through language model.

Utilizing estimated language model for document re-ranking. We applied the following procedure to re-rank relevant documents for each query in the second session: for each to-be-ranked candidate document, we estimated its document language model. The matching between a candidate document and the click-through was measured by the KL divergence between their language models. The matching between a candidate document and the given query was measured by Google rank position. The two scores were combined for a final document ranking. Instead of using linear interpolation as [18], we used LambaMART in RankLib (http://sourceforge.net/p/lemur/wiki/RankLib/) because of its better performance and easy parameter tuning. Therefore, for each query-document pair, we applied two ranking features: the Google rank position, and the KL-divergence between the click-through language model and candidate document language model. Depending on the methods we used for estimating the click-through language model, we had different variants of KL divergence.

Evaluation. The dataset was divided into the training and testing parts. The training part was used to estimate model parameters and the testing part was used to test the effectiveness of each method. Training-testing division was based on task type. Then, we can obtain three datasets: PD (PD for testing and the rest for training), PE (PE for testing and the rest for training) and NE (NE for testing and the rest for training). We chose nDCG@20 as an evaluation metric because our study is to support exploratory search tasks, where users usually explored the result space in depth and saved many relevant web pages. In our study, the participants saved $13.00(\pm6.65)$ documents on M-D and $15.88(\pm9.77)$ documents

on D-D. This is different from studying the navigational information needs, where nDCG at lower cutoffs were often used. Besides, in another study [10] with similar experiment setting, we found that using nDCG at different cutoffs produced consistent results. Therefore, we only reported nDCG@20 results.

4   Results Analysis and Discussions

This section begins with testing the utility of applying social navigation-based MTIs in M-D. Then, a similar approach is applied for D-D, in which the MTIs from M-D are adopted. The compared baselines and context-sensitive retrieval models include: (1) pure Google search results (G); (2) Google ranks with click-through language model estimated from the full-text of click-through documents $(G+H_F)$; (3) Google ranks with click-through language model estimated from the self MTI-based document chunks $(G+H_{MTI-S})$; and (4) Google ranking position with social navigation MTI-based document chunks $(G+H_{MTI-SN})$.

4.1   Applying Social Navigation-based MTIs in M-D

Overall results. As the results shown in Table 1, $G+H_{MTI-SN}$ achieves the best performance, which demonstrates the usefulness of social navigation MTIs. Comparing to the use of full text, MTIs indeed help locate more relevant content: both $G+H_{MTI-S}$ and $G+H_{MTI-SN}$ are significantly better than $G+H_F$. The social navigation MTIs $(G+H_{MTI-SN})$ further improve the search performance over the self-MTIs. The reasons might be that the social navigation MTIs are less likely to encounter the data sparseness problem and the document chunks identified by social navigation MTIs can also vote for the most relevant information.

Table 1. nDCG of different runs on M-D. $\uparrow/\downarrow$ mean the increase/decrease; $\sim$Hx means a significance test comparing to Hx. p-values are based on Wilcoxon signed-rank test.

	nDCG@20	Sig.($\sim H_F$)	Sig.($\sim H_{MTI-S}$)
G	0.3974	$\downarrow$,p<0.001	$\downarrow$,p<0.001
$G+H_F$	0.4298	-	$\downarrow$,p=0.004
$G+H_{MTI-S}$	0.4421	$\uparrow$,p=0.004	-
$G+H_{MTI-SN}$	0.4497	$\uparrow$,p<0.001	$\uparrow$,p<0.003

Task effects. We separate the search performance by task to understand the task effect. The results are provided in Table 2. We find that PE and NE tasks show the same trends as the overall nDCG changes while PD presents fewer significant differences. We think it may be related to the task nature: some tasks may need to explore diverse topics while the others tend to find more similar results. If PD belongs to the former task type, simply applying relevant information from search history may not work. To test this hypothesis, we compute the overall query similarity between the first and second search session. If our hypothesis is correct, the query similarity of PD should be smaller.

Query similarity is measured by the cosine similarity (with tf-idf weights) of their corresponding Google search result pages. We did not use the query keyword matching due to the word mismatching problem. The final query similarity

is averaged over all of the possible query pairs between the first and the second session. We find that the average query similarity for PD is 0.0146, which is statistically significantly lower than 0.0199 for NE (p-value=0.041) and 0.0194 for PE (p-value=0.039). This indicates that users are more likely to explore different information in PD. This result suggests the importance of considering task nature when utilizing MTI-based relevant content.

Table 2. nDCG of different runs on M-D. Numbers in bold (italics) indicates p<0.05 compare with H_F (H_{MTI-S}) using Wilcoxon signed-rank test.

	PD	**PE**	**NE**
G	**0.4012↓**	**0.4648↓**	**0.3286↓**
G+H_F	0.4300	0.4835	0.3773
G+H_{MTI-S}	0.4276	**0.5077↑**	0.3902
G+H_{MTI-SN}	0.4249	*0.5197↑*	0.4019↑

Impacts of click-through. We also examine the impacts of the quantity and quality of click-through documents to the search performance. First, we categorize the to-be-supported queries into three groups based on the number of corresponding click-through documents: "1-7", "8-10" and "11+". Each group has about equal number (∼110) of queries. To achieve a better understanding of how different feedback information performs in each click-through group, we compute nDCG changes, ΔnDCG=nDCG(G+H_X)–nDCG(G), before (G) and after (G+H_X) applying the click-through in each group and then plot them in Figure 2. A positive value means that applying the click-through is better than pure Google results. Here, H_X refers to different types of feedback information, which can either be H_F, H_{MTI-S} or H_{MTI-SN}. We name the corresponding nDCG changes as ΔNDCG_HF, ΔNDCG_HMTI_S and ΔNDCG_HMTI_SN.
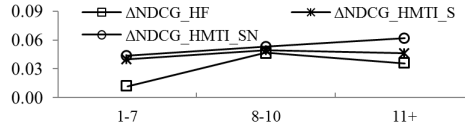
Fig. 2. ΔNDCG (Y axis) against the quantity of click-through (X axis) on M-D

All three curves in Fig. 2 are above zero, which shows the effectiveness of applying click-through information. In the ideal case, increasing the click-through quantity can result in a better performance because of having more information. However, although nDCG of G+H_F is positive in "11+", it is smaller than that in "8-10". We think the quality of click-through may also play important role here. Indeed, the average relevance for the click-through documents in "11+" is 2.11, while it is 2.40 in "8-10". The two values have significant difference (p=0.01).

One way to improve search performance is to employ the relevant subdocument content instead of the full text. In the group "11+", we indeed observe that the nDCG of G+H_{MTI-S} (p-value=0.034) and G+H_{MTI-SN} (p-value<0.001) are statistically significantly higher than the nDCG of G+H_F. The same trend is also observed in the group "1-7". This shows the robustness of MTI-based approaches. However, we do not observe significant differences in the group "8-10",

probably because they have already achieved very good performance. Furthermore, the performance of the self-MTI method does not always increase as the click-through quantity increases while that of the social navigation-based method keeps increasing. This also shows the robustness of social navigation-based MTIs in comparison with self-MTIs.

4.2 Applying Social Navigation-based MTIs in D-D

Overall results. Social navigation MTI-based method does not use information from the current querying user, which enables us to employ MTIs from mobile devices (there are no MTIs on desktop) to support the search queries in D-D. Suppose that we know a user U_i's click-through $C(c_1, c_2, ..., c_n)$ in the first session of D-D. For each to-be-supported query q_m of user U_i, when applying the click-through C, we can use MTIs from M-D on c_i to infer the relevant subdocument chunks and then apply them to estimate the click-through language models for document ranking. Given the best performance of social navigation MTI-based approach on M-D, we expect the same effects on D-D.

Yet, applying the same method as we did on M-D does not result in performance improvement as we expected (see Table 3). There are no significant difference between $G+H_{MTI-SN}$ and $G+H_F$. Further studies on the impact of click-through quantity/quality and task effect (see later sections) also do not reveal any difference. We think this is due to the device difference between M-D and D-D in the first session. Since there are more click-through documents in D-D, the social navigation based method may potentially introduce more noise, particularly when considering that users who visited the same documents may not have the same purposes. Therefore, we conduct further studies to understand whether it is possible to identify and differentiate users' search purposes.

Table 3. nDCG of different runs on D-D. Notations are the same as Table 1.

	nDCG@20	Sig.($\sim H_F$)	Sig.($\sim H_{MTI-SN}$)
G	0.4068	↓,p<0.001	↓,p<0.001
G+H$_F$	0.4452	-	↓,p=0.610
G+H$_{MTI-SN}$	0.4491	↑,p=0.610	-
G+H$_{MTI-SN-Q}$	0.4549	↑,p=0.021	↑,p=0.006

Utilizing MTIs from similar search queries. To find relevant documents, users need to issue queries, click webpages and interact with these webpages. Thus, each webpage and each interaction are associated with at least one query. We hypothesize that the same webpage under the same/similar queries may better reveal users' similar search purposes. In this case, when applying the social navigation MTI-based approach, we can use MTIs that are not only from the same webpage but also associated with similar queries. Query similarity is computed using cosine similarity (with tf-idf weights) of their corresponding Google search result pages. If cosine similarity is higher than 0.125, we define the two queries as similar. Based on this, we can build a new MTI-based click-through language model, where we still use social navigation MTI-based document chunks but

filter out the MTIs that are not associated with similar queries. We named this feedback information as $H_{MTI\text{-}SN\text{-}Q}$. Results in Table 3 show that $G+H_{MTI\text{-}SN\text{-}Q}$ does improve the search performance: it is significantly better than both $G+H_F$, and $G+H_{MTI\text{-}SN}$. This indicates the importance of differentiating the MTIs under similar search queries when applying the social navigation based approach.

Task effects. Table 4 presents the performance for each task, which shows that the effectiveness of social navigation approach is different for different tasks. Applying the query-based differentiation on MTIs ($G+H_{MTI\text{-}SN\text{-}Q}$) does improve nDCG, but its effect is task-dependent. It improves nDCG on NE, and has a trend on PE. However, it still has no impact on PD. The explanation may be the same as we hypothesized for M-D (see Table 2). This, again, illustrates the importance of considering task effects when applying such method.

Table 4. nDCG on the use of different runs on D-D. Numbers in bold (italics) indicates p<0.05 compare with H_F($H_{MTI\text{-}SN}$) using Wilcoxon signed-rank Test.

	PD	**PE**	**NE**
G+H$_F$	0.3612	0.5420	0.4308
G+H$_{MTI\text{-}SN}$	0.3575	0.5462	0.4498
G+H$_{MTI\text{-}SN\text{-}Q}$	0.3608	0.5529↑(p=0.06)	*0.4576↑*

Impacts of click-through. We also analyze the impact of click-through to the search performance. Similar to our M-D analysis, we divide the to-be-supported queries into three groups (each has ∼100 queries): "4-10", "11-17" and "18+". In each group, we compute ΔNDCG: the nDCG of each approach minus the nDCG of pure Google search (G). The results are plotted in Figure 3. $G+H_F$ shows the same trend as that on M-D: ΔNDCG is growing from the "smallest" group to the "medium" group but is decreasing from the "medium" to the "largest". However, it is different from the trend on M-D for $G+H_{MTI\text{-}SN}$: the performance does not grow from the "medium" group to the "largest" group. Thus, we analyze the change of relevance for the click-through documents from the "medium" group to the "largest" one on both M-D and D-D. The value is calculated using the "largest" group minus the "medium". We find that the change is -0.48 for D-D and -0.29 for M-D. The drop of click-through document quality may account for decreasing of ΔNDCG from the "medium" group to the "largest" group in D-D.
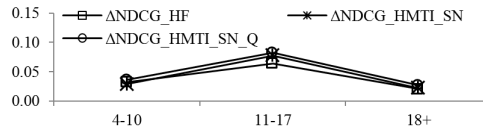
Fig. 3. ΔNDCG (Y axis) against the quantity of click-through (X axis) on D-D

4.3   Summarizations, discussions and implications

Consistent with our previous study [10], we observed that the self MTI-based subdocument relevance feedback outperformed both the initial Google ranking and the full text-based relevance feedback. This study went further by proving that the social navigation-based relevance feedback could achieve even better results than self MTI-based relevance feedback. The effectiveness of social

navigation-based approach may come from two aspects: (1) it relieves the data sparseness problem when users' own behaviors provide too little information; and (2) MTIs from different users can vote for the most relevant subdocument chunks. More importantly, we demonstrated that the social navigation-based MTI can be applied to both D-D and M-D conditions. This would be a very exciting finding for modern search engines since they have search logs about users from both their mobile and desktop search activites. Rich and fine-grained interactions obtained from MTIs enable search engines to accurately infer user interests, which in turn could better support both mobile and desktop searches.

A critical prerequisite for applying social navigation is that users share common task goals. Our study showed that this might not be enough. Different search queries might focus on different aspects even within the same task. Therefore, it is critically needed to have proper search intention discernment mechanisms such as using query similarity to differentiate social navigation information.

5   Conclusion and Future Work

With the wide adoption of smartphones, it is more common for people to conduct cross-device web search. Therefore, more and more users' mobile touch interactions (MTI) are collected when they access information through mobile devices. In this paper, we study whether the social navigation-based MTIs (MTIs collected from other users who visited the same webpages) can be employed to support the current user's cross-device web search. The MTIs used in this paper were collected through a lab study with 24 participants working on six cross-device web search tasks. Using the collected data, we then setup a relevance feedback experiment to evaluate our proposed approach. We find that the social navigation-based MTIs indeed help to improve search performance for cross-device web search over the methods of not using MTIs or only using self-performed MTIs. Further studies show that the effectiveness of social navigational based MTI is relates to many factors such as the quantity/quality of click-through and the task nature.

There could be several topics to be explored further. Firstly, we only considered the support of cross-device search queries in the late search stages, while it is also interesting to support queries at the beginning stage. This would be more challenging because of insufficient search history. Secondly, we found that both self-MTI and social navigation-based MTI can lead to the performance improvements. However, we have not studied whether they could be combined to further improve the search performance. Thirdly, we recruited only 24 participants and performed our study in a lab-controlled environment, which may not trully reflect users' interactions in a live search environment. Thus, a large-scale analysis of our approach from real search engine logs might be needed to confirm some of our findings. Finally, although we focused on cross-device web search in this paper, our approach can be easily adapted to other search scenarios such as pure mobile search or desktop search. Overall, MTI provides us many interesting challenges and opportunities.

References

1. Bennett, P. N., White, R., Chu, W., Dumais, S. T., Bailey, P., Borisyuk, F. and Cui, X.: Modeling the impact of short-and long-term behavior on search personalization. In SIGIR. 185-194, (2012).
2. Biedert, R., Dengel, A., Buscher, G. and Vartan, A.: Reading and estimating gaze on smart phones. In ETRA. 385-388 (2012).
3. Brusilovsky, P., Chavan, G. and Farzan, R.: Social adaptive navigation support for open corpus electronic textbooks. Adaptive Hypermedia and Adaptive Web-Based Systems. 24-33, (2004).
4. Buscher, G., Dengel, A. and van Elst, L.: Query expansion using gaze-based feedback on the subdocument level. In SIGIR. 387-394, (2008).
5. De Bra, P., Brusilovsky, P., and Houben, G.: Adaptive hypermedia: from systems to framework. ACM Computing Surveys (CSUR). 31, 4es, 12 (1999).
6. Dourish, P. and Chalmers, M.: Running out of space: Models of information navigation. Short paper presented at HCI. 94, 23-26, (1994).
7. Evans, B. M. and Chi, Ed. H.: An elaborated model of social search. Information Processing & Management. 46,6: 656-678(2010).
8. Farzan, R.: A study of social navigation support under different situational and personal factors. Doctoral Thesis of University of Pittsburgh. (2009).
9. Guo, Q., Jin, H., Lagun, D., Yuan, S. and Agichtein, E.: Mining touch interaction data on mobile devices to predict web search result relevance. In SIGIR. 153-162 (2013).
10. Han, S., He, D. and Yue, Z.: Understanding and Supporting Cross-Device Web Search for Exploratory Tasks with Mobile Touch Interactions. TOIS. (2015).
11. Han, S., Hsiao, I. and Parra, D.: A Study of Mobile Information Exploration with Multi-touch Interactions. In SBP. 269-276 (2014), Springer.
12. Jiang, J., Han, S., Wu, J. and He, D.: PITT at TREC 2011 Session Track (2011).
13. Kamvar, M. and Baluja, S. : Deciphering trends in mobile search. Computer, IEEE. 8: 58-62 (2007).
14. Liu, X. and Croft, W. B.: Passage retrieval based on language models. In CIKM. 375-382, (2002).
15. Marchionini, G.: Exploratory search: from finding to understanding. Communications of the ACM. 49, 4: 41-16(2006).
16. Montañez, G. D., White, R. W., and Huang, X.: Cross-Device Search. In CIKM. 1669-1678 (2014).
17. Raman, K., Bennett, P.N. and Collins-Thompson, K.: Toward whole-session relevance: Exploring intrinsic diversity in web search. In SIGIR. 463-472, (2013).
18. Shen, X., Tan, B., and Zhai, C.: Context-sensitive information retrieval using implicit feedback. In SIGIR. 43-50 (2005).
19. Smyth, B.: A community-based approach to personalizing web search. Computer, IEEE. 40, 8, 42-50, (2007).
20. Song, Y., Ma, H., Wang, H. and Wang, K.: Exploring and exploiting user search behavior on mobile and tablet devices to improve search relevance. In WWW. 1201-1212, (2013).
21. Särkkä, S.: Bayesian filtering and smoothing. Cambridge University Press. 3 (2013).
22. Wang, Y., Xiao, H.and White, R.: Characterizing and supporting cross-device search tasks. In WSDM. 707-716 (2013).
23. White, R. W., Chu, W., Hassan, A., He, X., Song, Y., and Wang, H.: Enhancing personalized search by mining and modeling task behavior. In WWW. 1411-1420 (2013).