

# 基于吸收态随机行走的两阶段效用性查询推荐方法

朱小飞 郭嘉丰 程学旗 兰艳艳

(中国科学院计算技术研究所 北京 100190)

(zhuxiaofei@software.ict.ac.cn)

## A Two-Step Utility Query Recommendation Method Based on Absorbing Random Walk

Zhu Xiaofei, Guo Jiafeng, Cheng Xueqi, and Lan Yanyan

(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190)

**Abstract** Search engine has become an essential way for satisfying users' daily information needs, however, formulating a proper query for search is difficult for users. To alleviate users' search burden, query recommendation has been proposed and considered as a prominent ingredient of modern search engines. Traditional recommendation approaches have paid great attention to recommend relevant queries, which attempt to find alternative queries with close search intent to the original query. However, the ultimate goal of query recommendation is to assist users to accomplish their search task successfully, while not just find relevant queries in spite of they can sometimes produce useful search results. To better match user search objective in the real world, a more straight way of query recommendation is to recommend users high utility query, i. e., queries that can better satisfy users' information needs. In this paper, we propose a two-step utility query recommendation method based on absorbing random walk, which can infer query's utility by simultaneously modeling both users' reformulation behaviors and click behaviors. Extensively experiments are conducted on a real query log, and the results show that this method significantly outperforms five baseline methods under the evaluation metric query relevant ratio (QRR) and mean relevant document (MRD).

**Key words** query recommendation; query log; absorbing random walk; query flow graph; utility

**摘 要** 搜索引擎已经成为人们获取信息的重要途径,然而对于用户而言如何构造一个合适的查询仍然是一项困难的工作.为了减轻用户搜索信息的负担,查询推荐技术应运而生并且已经成为当今搜索引擎不可或缺的组成部分.传统的查询推荐方法主要关注向用户推荐相关性查询,即推荐与源查询具有相近搜索意图的其他查询.然而查询推荐的根本目标是帮助用户成功完成其搜索任务,而不仅仅是找到相关性查询,尽管相关性查询有时也能得到有用的搜索结果.为了更好地满足用户的搜索目标,一种更直接的查询推荐方式是向用户推荐高效用性查询,即能够更好满足用户信息需求的查询.提出了一个基于吸收态随机行走的2阶段效用性查询推荐方法,该方法能够同时对用户的查询重构行为和查询点击行为进行建模并推导出查询的效用.在真实查询日志上的实验结果表明:新方法在评价指标查询相关率(query relevant ratio, QRR)和平均相关文档数(mean relevant document, MRD)上要显著优于其他5种基准方法.

收稿日期:2012-05-14;修回日期:2013-03-05

基金项目:国家“九七三”重点基础研究发展计划基金项目(2012CB316303,2013CB329602);国家自然科学基金项目(60933005,61232010,61003166,61203298);国家科技支撑计划基金项目(2012BAH39B02)

**关键词** 查询推荐;查询日志;吸收态随机行走;查询流程图;效用性

**中图法分类号** TP391

搜索引擎已经成为人们从海量 Web 数据中获取信息的一个不可或缺的工具. 为了有效地使用搜索引擎, 用户需要事先构造一个恰当的查询, 然而这对用户来说是一件困难的事情, 因为查询词通常比较短且富有歧义性. 事实上, 用户需要通过查询重构才能完成其搜索任务. 我们对 Yandex 查询日志 (<http://imat-relpred.yandex.ru/en/>, 共包含超过 4 000 万个查询会话数据) 的分析发现, 有超过 60% 的查询任务需要进行查询重构. 为了帮助用户构造查询, 目前主流的搜索引擎都向用户提供了查询推荐<sup>[1]</sup>的功能, 即在搜索引擎的上方或侧面显示与用户输入查询相似的其他查询来帮助用户进行重构, 这种方式可以有效减轻用户查询重构的负担.

现有查询推荐方法主要关注向用户推荐相关性查询, 即推荐与源查询相似的查询来替代源查询进行搜索. 查询日志中用户的行为信息(如与源查询含有相同词<sup>[2]</sup>, 点击相同的文档<sup>[3]</sup>, 或者在相同的查询会话中共现<sup>[4]</sup>)常常被用来找出哪些查询与源查询相似. 然而, 查询推荐的本质目标是帮助用户成功地找到其想要的信息, 而不是仅仅帮助其找到相似的查询, 因此, 传统的查询推荐方法很显然不能直接满足用户查询的本质目标. 例如, 用户初始提交一个查询“iphone available time market”, 其目的是要找“what’s the time of iphone to sell on the market”的信息, 相关的候选查询可能包括“iphone market sale time”, “iphone start selling market”和“iphone release date”. 这些候选查询都与源查询相关, 而且前 2 个候选查询看上去要比最后一个查询与源查询更相似. 但是实际使用结果表明, 最后一个候选查询能够更好地帮助用户找到其需要的信息, 因此若向用户推荐最相关的查询, 如向用户推荐前 2 个候选查询, 则可能会导致用户信息需求得不到满足.

在本文的研究工作中, 我们提出了应该向用户推荐高效用性 (utility) 查询, 即该查询能够更好地满足用户的信息需求. 查询的效用性定义为: 用户能够从该查询的搜索结果中获得的有用信息. 值得注意的是, 尽管效用性的概念已经出现在一些查询推荐的研究工作中<sup>[5-6]</sup>, 但是这些工作与我们研究工作中有着本质的不同. 在文献<sup>[5]</sup>中主要关注如何最优化推荐查询集的全局效用, 其提出的查询效用概念类实际上等价于边际相关性 (marginal relevance); 而

文献<sup>[6]</sup>则直接使用查询的 Click-through Rate 来作为查询的近似效用, 目的是研究如何使得推荐查询集期望的 Click-through Rate 最大化. 这些研究工作的共同特点在于其并没有真正挖掘查询的效用. 而我们的研究工作则是关注如何从用户的搜索行为中挖掘出查询的真实效用, 并基于这些学习到的效用来进行查询推荐.

效用性查询推荐研究最大的挑战在于如何挖掘各个查询的效用. 我们通过分析发现, 用户的搜索行为, 尤其是用户的查询重构行为和查询点击行为, 包含了大量有价值的查询效用信息. 因此, 在本研究工作中, 我们提出了一种新的方法, 即基于吸收态随机行走的 2 阶段效用性查询推荐方法 (two-step utility query recommendation method based on absorbing random walk, TARW) 同时对用户搜索过程中的查询重构行为和查询点击行为进行建模, 并根据学习到的各个查询的效用进行查询推荐.

为了评价 TARW 方法的表现, 我们在一个公开查询日志上将其与其他 5 种基准查询推荐方法进行了比较. 此外, 我们提出了 2 种客观的效用性评价指标: 查询相关率 (query relevant ratio, QRR) 和平均相关文档数 (mean relevant document, MRD), 对各种方法的查询推荐结果进行自动评价. 实验结果表明, 和其他方法相比, 我们提出的 TARW 方法在所有的评价指标上显著优于所有基准方法.

1 相关工作概述

现有的查询推荐研究工作主要基于用户查询日志来挖掘查询之间的相似关系, 其中查询 Click-through 信息<sup>[3]</sup>、查询会话信息<sup>[4]</sup>是最常使用的信息类型. 传统的查询推荐方法主要关注向用户推荐相关查询, 以帮助用户找到其所需要的信息, 例如 Wen 等人<sup>[2]</sup>同时考虑使用 Click-through 数据, 查询文档的内容信息来发现相似查询; Beeferman 等人<sup>[3]</sup>通过在 Query-URL 二部图上使用凝聚聚类算法来发现相关查询; Li 等人<sup>[7]</sup>提出了一个 2 阶段的查询推荐方法: 发现阶段和排序阶段. 在发现阶段, 使用基于 Query-URL 向量来计算查询之间的相似度; 在排序阶段, 使用层次凝聚聚类算法来对相似查询进行排序; Ma 等人<sup>[8]</sup>基于 Click-through 数据构

建了 2 个二部图,即 User-Query 二部图和 Query-URL 二部图,然后提出了一种 2 层的查询推荐方法;Zhang 等人<sup>[4]</sup>首次提出将用户的顺序查询行为建模为查询图(query graph),并根据查询会话信息来计算查询之间的相似性;Boldi 等人<sup>[9]</sup>进一步基于用户的查询会话信息提出查询流程图(query flow graph)的概念,并在该图上执行随机行走方法来找到相似的查询;Mei 等人<sup>[10]</sup>基于 Click-through 数据构造 Query-URL 二部图,并在该图上使用随机行走方法,依据各候选查询到源查询的命中时间(hitting time)来对查询进行相关性排序.近年来,一些研究工作<sup>[11-12]</sup>开始关注差异性查询推荐,但是其本质也是在相关性基础上进一步考虑减少推荐结果的冗余性来实现差异性推荐.例如 Zhu 等人<sup>[11]</sup>提出扩展流型排序方法来进行差异性查询推荐;Guo 等人<sup>[12]</sup>提出结构化查询推荐方法,将推荐的查询进一步按照搜索意图分类后展示给用户.

随机行走方法目前已经被广泛应用<sup>[9]</sup>,并取得了很好的实际应用效果.近年来,研究人员开始在随机行走过程中引入吸收态(absorbing state),提出了一类新的随机行走模型,即吸收态随机行走模型,并开始应用到多个领域<sup>[13-15]</sup>.例如 Singh 等人<sup>[13]</sup>提出在 User-User 和 User-Item 扩展的二部图上,通过将 item 结点置为吸收态结点来进行社会推荐;Sadikov 等人<sup>[14]</sup>则是将文档看成吸收态结点,计算得到各个查询的文档表示,并用来进行查询聚类;Zhu 等人<sup>[15]</sup>通过迭代地将已选取的结点置为吸收态来实现差异性排序.我们的方法与现有方法的区别在于:1)研究问题不同,我们首次将吸收态随机行走模型应用到效用性查询推荐研究问题中;2)我们的方法包含 2 个阶段:在正向阶段中,使用吸收态随机行走过程计算出各个点击文档的效用,然后在反向阶段中,根据学习到的点击文档效用来计算出查询的效用.

## 2 扩展查询流程图

Boldi 等人<sup>[9]</sup>首次提出了查询流程图的概念,用来描述用户的查询重构行为.在我们的研究工作中,为了能够同时描述用户的查询重构行为和查询点击行为,我们进一步对查询流程图进行扩展,在查询流程图中引入一类新的结点,即文档结点,用来描述用户点击行为,从而得到一种新的扩展的查询流程图(在文献<sup>[14]</sup>中,Sadikov 等人也使用相似的方式来

表示查询).在该扩展查询流程图上,查询结点之间的连边 $(q, q')$ 表示查询  $q'$  对查询  $q$  进行了重构,而查询与文档之间的连边 $(q, d)$ 表示用户点击了查询  $q$  搜索结果中的文档  $d$ .扩展查询流程图的形式化定义如下:

给定查询集合  $Q = \{q_1, q_2, \dots, q_n\}$  和文档集合  $D = \{d_1, d_2, \dots, d_m\}$ ,我们构造图  $G = (Q, D, E, F)$ ,其中,  $E$  表示查询与其重构查询之间的连边集合,  $F$  表示查询与其点击文档之间的连边集合,即:

$$E \triangleq \{(q, q') : q \in Q, q' \in R(q)\},$$

$$F \triangleq \{(q, d) : q \in Q, d \in C(q)\},$$

其中,  $R(q)$  表示查询  $q$  的重构查询集合,  $C(q)$  表示用户点击查询  $q$  搜索结果的文档集合.

我们进一步得到扩展查询流程图的邻接矩阵  $W$ , 表示为

$$W = \begin{bmatrix} W_E & W_F \\ W_F^T & \theta \end{bmatrix},$$

其中,  $W_E$  和  $W_F$  分别对应集合  $E$  和  $F$  的邻接矩阵,边  $w(q, q') \in W_E$  的权重定义为查询  $q'$  对查询  $q$  进行重构的次数,边  $w(q, d) \in W_F$  的权重定义为用户点击查询  $q$  搜索结果中文档  $d$  的次数.

## 3 基于吸收态随机行走的两阶段模型

基于以上得到的扩展查询流程图,我们可以将其转换成一个 Markov 链,并通过引入吸收态结点,提出了一种新的基于吸收态随机行走的 2 阶段模型,该模型包括正向效用传播阶段和反向效用传播阶段(如图 1 所示).在正向效用传播阶段,我们将扩展查询流程图中的查询结点看成是瞬态(transient state)结点,而将文档结点看成是吸收态(absorbing state)结点,从而得到一个正向的吸收态 Markov 链,即效用得分从源查询经重构查询并最终被文档结点吸收,在这一阶段中,我们能够学习到各个文档的效用.在反向效用传播阶段,我们根据已经学习到的文档效用,将其反向传播到查询结点,从而得到各个查询的效用,并以此来进行查询推荐.

该方法可以直观地理解为:在正向的吸收态随机行走过程中(即效用从源查询到文档的传播过程),随机行走者从源查询出发,其代表了从源查询释放出一个特定的查询需求,这个查询需求通过查询重构结点进行传播,并最终被各个文档结点吸收,文档结点的得分就表示该文档结点向用户提供多少有价值的信息,其得分越高表示该文档效用越大.其次,在反向效用传播阶段,文档结点将其在前一阶段

中获得的效用,即该文档所能提供的有用信息,反向传播给查询结点,从而得到各个查询结点的得分,该

得分反映了用户使用该查询结点能够找到多少有用的信息,对应于用户能够点击到哪些有用的文档。

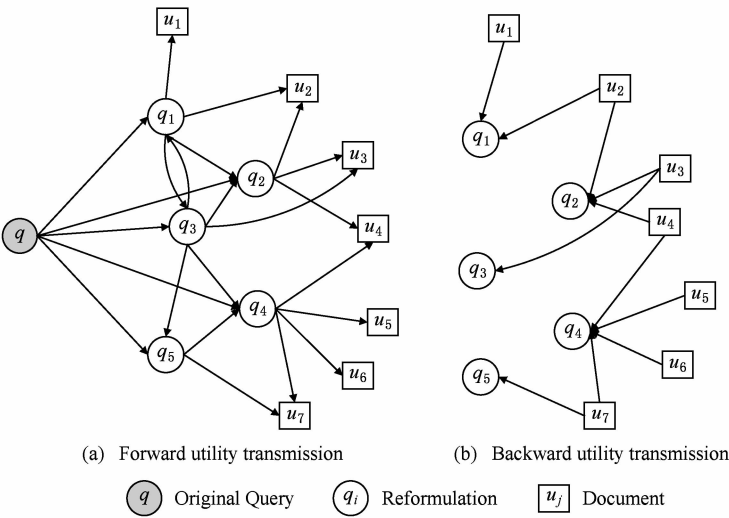


Fig. 1 A two-step model based on absorbing random walk.  
图1 基于吸收态随机行走的两阶段模型

以下我们分别给出正向效用传播阶段和反向效用传播阶段详细计算过程。

3.1 正向效用传播阶段

在正向效用传播过程中(如图 1(a)所示),随机行走者从源查询结点(代表用户查询意图)出发,并以  $\alpha$  的概率行走到其邻居查询结点,以  $(1-\alpha)$  的概率行走到其点击文档结点. 我们可以将  $\alpha$  理解为重构查询的得分比重,  $\alpha$  越大表示随机行走者将以更大的概率行走到其重构查询,反映了随机行走者更相信重构查询要比当前的查询更能满足用户的查询信息需求. 而  $\alpha$  越小表示随机行走者将以更大的概率行走到当前查询的点击文档,反映了随机行走者认为当前的查询更能满足用户的信息需求. 我们在实验中分析了  $\alpha$  的选择对结果的影响,总的来讲,  $\alpha$  取较大值时,我们的方法效果更好(在本文中,我们经验设置参数  $\alpha=0.95$ ). 这也与实际用户搜索体验一致,即用户通常因为对当前查询的搜索结果不满意才会选择进行查询重构,因此重构查询通常能够比当前查询提供更多有用的信息。

当随机行走者从查询  $q_i$  转移到其邻居查询  $q_j$  时,相应的转移概率为查询  $q_i$  被查询  $q_j$  重构的次数占查询  $q_i$  被所有查询重构的次数比例,形式化为

$$P(q_j | q_i) = \alpha \times \frac{\omega(q_i, q_j)}{\sum_{q_k \in R(q_i)} \omega(q_i, q_k)},$$

其中,  $\omega(q_i, q_j)$  表示查询  $q_i$  被查询  $q_j$  重构的次数。

当随机行走者从查询  $q_i$  转移到其点击文档  $d_j$

时,其转移概率为文档  $d_j$  被点击次数所占查询  $q_i$  所有文档被点击次数的比例,形式化为

$$P(d_j | q_i) = (1 - \alpha) \times \frac{\omega(q_i, d_j)}{\sum_{d_k \in C(q_i)} \omega(q_i, d_k)},$$

其中,  $\omega(q_i, d_j)$  表示用户使用查询  $q_i$  时点击文档  $d_j$  的次数。

由于各个文档结点  $d_j$  是吸收态结点,因此传播到文档结点的得分将不再传播给任何其他结点,形式化为

$$P(d_j | d_j) = 1.$$

此外,为了保证每个结点都能够至少到达一个吸收态结点,我们对没有任何文档点击的查询进行处理,使其  $(1-\alpha)$  的得分以均匀概率跳转到各个文档结点. 同时我们还对没有任何重构的查询进行处理,使其  $\alpha$  的得分以均匀概率跳转到任意查询结点。

3.2 反向效用传播阶段

在反向效用传播阶段(如图 1(b)所示),各个文档结点将其获得的得分反向传播给各个查询结点。

我们可以定义多种文档到查询的得分传播方式,例如文档之间可以按照相关查询的权重分配进行传播,或者直接使用查询点击文档的效用之和作为查询的效用. 在本文中,为了简单考虑,我们使用第 2 种方式来计算查询的效用,这样做是因为查询的效用取决于其点击文档所能提供的有用信息,因此用点击文档得分总和来表示查询的效用是合适的。

这种处理方式虽然简单,但是实验结果显示该方式能够取得很好的查询推荐效果.此外,由于源查询并不推荐给用户,因此在此过程中不需要计算源查询结点的效用.

3.3 计算吸收态分布

在正向效用传播阶段中,对应的转移概率矩阵表示为

$$P = \begin{bmatrix} P_Q & P_D \\ \mathbf{0} & I \end{bmatrix},$$

其中,  $P_Q$  为查询之间的概率转移矩阵,  $P_D$  为查询到文档的概率转移矩阵,  $I$  是单位阵. 因为该矩阵可约简(reducible),即不存在稳态分布,因此要计算各个吸收态得分分布,最常用的方法就是使用迭代方式直接计算:

$$P^t = \begin{bmatrix} P_Q & \sum_{k=0}^{t-1} P_Q P_D \\ \mathbf{0} & I \end{bmatrix},$$

其中,元素  $P^t[i, j]$  就表示结点  $i$  在经过  $t$  步到达结点  $j$  的概率. 由于我们只需要计算查询到文档结点的概率,即只需计算  $P^t$  右上角矩阵  $\sum_{k=0}^{t-1} P_Q^k P_D$ , 其计算复杂度为  $O(tn^3 + n^2m)$ , 其中  $n$  为查询结点数目,  $m$  为文档结点数目. 此外,在我们的研究工作中,我们只关注源查询到各个文档结点的概率,即只需要源查询所对应行中文档的概率分布. 令  $v$  为  $P_Q$  中源查询所对应的行,即  $v$  是一个  $1 \times n$  的行向量,因此我们可以只需要计算  $vP_Q^{t-1}$ , 而不再需要计算  $P_Q^t$ , 因此最终的计算复杂度为  $O(tn^2 + nm)$ .

在每轮传播过程中,将有  $1 - \alpha$  的得分会被文档结点吸收,因此在  $t$  轮迭代之后,将有  $(1 - \alpha^t)$  的得分被文档结点吸收,例如当  $\alpha = 0.8$  时只需要迭代 20 次,即  $t = 20$  就可以有 99% 的得分被文档结点吸收.

4 实验结果及讨论

在本节中我们对文中提出的效用性查询推荐模型进行评价. 我们首先介绍了实验中所使用的数据集、评价指标和基准方法,然后比较不同方法的实验结果并进行讨论.

4.1 数据集

为了合理评价各种方法在效用性查询推荐方面的表现,我们需要事先知道在给定用户搜索意图的条件下各个查询的效用,即查询为满足用户查询需求能够提供多少有用的信息. 因此,我们在实验中所使用的数据集必须满足 2 个条件:1) 需要知道各个查询背后用户真实的搜索意图; 2) 查询返回的搜索结果中显式地给出满足用户搜索意图的相关文档信息. 然而,在查询推荐常用的数据集,如 AOL<sup>①</sup> 和 Microsoft 2006 RFP<sup>②</sup> 中缺乏这些重要的信息. 在本实验中,我们使用埃默里大学(Emory University)在 2011 年公开的 UFindIt<sup>③</sup> 查询日志数据集. 该数据集共收集了 6 个月的用户查询日志,记录了 40 个查询任务,分别对应着 40 个不同的测试查询. 这些测试查询是从社区问答(community question answering)网站,如 wiki.answers.com 和 Yahoo!Answers 挑选出来的真实用户查询. 该数据集记录了用户的查询重构行为、查询点击行为以及用户答案提交行为,以及标注人员对用户所有点击的 URL 进行了人工标注的结果,标注是依据该 URL 是否与用户查询意图相关分为 2 个类别:相关和不相关. 为了模拟真实的搜索过程,我们需要忽略用户的答案提交行为,因为在真实的搜索过程中,当用户的信息需求得到满足后并不会向系统提交答案.

我们进一步对 UFindIt 数据集进行处理,包括过滤掉一些任务交叉的查询会话,即用户试图在一个查询会话中同时完成多个不同的查询任务. 此外,我们过滤掉没有重构查询或者不是以查询开始的会话(正常的查询会话应该是用户先提交查询,然后点击 URL). 在经过以上处理之后,我们得到数据集的统计信息如表 1 所示:

Table 1 Statistics of Dataset  
表 1 数据集统计信息

Name	Size
Number of search sessions	1 298
Number of distinct queries	1 086
Number of distinct URLs	1 555
Average number of search sessions for each test query	32
Average number of distinct reformulations for each test query	26

① <http://www.gregsadetsky.com/aol-data/>  
② <http://research.microsoft.com/users/nickcr/wscd2009>  
③ <http://ir-ub.math.cs.emory.edu/uFindIt/>

## 4.2 评价指标

与传统的相关性查询推荐的评价目标不同,效用性查询推荐评价的目标更加关注用户在使用查询之后的实际体验,即不仅评价所推荐的查询是否相关,而是进一步关注用户使用该查询之后是否能够从该查询的检索结果中得到有用的信息以满足其信息需求.传统的相关性查询推荐评价指标,例如 Precision 和 Recall<sup>[16]</sup>,通常根据人工对所推荐的查询进行相关性标注来进行评价<sup>[17]</sup>,这些指标已经不再适合用来评价效用性查询推荐结果.在本研究工作中,我们定义 2 个评价指标来对各种方法的实际推荐效果进行评价:QRR 和 MRD.

对于给定的信息需求,评价指标 QRR 定义为

$$QRR(q) = \frac{RQ(q)}{N(q)},$$

其中,  $RQ(q)$  表示查询  $q$  作为重构时,用户点击到相关文档的总重构次数,  $N(q)$  表示查询  $q$  作为查询重构的次数.该指标的含义是当用户使用查询  $q$  来实现其信息需求时,其有多大概率能够点击到相关结果.  $QRR$  的值越大表示用户越可能找到有用的结果<sup>①</sup>.

对于给定的信息需求,评价指标 MRD 定义为

$$MRD(q) = \frac{RD(q)}{N(q)},$$

其中,  $RD(q)$  表示查询  $q$  作为重构时,用户点击到的相关文档数目.该指标的含义是当用户使用查询  $q$  来实现其信息需求时,其平均点击到相关文档的数目.  $MRD$  的值越大表示用户找到的相关信息就越多.

## 4.3 基准方法

为了比较我们提出的 TARW 方法的表现,我们将其与 5 种基准方法进行对比,这些基准方法可以分为 3 类.

### 1) 基于频次的方法

Adjacency (ADJ): 基于查询会话中邻近重构的方法.给定测试查询  $q$ , 使用查询会话中  $q$  的重构查询出现频次来作为查询推荐的依据<sup>[18]</sup>.

Co-occurrence (CO): 基于查询会话中共现重构的方法.给定测试查询  $q$ , 使用查询会话中与  $q$  共现查询的频次来作为查询推荐的依据<sup>[19]</sup>.

### 2) 基于图的方法

Query Flow Graph (QF): 基于查询流程图的方法是由 Boldi 等人<sup>[9]</sup>提出的,是当前查询推荐研

究中具有代表性的方法.该方法首先基于查询会话构造查询流程图,然后在该图上执行随机行走,并以此来进行查询推荐.

Click-through Graph (CT): 基于 Click-through 图的方法是由 Mei 等人<sup>[10]</sup>提出的,也是当前查询推荐研究中具有代表性的方法.其基本思想是首先根据查询日志构造一个 Query-URL 的二部图,然后在该图上执行一个随机行走方法,并根据各个候选结点到源查询结点的命中时间(hitting time)来作为查询推荐的依据.

### 3) 基于 Click-through Rate 的方法

Click-through Rate (CTR): 该方法<sup>[6]</sup>根据查询搜索结果的期望 Click-through Rate 来进行查询推荐.其基本假设是认为用户更愿意点击其认为是相关的搜索结果,因此 CTR 反映了用户对该查询的感知效用.

## 4.4 在所有测试查询上的评价结果

图 2 展示了在 2 个评价指标 QRR 和 MRD 下,各种方法的查询推荐的结果对比情况.从图 2 可以看出,2 种基于频次的方法 ADJ 和 CO 的推荐效果最差.这说明了简单使用与源查询在相同查询会话中的频繁邻近出现或共现的候选查询(这些查询通常与用户源查询非常相关),并不能保证能够向用户推荐高效用性的查询以满足用户的信息需求.和基于频次的方法相比,基于图的方法如 QF 和 CT 表现出更好的推荐效果.这说明通过综合查询在图上局部的相关性,如在查询会话中共现或者查询重构相关性来得到查询之间全局的相关性,能够得到更好的查询推荐效果.CTR 方法能够获得与基于图的方法相似的推荐效果.但是查询的实际效用只有在用户浏览相应的文档内容之后才能确定该文档是否能够提供有用的信息,因此仅仅依据 Click-through Rate 并不能真正反映该查询是否真正有用,这也是为什么 CTR 方法在 2 个评价指标上并不能一致性地取得很好的表现.和所有的基准方法相比,我们提出的 TARW 方法一致地取得最好的评价结果.统计显著性测试( $T$ -Test,  $p$ -value  $\leq 0.05$ )结果表明 TARW 方法和基准方法相比,在所有指标上面都有显著性提高.这说明使用基于吸收态随机行走的 2 阶段模型来同时建模用户的查询重构行为和查询点击行为,可以更好地挖掘查询的实际效用,从而帮助用户找到其所需的信息.

① 在本文中,为了减小观察数据不均衡带来的影响,例如  $RQ(q)$  和  $N(q)$  分别为 (2, 2) 以及分别为 (20, 20) 时,得到的结果均为 1,但显然后一个结果更好.在实际应用中,我们对结果进行平滑,即  $QRR(q) = (RQ(q) + 1) / ((RQ(q) + 1) + (UQ(q) + 1)) = (RQ(q) + 1) / (N(q) + 2)$ , 其中  $UQ(q)$  表示查询  $q$  作为重构查询时,用户未点击到相关文档的重构次数.同样,我们对指标 MRD 也进行相似的处理.

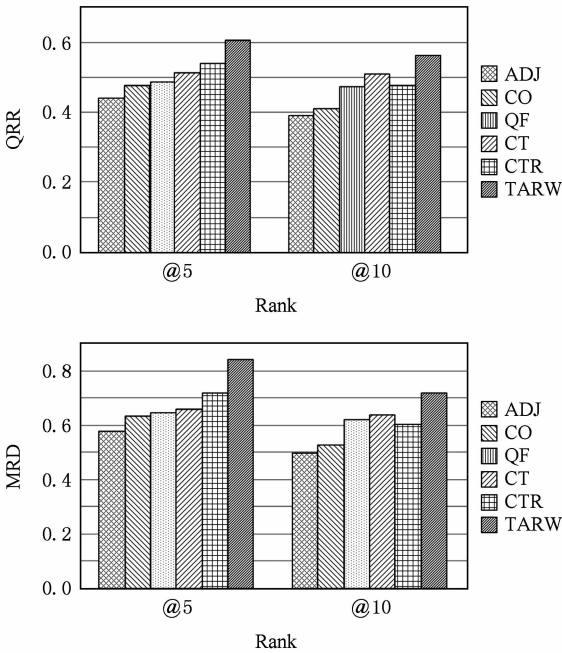


Fig. 2 Comparison of the performance of all approaches in terms of QRR and MRD.

图 2 各方法在评价指标 QRR 和 MRD 上表现比较

4.5 按查询难度划分的评价结果

为了进一步分析 TARW 方法在推荐高效用性查询方面的表现,我们将测试查询按照其困难程度分为 3 个类别:简单查询、中等查询和困难查询,然后对不同方法在各个查询难度上面的效用性推荐的表现进行比较.表 2 中展示了在评价指标 QRR 和 MRD 上,各种方法在不同查询难度上的表现.其中括号中的百分比表示我们的方法与相应方法相比在该评价指标上提高的比率.从实验结果可以看出,随着查询困难程度的增加,所有方法的表现都呈现出下降的趋势.但是,同时也可以看到我们提出的 TARW 方法在所有的查询难度上面都一致性地优于所有基准方法.而且,一个有趣的现象就是,随着查询困难程度的增加,TARW 方法和基准方法相比提高的幅度更大.例如,在评价指标 MRD@10 上,当查询简单时,基于频次的方法的值接近于 0.7,基于图的方法和基于 Click-through Rate 的方法都在 0.8 左右,而 TARW 方法则达到 0.844.这说明对于简单查询而言,大部分的查询推荐方法表现的差别并不大,基准方法也能够推荐较好的查询.然而,随着查

Table 2 Performance Comparison of All Methods under Different Query Difficulty Levels

表 2 所有方法在不同查询难度上的表现比较

Query Difficulty	Method	QRR		MRD	
		@5(num/%)	@10(num/%)	@5(num/%)	@10(num/%)
Easy	ADJ	0.588(16.79)	0.526(25.97)	0.771(28.11)	0.674(27.86)
	CO	0.609(12.76)	0.529(25.31)	0.830(19.04)	0.687(25.48)
	QF	0.618(11.17)	0.604(9.61)	0.846(16.77)	0.806(6.90)
	HT	0.654(4.95)	0.635(4.38)	0.836(18.21)	0.805(7.00)
	CTR	0.656(4.71)	0.611(8.46)	0.889(11.11)	0.798(8.02)
	TARW	<b>0.687</b>	<b>0.663</b>	<b>0.988</b>	<b>0.862</b>
Medium	ADJ	0.460(36.59)	0.429(41.76)	0.596(44.16)	0.527(44.30)
	CO	0.495(26.93)	0.441(38.00)	0.640(34.39)	0.550(38.19)
	QF	0.511(23.00)	0.500(21.76)	0.615(39.85)	0.630(20.60)
	HT	0.534(17.75)	0.549(10.72)	0.689(24.88)	0.692(9.83)
	CTR	0.544(15.48)	0.485(25.32)	0.703(22.29)	0.588(29.19)
	TARW	<b>0.628</b>	<b>0.608</b>	<b>0.860</b>	<b>0.760</b>
Hard	ADJ	0.259(92.00)	0.216(94.29)	0.351(87.62)	0.284(89.87)
	CO	0.314(58.32)	0.261(60.74)	0.412(59.98)	0.340(58.52)
	QF	0.324(53.43)	0.312(34.35)	0.441(49.42)	0.414(30.43)
	HT	0.334(48.79)	0.343(22.13)	0.437(50.89)	0.424(27.29)
	CTR	0.404(23.02)	0.324(29.13)	0.534(23.40)	0.413(30.69)
	TARW	<b>0.497</b>	<b>0.419</b>	<b>0.659</b>	<b>0.540</b>

Note: The number in parentheses presents the improvements of TARW over the corresponding methods

询难度的提高,基准方法的表现和 TARW 方法相比下降的幅度更大. 例如基于频次的方法,即 ADJ 和 CO 在评价指标 MRD@10 上分别下降到 0.284 和 0.34,基于图的方法,QF 和 CT 分别下降到 0.414 和 0.424,基于 Click-through Rate 的方法,即 CTR 下降到 0.413,而我们提出的 TARW 方法在指标 MRD@10 上仍然达到 0.540. 与基准方法相比,TARW 方法比基于频次的方法提高的幅度超过 50%,比基于图的方法和基于 Click-through Rate 的方法提高的幅度分别达到 30%左右. 这说明当用户的查询任务困难时,我们的 TARW 方法能够更好地挖掘查询的效用. 在评价指标 MRD@5, QRR@5 和 QRR@10 上,我们也得到相似的评价结果.

## 5 结束语

本文中,我们针对查询推荐的最终目标是满足用户信息需求,首次提出应该向用户推荐高效用性查询,而不仅仅只关注向用户推荐相关性查询. 我们基于查询日志中的用户查询重构行为和查询点击行为,提出了一种新的模型来同时对这 2 种信息进行建模并挖掘出查询的效用. 此外,我们还提出了 2 个的查询效用性度量指标来自动评价各种方法的查询推荐效果. 在真实查询日志上的实验结果表明,和基准方法相比,我们的方法能够更好地挖掘查询的效用,从而更好地帮助用户实现其查询需求.

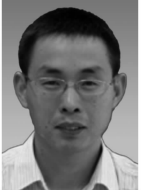
在我们的模型中,如何选择参数  $\alpha$  是一个值得研究的问题,在下一步研究工作中,我们将提出一种新的方法来自动进行参数  $\alpha$  的选择. 此外,如何实现所推荐的一组查询的效用最大化,即减小推荐查询之间的冗余效用也是一个非常重要的研究问题.

## 参 考 文 献

- [1] Li Yanan, Wang Bin, Li Jintao. A survey of query suggestion in search engine [J]. Journal of Chinese Information Processing, 2010, 24(6): 75-84 (in Chinese)  
(李亚楠, 王斌, 李锦涛. 搜索引擎查询推荐技术综述[J]. 中文信息学报, 2010, 24(6): 75-84)
- [2] Wen Jirong, Nie Jianyun, Zhang Hongjiang. Clustering user queries of a search engine [C] //Proc of the 10th Int Conf on World Wide Web. New York: ACM, 2001: 162-168
- [3] Beeferman D, Berger A. Agglomerative clustering of a search engine query log [C] //Proc of the 6th ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2000: 407-416
- [4] Zhang Zhiyong, Nasraoui O. Mining search engine query logs for query recommendation [C] //Proc of the 15th Int Conf on World Wide Web. New York: ACM, 2006: 1039-1040
- [5] Jain A, Ozertem U, Velipasaoglu E. Synthesizing high utility suggestions for rare web search queries [C] //Proc of the 34th Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2011: 805-814
- [6] Anagnostopoulos A, Becchetti L, Castillo C, et al. An optimization framework for query recommendation [C] //Proc of the 3rd ACM Int Conf on Web Search and Data Mining. New York: ACM, 2010: 161-170
- [7] Li Lin, Yang Zhenglu, Liu Ling, et al. Query-url bipartite based approach to personalized query recommendation [C] //Proc of the 23rd National Conf on Artificial Intelligence. Menlo Park, CA: AAAI, 2008: 1189-1194
- [8] Ma Hao, Yang Haixuan, King I, et al. Learning latent semantic relations from clickthrough data for query suggestion [C] //Proc of the 17th ACM Conf on Information and Knowledge Management. New York: ACM, 2008: 709-718
- [9] Boldi P, Bonchi F, Castillo C, et al. The query-flow graph: Model and applications [C] //Proc of the 17th ACM Conf on Information and Knowledge Management. New York: ACM, 2008: 609-618
- [10] Mei Qiaozhu, Zhou Dengyong, Church K. Query suggestion using hitting time [C] //Proc of the 17th ACM Conf on Information and Knowledge Management. New York: ACM, 2008: 469-477
- [11] Zhu Xiaofei, Guo Jiafeng, Cheng Xueqi, et al. A unified framework for recommending diverse and relevant queries [C] //Proc of the 20th Int Conf on World Wide Web. New York: ACM, 2011: 37-46
- [12] Guo Jiafeng, Cheng Xueqi, Xu Gu, et al. A structured approach to query recommendation with social annotation data [C] //Proc of the 19th ACM Int Conf on Information and Knowledge Management. New York: ACM, 2010: 619-628
- [13] Singh A P, Gunawardana A, Meek C, et al. Recommendations using absorbing random walks [C/OL] //Proc of the 2nd North East Student Colloquium on Artificial Intelligence (NESCAI07), 2007. [2011-05-17]. <http://research.microsoft.com/pubs/76151/singh2007.pdf>
- [14] Sadikov E, Madhavan J, Wang Lu, et al. Clustering query refinements by user intent [C] //Proc of the 19th Int Conf on World Wide Web. New York: ACM, 2010: 841-850
- [15] Zhu Xiaojin, Goldberg A B, Gael J V, et al. Improving diversity in ranking using absorbing random walks [C] //Proc of the 2007 Conf on the North American Chapter of the Association for Computational Linguistics Human Language Technologies. Stroudsburg: ACL, 2007: 97-104

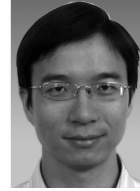


- [16] He Qi, Jiang Daxin, Liao Zhen, et al. Web query recommendation via sequential query prediction [C] //Proc of the 2009 IEEE Int Conf on Data Engineering. Los Alamitos, CA: IEEE Computer Society, 2009: 1443-1454
- [17] Cao Huanhuan, Jiang Daxin, Pei Jian, et al. Context-aware query suggestion by mining click-through and session data [C] //Proc of the 14th ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2008: 875-883
- [18] Jones R, Rey B, Madani O, et al. Generating query substitutions [C] //Proc of the 15th Int Conf on World Wide Web. New York: ACM, 2006: 387-396
- [19] Huang Chienkang, Chien Leefeng, Oyang Yenjen. Relevant term suggestion in interactive web search based on contextual information in query session logs [J]. Journal of the American Society for Information Science and Technology, 2003, 54(7): 638-649



**Zhu Xiaofei**, born in 1979. PhD. He is currently a postdoctoral researcher at L3S Research Center, Leibniz Universitaet Hannover, Germany. His major research interests include information retrieval and

data mining.



**Guo Jiafeng**, born in 1980. PhD. He is currently an associate professor in Institute of Computing Technology, Chinese Academy of Sciences. His major research interests include web search and mining, user data mining, machine learning, and social network (guojiafeng@ict.ac.cn).



**Cheng Xueqi**, born in 1971. PhD. He is currently a professor and PhD supervisor in Institute of Computing Technology, Chinese Academy of Sciences. His major research interests include network information security, large-scale information retrieval and knowledge mining(cxq@ict.ac.cn).



**Lan Yanyan**, born in 1982. PhD. She is currently an assistant researcher in Institute of Computing Technology, Chinese Academy of Sciences. Her major research interests include machine learning and information retrieval(lanyanyan@ict.ac.cn).