

信息类、导航类与事务类查询个性化潜力的对比析究^{*}

张晓娟

(西南大学计算机与信息科学学院, 重庆 400715)

摘要: 首先分别利用显式与隐式评测指标分析信息类、导航类与事务类查询的个性化潜力, 然后通过对基于人工评测的显式指标与基于用户点击行为的隐式指标间的相关性分析, 来验证各隐式指标的有效性。在此基础上, 本文针对不同查询意图, 分析各特征表征个性化潜力的有效性。最后, 根据实验结果对搜索引擎性能优化提供相关建议。

关键词: 信息类查询; 导航类查询; 事务类查询; 个性化潜力

中图分类号: G353.4

DOI: 10.3772/j.issn.1673-2286.2017.09.006

1 引言

自Broder按照用户意图(或用户任务)将查询划分为信息类、导航类与事务类三大类别^[1], 学界对如何选择分类特征实现三大查询类别间的有效区分进行大量研究^[2-4]。对查询进行分类的最终目的是使搜索引擎根据不同类别查询特征返回不同检索结果; 而如何对归类后查询的特征进行分析, 并根据其特征为搜索引擎性能优化提供相关依据是重要研究方向。

由于网络用户对搜索引擎的返回结果呈现个性化需求趋势, 基于访问大众性的检索排序方法已不能满足用户的信息需求^[5]。为更大程度地提高用户满意度, 搜索引擎尝试根据用户个人背景与兴趣爱好提供个性化检索服务。但当前搜索引擎的个性化检索建模是对所有检索词采用相同的检索技术, 未考虑某些查询是否适合个性化检索。Dou等研究发现, 个性化检索技术并非适合所有查询, 某些查询采用个性化检索技术反而会降低检索结果的用户满意度^[6]。如查询“百度”, 用户更多关注百度网站主页, 此类查询适合采用基于大众访问的排序方法; 查询“苹果”, 不同用户对该查询

结果感兴趣的内容存在差异, 包括“苹果沙拉”“苹果手机6s”“苹果电脑”等不同方面, 此类查询适合采用个性化的排序方法。Dou等将用户在查询中表现的个性化差异称为查询的个性化潜力, 该差异值表明用户查询从个性化中获益的程度^[6], 一般情况下, 该差异值越大, 查询越适合个性化检索排序算法。

在信息检索中, 用户任务类别是很重要的情景因素, 直接影响用户获取信息的途径以及想要获取信息的类型。对不同任务类别查询的个性潜力进行比较分析, 有助于搜索引擎判断用户在不同任务情景中的个性化潜力情况, 并以此作为优化搜索引擎性能的重要依据, 为搜索算法的改进和检索性能测评提供方向。本文尝试对信息类、导航类和事务类查询的个性化潜力进行对比分析, 并探讨三类个性化潜力特征的异同。

2 相关研究现状

2.1 查询意图相关研究

Broder依据用户查询过程中可能包含的潜在意图

^{*} 本研究得到国家社会科学基金青年项目“融合用户个性化与实时性意图的查询推荐模型研究”(编号: 15CTQ019)资助。

(或任务)将查询分为信息类、导航类和事务类^[1]。信息类查询指用户提交此类查询时想获得某主题的相关信息,其信息需求既可以是精确的又可以是模糊的,如查询“如何网上购物”;导航类查询指用户提交此类查询时想查找特定网站,该网站可以是个人网站也可以是组织网站等,如查询“武汉大学主页”;事务类查询指用户提交此类查询时想获取资源或网络服务。在Broder研究基础上,此领域的主流研究是探讨如何选取有效分类特征对三大意图类别查询进行自动分类^[2-4]。

还有学者尝试对信息类、导航类和事务类查询进行比较分析,以此提出各类查询所适合的检索模型。如Craswell等通过实证研究发现导航类查询更适合基于链接排序的检索模型^[7];Li等研究发现识别具有事务意图的网页,有利于改善事务类查询的检索结果^[8];Fujii研究发现信息类查询适合基于内容的检索方法,而导航类查询适合于基于锚文本的检索方法^[9]。

2.2 查询个性化潜力的相关研究

Dou等提出“查询个性化潜力”概念后,如何采用相关方法来对其进行衡量成为热点研究问题^[6]。宋超采用Kappa一致性检验衡量查询词的个性化潜力,并对查询词的个性化潜力分布进行分析,结果表明,大多数查询词的个性化潜力较大^[5];Chen等探讨自动分类方法与回归方法哪种更适合查询词的个性化潜力识别研究,结果表明自动分类方法优于回归方法^[10];陈晨等利用Wikipedia挖掘查询中的语言学特征,以此预测查询的个性化潜力,降低查询日志数据稀疏带来的影响^[11];Teewan等提出基于人工评测的显式指标与基于用户点击的隐式指标,并通过对二者的相关性分析,探讨隐式指标的有效性^[12];Liu等提出应针对不同用户任务构建不同的个性化检索模型^[13];Li等基于中英文查询日志数据集,将用户重复提交的查询分为共同查询(不同用户提交同一查询后点击相同的URL)与个体查询(不同用户提交同一查询后点击不同的URL),再计算各类查询的个性化潜力,结果表明个体查询更能从个性化排序算法中获益^[14]。

大多数情况下,语义模糊(或歧义)查询更能从个性化中获益^[15],故查询模糊性(或歧义性)识别是个性化潜力识别的另一相关研究。Song等在探讨查询模糊性时,将查询分为模糊查询、宽泛查询以及清晰查询三类,且在人工标注的250个查询基础上,通过采用不同

计算方法来衡量查询文档间的相似度来获取相关分类特征,采用监督式学习方法实现查询模糊类与非模糊类(包含宽泛查询与清晰查询)的区分^[15-16];Wang等提出基于点击熵的相关特征,分别利用朴素贝叶斯、逻辑回归支持向量机模型来训练分类模型识别模糊性查询^[17];Sanderson等探讨Wikipedia与WordNet中模糊性查询词的特征,结果发现就非模糊性查询词而言,Wikipedia中的模糊性查询包含较多参照网页,而WordNet中的模糊查询包含更多的释义^[18];Luo等提出从查询表达式、用户点击行为和查询会话中选取相关特征,利用SVM分类器训练分类模型来识别模糊性查询^[19]。综合已有研究可知,目前没有对信息类、导航类与事务类查询的个性化潜力进行对比分析的相关研究。

3 查询个性化潜力指标的操作方法

本文采用两种方法衡量查询个性化潜力^[12]:显式评测方法,让用户根据其信息需求来对检索结果的相关性进行人工评判;隐式评测方法,即从查询日志中自动收集与用户兴趣相关的点击信息,并将其视为用户对查询结果的相关性评测依据。

3.1 显式评测方法

(1) 基于人工评测的Potential@N指标。Potential@N指标的主要思想:首先让一组人数为N的评测组对某查询返回结果的相关性进行评测,然后利用NDCG指标衡量每位评测者的评测结果,最后计算个体评测者最佳排序的NDCG值与团体评测者最佳排序的NDCG值的差值,该差值即Potential@N^[12]。NDCG值计算方法参见公式(1)和公式(2)所示。在公式(1)中, i 表示文档在实际结果集中的排序, $G(i)$ 表示评测者对在结果集中排序为 i 的文档给予的相关性分数, $G(i)$ 取值可为2(高度相关)、1(相关)与0(不相关)。 $1/\log(i)$ 表示惩罚因子,如排 $i=2$ 时,其惩罚因子为1;排序为1 024时,其惩罚因子为1/10。公式(2)中, $IDCG$ (Ideal DCG)表示理想的DCG,即检索结果均按照相关性从高到低进行排列,说明该系统取得最优排序结果,并获得最大DCG值。

$$DCG(i) = \begin{cases} G(1) & \text{if } i=1, \\ DCG(i-1) + \frac{G(i)}{\log(i)} & \text{otherwise} \end{cases} \quad (1)$$

$$NDCG = \frac{DCG}{IDCG} \quad (NDCG \text{ 的取值在 } [0,1]) \quad (2)$$

(2) 基于人工评测的Kappa一致性指标。本文采用Fleiss等提出的Kappa指标,用于衡量固定个数评测者间评测结果一致性^[20],计算方法参见公式(3)。其中, $P(O)$ 表示评测者间实际评测结果的一致性概率,而 $P(C)$ 表示评测者间评测结果被期望的一致性概率。 k 的取值为 $[-1, 1]$,1表示二者判断结果完全一致,0表示二者判断结果是偶然性造成的,-1表示二者判断结果完全不一致。Kappa值越大,说明用户对搜索引擎返回结果的评价一致程度较高,即用户在该查询上的兴趣差异较小,查询的个性化潜力越小;反之,该值越小,查询的个性化潜力越大。

$$k = \frac{P(O) - P(C)}{1 - P(C)} \quad (3)$$

3.2 隐式评测方法

用户点击行为在一定程度上能间接反应用户兴趣差异^[21],学者也提出了基于用户点击信用的个性化潜力评测指标,此类指标被称为隐式评测指标。

(1) Dice指标。主要表示两个集合的重复情况,计算方法参见公式(4)。其中, URL_1 表示用户1的URL点击集合, URL_2 表示用户2的URL点击集合, URL_{int} 表示两用户点击集合的交集。该公式主要用于衡量两个用户对同一查询结果的点击一致性。当判断多个用户的点击结果一致性时,需分别计算任意两用户间的 Q 值,再对其进行平均,计算方法参见公式(5)。其中, $\bar{Q}$ 值越大,表明用户间兴趣一致性越高,该查询的个性化潜力越小。

$$Q = \frac{2|URL_{int}|}{|URL_1| + |URL_2|} \quad (4)$$

$$\bar{Q} = \frac{1}{n} \sum_{i=1}^n Q_i \quad (5)$$

(2) 点击熵。本文采用点击熵来衡量查询的个性

化潜力,计算方法参见公式(6)。其中, $P(q)$ 表示用户提交查询 q 后所点击的文档集合, $P(p|q)$ 表示用户提交查询 p 后点击文档 q 的概率。某查询的整体点击熵值越大,表明该查询的个性化潜力越大;反之,该查询的个性化潜力越小。

$$OverClickEntropy(q) = \sum_{p \in P(q)} -P(p|q) \log_2 P(p|q) \quad (6)$$

(3) 基于用户点击的Potential@N指标。用户点击行为是对查询结果进行相关性评判的过程,鉴于此,本文利用基于点击行为的Potentail@N($N=10$)指标计算步骤为:①从查询日志中获得某查询1天内(该查询当天至少被10个用户提交)点击量排序前20的网页,若在查询日志中缺失某排序位置上的URL,则认为该排序位置上的URL与信息需求无关;②若用户点击某文档,认为该文档与用户信息需求相关,相关性评分为“1”,反之,则认为该文档与用户信息需求无关,相关性评分为“0”;③利用个人评测结果最优NDCG值与团体评测最优NDCG值的差值获得该查询的Potential@10值。

4 实验数据

本文首先在查询日志数据中选取样本查询,再利用样本查询获得能对查询个性化潜力进行显式与隐式评测的实验数据。

4.1 查询日志数据集

数据集采用Sogou实验室于2008年6月发布的查询日志数据^[22-23],其数据格式如图1所示。从左到右代表的含义分别为用户访问时间、用户ID、查询、该URL在返回结果中的排名、用户点击的顺序号,以及用户点击的URL。

00:00:03	35804326352621896	[免费取名]	3 1	huaxia.wangzhan8.com/
00:00:03	07321773511158924	[欧美金发女郎]	2 4	a.se2222.com/Html/OPIC/index.html
00:00:03	43080219994871455	[google]	1 1	www.google.com/
.....				

图1 Sogou查询日志数据格式样例

受人力及时间限制,本文无法对查询日志数据进行逐项分析,故首先利用泊松抽样方法^[24]从Sogou日志数据集中抽取3 000个查询。这些查询满足两个条件:

(1) 在查询日志中出现频次不少于2 000;(2) 将每个查询每天所点击的URL集合求交集,交集中元素个数不少于15,将交集集中的URL称为“共有URL”。其次,要求

3位标注者对3 000个查询应属类别(信息类、导航类与事务类)进行人工标注。考虑到一个查询可能属于多种类别,因此,要求标注者标记出查询在大多数情况下应该属于的类别。在标注过程中,若某查询的用户意图类别难以判断,则由三位标注者共同商讨决定。经过人工标注,分别获得305个导航类查询,752个事务类查询和1 943个信息类查询。最后,通过随机抽样方法分别从每种查询类别中随机抽取100个样本查询用于实验分析。

4.2 人工评测与隐式评测数据集

由于用户的学科知识背景可能造成其对检索结果存在不同偏好,因此本文选取来自化学、生物、计算机、图书情报、法学以及英语专业的30名学生作为评测者(每个专业各5名)分析不同意图类别查询所表现的个性化潜力差异,首先将300个样本查询分别提交至Sogou搜索引擎,记录每次查询返回的排名前20查询结果的网页标题、结果片段以及URL信息。然后,让评测者为每次查询的前20个结果进行相关性评测打分。其中,相关性分数可为2(高度相关)、1(相关)和0(不相关)。为使得用户的评测结果不受Sogou返回结果排序的影响,将每次查询前20个查询结果以随机顺序呈现给评测者。

本文在搜集隐式评测数据集时,为探讨不同用户对同一查询结果的信息需求差异性,将300个样本查询分别从查询日志中获取其点击信息与用户信息。显式与隐式评测数据如表1所示。

表1 显式评测与隐式评测方法中的查询数、用户数以及相关性评测指标

相关性评测	用户数/人	不同查询数/个	相关性指标
显式评测	30	150	2 高度相关
			1 相关
			0 不相关
隐式评测	9 245	150	1 若文档被点击
			0 若文档没被点击

5 实验结果分析

5.1 信息类、导航类与事务类的个性化潜力分析

首先,比较信息类、导航类与事务类查询不同个性

化潜力指标的平均值(见表2)。在隐式评测指标和显式评测指标中(除Kappa值外),信息类查询的个性化潜力均大于导航类和事务类两类查询,即用户提交信息类查询时,更期望获得个性化的检索结果。主要因为用户提交导航类或者事务类查询时,用户意图比较单一。如用户提交导航类查询时,用户一般想去访问网站主页;用户提交事务查询时,用户一般想去从事某事务类查询,如下载、注册等;而用户在提交信息类查询时,想获得某个主题所有相关信息,故不同用户检索需求不同,个体兴趣差异较大。

为进一步探讨信息类、导航类与事务类查询个性化潜力的差异性,本文采用两个独立样本进行 t 检验分析(见表3)。信息类与导航类查询的个性化潜力在显式和隐式指标中均存在显著性差异;导航类查询与事务类查询无显著性差异;事务类查询与信息类查询在显式指标中存在显著性差异,在隐式指标Potential@10中存在显著性差异。

表2 信息类、导航类与事务类查询在不同指标下的平均值

相关评测指标		信息类查询	导航类查询	事务类查询
显式评测指标	基于人工评测的Potential@10值	0.550	0.120	0.220
	Kappa值	0.006	0.090	0.010
隐式评测指标	基于点击的Potential@10值	0.020	0.012	0.013
	用户点击熵	0.450	0.030	0.040
	点击Dice值	0.040	0.010	0.020

表3 信息类、导航类与事务类查询的个性化潜力差异性

评测指标		信息类与导航类(t 统计量观测值)	导航类与事务类(t 统计量观测值)	事务类与信息类(t 统计量观测值)
显式评测指标	Potential@10	29.05*	2.03	10.11*
	Kappa值	27.00*	1.05	9.10*
隐式评测指标	Potential@10	17.82*	6.93	5.91*
	Dice值	13.78*	7.32	1.23
	用户点击熵	27.22*	2.60	5.63*

注: “*”表示显著性水平, $p < 0.05$ 。

5.2 显式评测与隐式评测子指标的相关性分析

显式评测指标虽能真实反映用户的个性化差异,但需花费大量人力,应用范围具有一定局限性;隐式评测指标虽能从更大范围内获取用户数据,但其合理性还需进一步验证。基于此,本文拟通过对隐式评测指标(基

于用户点击的Dice距离值、点击熵及其Potential@10)与显式评测指标(基于用户评测的Potential@10及Kappa值)的相关性分析来验证隐式指标的有效性(见表4)。三个隐式评测指标均与两个显式评测指标间存在相关性。其中,基于隐式评测的Potentail@10指标与两个显式评测指标的相关性强度显著,用户点击熵与显式评测指标的相关性强度次之。隐式指标与Kappa指标呈现负相关的原因在于Kappa是衡量评测者之间对评测结果的相同性,而表5中的三种隐式指标是用于衡量用户针对相同查询结果的差异性。

表4 显式指标与隐式指标的相关性分析
(皮尔逊相关系数)

隐式评测指标	显示评测指标	
	Potential@10	Kappa
点击Dice值	0.45*	-0.76*
Potential@10	0.75*	-0.56*
用户点击熵	0.65*	-0.76*

注: “*”表示显著性水平, $p < 0.05$ 。

5.3 信息类、导航类与事务查询中各特征的个性化潜力分析

本文也尝试探讨能表征查询个性化潜力的特征,且进一步分析针对不同查询意图,各特征表征个性化

潜力的有效性。个性化检索的难点在于分辨用户提交查询的歧义性(模糊性),以此来确定查询在哪种程度上获益于个性化检索。对于歧义性小的查询,大多数关注信息的兴趣差异度小,查询个性化潜力小;反之,用户兴趣差异大,查询个性化潜力大。因此,学者尝试利用一些衡量查询模糊性的特征来判断查询个性化潜力。本文选取衡量查询模糊性的常用特征^[10],通过这些特征与两隐式指标间的相关性获得各指标用于衡量查询个性化潜力的有效性(见表5)。其获取查询清晰度特征,首先将查询提交至Sogou搜索引擎,获得每次查询排名前20结果的标题、摘要信息,利用公式(7)计算其Clarity(q)值。其中, $P(t|q)$ 表示查询词 t 出现在查询返回结果集中的频次, $P(t)$ 表示该词出现在索引中的概率。Clarity(q)值越大,表明该查询的模糊性越小。

$$Clarity(q) = - \sum_{Terms \ t} P(t|q) \times \log_2 \left(\frac{P(t|q)}{P(t)} \right) \quad (7)$$

从表5可见,不同查询特征与隐式指标间相关程度不同,除查询中词数外,其他特征与个性化潜力呈正相关关系。查询中包含的词数与查询个性化潜力成反比关系,说明查询中包含的词越多,其个性化潜力越大。其主要原因在于,包含词越多的查询所描述的用户信息越明确,不同的用户提交该类查询,用户间信息需求的差异性较小,此类查询的个性化潜力较小。反之,不同用户提交此类查询时,信息需求差异较大,此类查询的个性化潜力较大。

表5 信息类、导航类与事务类查询中,各隐式指标与查询特征的相关关系(皮尔逊相关系数)

	用户点击熵			Potential@10		
	信息类	导航类	事务类	信息类	导航类	事务类
查询中词数	-0.51*	-0.20*	-0.51*	-0.47*	-0.20*	-0.43*
# 查询被用户提交的次数	0.12*	-0.45	-0.06	0.22*	-0.04	-0.06
# 提交查询的不同用户数	0.23*	0.00*	0.35*	0.22*	-0.01*	0.43*
域名点击熵	0.12*	0.01*	0.02*	0.03	-0.01	-0.02
点击熵	1.00*	1.00*	1.00*	0.22*	0.87*	0.33*
Potential@10	-	-	-	1.00*	1.00*	1.00*
用户平均点击网页数	0.43*	0.03*	0.40*	0.32*	0.10*	0.30*
点击网页的平均排序	0.55*	0.13*	0.52*	0.56*	0.23*	0.22*
查询清晰度	-0.56*	-0.42*	-0.42*	-0.34*	-0.33*	-0.23*

注: “*”表示显著性水平, $p < 0.05$ 。

查询点击网页的平均排序与个性化潜力呈正相关关系,即不同用户提交同一查询后,所点击网页的排序

值越大,该查询个性化潜力越大。其主要原因在于,在不考虑相关性的情况下,用户倾向于点击排序靠前的网

页^[17]。排序靠前的网页能满足与查询相关的一般信息需求,而当用户针对查询有特定信息需求时,通常查看排序靠后的网页,因此,查询的个性化潜力越大,用户越可能点击排序靠后的网页。另外,每个用户平均点击网页数与查询的个性化潜力呈正相关关系,即用户点击的网页数越多,表明该查询的个性化潜力越大。点击熵与Potential@10也存在正相关关系。查询明晰度与隐式指标呈负相关关系,说明某查询结果质量越好,该查询的个性化潜力越小。

整体而言,各查询特征在信息类和事务类查询中与两隐式指标的相关性相差不大。而在某些特征中,导航类与信息类和事务类间存在差异性。针对导航类查询而言,事务类和信息类的查询中词数更能表明其个性化潜力;对于提交查询的不同用户数而言,该特征在事务类和信息类中与两隐式评测指标呈正相关关系,而在导航类查询中呈现负相关关系或无相关关系,其主要原因可能是用户提交导航类查询的目标较单一,用户对检索结果的偏好大体一致。因此,不同用户数不能有效表征导航类查询的个性化潜力。在导航类查询中,点击熵与Potential@10的相关性值高于在其他两类查询的相关性值,其主要原因可能在于,用户提交导航类查询的目标较明确,点击稳定倾向于与查询相关的网页主页。计算信息类查询的点击熵值时,由于用户倾向点击与查询相关的不同文档,若一半用户点击某一网页,而其他用户点击另外一网页时,其点击熵值与所有用户点击同一网页的点击熵相等。在这两个例子中,用户的相关性网页不同,即前一个例子中有两个不同网页与用户信息需求相关,后一个例子中只有一个网页与用户信息需求相关,故针对信息类查询,点击熵值不能完全有效地表明查询个性化潜力。

点击网页的平均排序在导航类中与两隐式指标(用户点击熵与Potential@10)的相关性低于在其他两类查询中的相关性。其主要原因在于,导航类网页一般在查询结果排序靠前,点击靠后网页有可能是用户漫无目的的点击行为,产生噪声数据,不能有效地描述用户信息需求。而对于信息类或者事务类查询,用户的信息需求更广泛,除排序靠前的文档外,排序靠后的文档也可能满足用户信息需求,因此,即使点击排序靠后的文档,也可表明网页与用户信息需求相关。用户平均点击网页数更能衡量事务类和信息类的用户意图,主要原因在于,用户提交导航类查询的目标较明确,但不排除存在用户点击较多网页的情况,而对于信息类或事务类查

询而言,其意图是多方面的,点击的网页能满足用户意图,因此点击网页数与用户信息需求相关。

6 总结与展望

本文首先利用隐式指标(基于用户点击的Dice值、Potential@10与用户点击熵)和显式指标(基于用户评测的Potential@10与Kappa值)分析信息类、导航类和事务类查询的个性化潜力,在此基础上,分析不同意图类别查询中能表征其个性化潜力的有效特征。结果表明,信息类查询相对其他两类查询的个性化潜力更大;词数、用户平均点击网页排序、用户平均点击网页数以及查询清晰度都能表征查询的个性化潜力;各查询特征在信息类和事务类查询中与两隐式指标的相关性相差不大,而导航类与信息类、事务类间存在差异性。如查询中词个数更能表征事务类和信息类查询的个性化潜力,点击熵更能表征导航类查询的个性化潜力,点击网页的平均排序特征更能表征信息类查询的个性化潜力。基于上述结果,本文对搜索引擎个性化性能优化提出建议。(1)导航类查询更适合共性化的检索排序。一般可通过用户对该查询返回结果网页的点击次数和停留时间来确定哪些网页能适合用户普遍的信息需求,并以此作为排序依据。(2)信息类查询一般适合个性化检索排序,可进一步通过查询的模糊性或歧义性来判断该查询适合个性化的可能性大小。(3)搜索引擎在对查询个性化潜力进行判断时,对信息类和事务类查询可从查询长度、用户点击网页的平均排序来衡量其个性化潜力,对导航类查询可多从点击熵和用户平均点击网页数来衡量其个性化潜力。

本文对信息类、导航类与事务类查询的个性化潜力进行对比分析,但仍存在不足,未来研究可尝试采取更多有效的隐式指标对查询个性化潜力进行分析;尝试采用机器学习思想,利用文本查询分类通过选取特征判断某查询是否具有个性化意图,进一步判断不同特征的有效性。

参考文献

- [1] BRODER A. A taxonomy of web search[J]. SIGIR Forum, 2002, 36(2): 3-10.
- [2] 伍大勇, 赵世奇, 刘挺, 等. 融合多类特征的Web查询意图识别[J]. 模式识别与人工智能, 2012, 25(3): 500-505.
- [3] BRENES D J, GAYO-AVELLO D, PÉREZ-GONZÁLEZ K. Survey and evaluation of query intent detection methods[C]// Proceedings of

- the 2009 Workshop on Web Search Click Data.[S.1.]:[s.n.],2009.
- [4] LIU Y,ZHANG M,RU L,et al. Automatic query type identification based on click through information[C]//Asia Conference on Information Retrieval Technology.[S.1.]:Springer Berlin,2006:593-600.
- [5] 宋超.基于Sogou日志的个性化信息检索分析与建模[D].哈尔滨:哈尔滨工业大学,2012.
- [6] DOU Z,SONG R,WEN J R.A large-scale evaluation and analysis of personalized search strategies[C]//International Conference on World Wide Web.[S.1.]:[s.n.],2007:581-590.
- [7] CRASWELL N,HAWKING D,ROBERTSON S.Effective site finding using link anchor information[C]//Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.[S.1.]:[s.n.],2001:250-257.
- [8] LI Y,KRISHNAMURTHY R,VAITYANATHAN S,et al. Getting work done on the web:supporting transactional queries[C]//Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.[S.1.]:[s.n.],2006:557-564.
- [9] FUJII A. Modeling anchor text and classifying queries to enhance web document retrieval[C]//Proceedings of the 17th International Conference on World Wide Web.[S.1.]:[s.n.],2008:337-346.
- [10] CHEN C,YANG M Y,LI S,et al. Predicting query potential for personalization,classification or regression classification or regression?[C]//SIGIR.[S.1.]:[s.n.],2010:725-726.
- [11] 陈晨,赵铁军,李生,等.基于语言学知识的查询个性化潜力预测[J].中文信息学报,2013,6(11):11-18.
- [12] TEEVAN J,DUMAIS S T,LIEBLING D J.To personalize or not to personalize: modeling queries with variation in user intent[C]//Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.[S.1.]:[s.n.],2008:163-171.
- [13] LIU C,BELKIN N,COLE M,et al. Personalization of information retrieval in different types of tasks[C]//SIGIR.[S.1.]:[s.n.],2011:26-33.
- [14] LI D,YANG M,QI H,et al. Reexamination on potential for personalization in web search[C]//Proceeding of the 23rd International Conference on Computational Linguistics.[S.1.]:[s.n.],2010:701-709.
- [15] SONG R,LUO Z,WEN J R,et al. Identifying ambiguous queries in web search[C]//International Conference on World Wide Web.[S.1.]:[s.n.],2007:1169-1170.
- [16] SONG R,LUO Z,NIE J Y,et al. Identification of ambiguous queries in web search[J].Information Processing and Management An International Journal,2009,45(2):216-229.
- [17] WANG Y,AGICHTEIN E. Query ambiguity revisited:click through measures for distinguishing informational and ambiguous queries[EB/OL].[2017-08-01].<http://www.aclweb.org/anthology/N10-1055.pdf>.
- [18] SANDERSON M. Ambiguous queries:test collections need more sense[C]//Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development.[S.1.]:[s.n.],2008:499-506.
- [19] LUO C,LIU Y,ZHANG M,et al. Query ambiguity identification based on user behavior information[J].Journal of Information Retrieval Technology,2014,8870:36-47.
- [20] FLEISS J L. Measuring nominal scale agreement among many raters[J].Psychological Bulletin,1971,76(5):378-382.
- [21] KELLY D,TEEVAN J. Implicit feedback for inferring user preference:a bibliography[J].ACM SIGIR Forum,2003,37(2):18-28.
- [22] 搜狗查询日志[EB/OL].[2017-09-14].<http://www.sogou.com/labs/dl/q.html>
- [23] LIU Y,XIE X,WANG C,et al. Time-Aware click model[J].ACM Transactions on Information Systems,2016,35(3):16.
- [24] MCCREADIE R M C,MACDONALD C,OUNIS I. Crowdsourcing a news query classification dataset[C]//Proceedings of the SIGIR 2010 Workshop on Crowdsourcing for Search Evaluation.[S.1.]:[s.n.],2010:31-38.

作者简介

张娟娟,女,1985年生,博士,副教授,研究方向:信息检索,E-mail:zxj0614@swu.edu.cn。

Comparative Analysis of the Personalized Potential of Informational, Navigational and Transactional Queries

ZHANG XiaoJuan

(School of Computer and Information Science at Southwest University, Chongqing 400715, China)

Abstract: The personalized potential of informational, navigational and transactional queries is firstly measured by explicit and implicit measures respectively. And then, the effective implicit measures are obtained by analyzing the correlation between implicit measures and explicit measures. On light of these, for each category of query intent, this paper also analyzes which query feature can effectively characterize their personalized potential. Finally, some suggestions about the performance optimization of search engines are provided according to the experimental results.

Keywords: Informational Queries; Navigational Queries; Transactional Queries; Personalized Potential

(收稿日期: 2017-08-08)