作者简介及博士学位论文中英文摘要

论文题目：搜索引擎点击模型构建

作者简介：王超，男，1989 年 2 月出生，2011 年 9 月师从于清华大学马少平教授，于 2016 年 7 月获博士学位。

中 文 摘 要

搜索引擎用户在与搜索引擎的交互过程中反映出的隐性反馈信息（主要是点击行为信息）是搜索引擎用来改进结果排序的重要影响因素。然而，由于结果位置、展现形式等各种因素的影响，将反馈信息直接应用于搜索排序任务往往难以取得较好的效果。针对这一问题，研究人员提出了构建描述用户点击行为的点击模型（Click Model），并基于不同的点击模型估计用户对展现结果的浏览概率，进而尝试去除结果展现位置等因素对用户行为的偏置性影响，以达到更好利用隐性反馈信息的目的。

作为一种用户交互信息的有效利用方法，点击模型在学术界得到了充分关注，并在工业界得到了广泛的应用。然而，已有的点击模型主要针对于传统同质化的搜索页面（搜索页面中的结果均采用相近的文本形式展现，结果之间除了文字内容不同外并没有明显的展现形式差异）进行设计。随着 Web2.0 时代的到来，富媒体展现形式被越来越多的应用于搜索交互界面，搜索结果也变得越来越异质化，这些变化使得用户的检验行为（注意力分布偏好，浏览顺序等）发生了明显的改变，已有的点击模型已经不能正确的描述用户的真实行为，相应的排序方法也难以取得较优的性能。针对这一问题，本文主要研究异质环境下搜索引擎用户的浏览行为变化，提出有效的新型点击模型描述用户行为，进而改进搜索引擎排序效果。

具体的，本文首先基于大规模真实搜索行为数据和实验室环境下的用户认知行为实验，针对搜索引擎用户检验认知行为开展了实证分析；进而基于上述分析结果开展了三项逐渐深化的搜索引擎用户点击模型构建工作，包括：具有异质描述能力的点击模型构建，具有非顺序行为描述能力的点击模型构建，以及融合内容与行为特征的点击模型构建。

在搜索引擎用户检验认知研究方面，传统的工作大多基于上世纪 80 年代认知心理学研究提出的“心眼一致”的理论假设，将眼动追踪中用户对于搜索结果的视觉注视（Fixation）行为等同于用户对于搜索结果的认知理解（Examination）行为。然而，近年的研究成果表明，用户的视觉注视行为和认知理解过程并非同时完成，因此传统的工作无法解释上述差异。针对上述问题，我们利用眼动视线追踪仪器设计用户交互实验证实：在搜索技术研究中广泛使用的“心眼一致”假设事实上并不适用于真实网络搜索环境。用户真实的检验行为是以“略读—精读”两阶段过程为主要特征的，这也首次证实了认知心理学研究中普遍认可的“选择性注意理论”同样适用于网络信息获取的应用场景。基于以上发现和对眼动注视数据的进一步分析，我们建立了用户两阶段检验行为模型，并提出了采用收集成本更加低廉的鼠标移动行为代替眼动注视信息进行检验概率预测的学习方法。实验显示使用鼠标移动行为特征能够较为准确的预测眼动注视信息，这为大规模应用提供了可行方案。

在具有异质描述能力的点击模型构建研究方面，传统的工作假设所有搜索结果具有相同的展现形式，因此无法正确描述不同展现形式对用户行为产生的影响。针对上述问题，我们利用眼动视线追踪仪器分析了不同展现形式的垂直搜索结果（传统结果，图片类结果，新闻类结果，应用类结果）对用户的视觉注视和点击偏好产生的影响，并构建了相应的点击模型

（VCM）。实验显示 VCM 能够比已有的模型更加准确的描述和预测用户的点击行为，并且通过 VCM 获得的不同搜索结果的相关性推断也更为准确，可以有效的提升搜索引擎的排序效果。

在具有非顺序行为描述能力的点击模型构建研究方面，传统的工作假设用户只会对搜索结果列表进行自上而下检验一遍的浏览过程，因此无法正确描述不满足上述检验（或点击）顺序的行为。针对上述问题，我们利用眼动视线追踪仪器对用户复杂的检验及点击顺序进行了深入研究，突破了已有模型只能针对用户自上而下检验一遍的假设限制，构建了具有更强描述能力的点击模型（PSCM）。实验显示上述模型能够比传统模型更准确的预测用户的点击行为。最后，在融合内容与行为特征的点击模型构建研究方面，传统的点击模型基于概率图模型构建，当越来越多的复杂用户行为规则以及内容特征加入后，图模型会变得过于复杂而难以计算，因此难以有效引入内容特征。针对上述问题，我们提出了基于卷积神经网络的新型点击模型框架。相比于传统点击模型框架只利用用户点击行为信息，该框架能够进一步利用大量的文本内容信息和上下文信息，同时利用神经网络分层递进处理的特性保证了用户浏览行为信息不会被大量的文本内容信息所“淹没”。该框架能够应用于大多数已有的点击模型，我们通过实验显示了该框架能够使得传统点击模型在用户点击预测和搜索结果相关性预测方面得到显著的改进。

在研究成果应用方面，本文提出的部分点击模型已在某中文商业搜索引擎系统中得到了真实应用，有效的提高了该搜索引擎搜索排序性能。相关工作也作为“群体智能支撑的互联网搜索技术及其应用”项目的一部分获得了 2015 年北京市科技进步一等奖。

关键词：点击模型；用户行为分析；眼动追踪；垂直搜索；卷积神经网络

A Study of Search Engine Click Model

WANG Chao

ABSTRACT

Modern search engines record user interactions and use them to improve search quality. In particular, user’s click-through has been successfully used to improve click-through rate (CTR), Web search ranking, query recommendation and suggestion, and soon.

Although click-through logs can provide implicit feedback of users’ click preferences, it is difficult to derive accurate absolute relevance judgments due to the existence of click noises and behavior biases. Previous studies showed that users’ clicking behaviors are biased towards many aspects such as “position” (user’s attention decreases from top to bottom), “trust” (Web site reputations will affect user’s judgment) and so on. To address these problems, researchers have proposed a number of click models to describe user’s practical browsing behavior and to obtain an unbiased estimation of result relevance.

While these existing click models have achieved much success in click/relevance prediction for Web search and sponsored search, the limitations of these models have been more and more serious. With SERPs being more and more complex, an increasing number of search result pages (SERPs) are federated from multiple specialized search engines (called verticals, such as Image or Video), search user behaviors also become more difficult to be described by simple behavior

assumptions (e.g. cascade assumption).According to our experiments, these models are not able to describe user's actual behavior for modern search environment. Therefore, we should make further analysis of user's new search behaviors and then improve click models to describe such new behaviors.

In this paper, we first try to analyze user's cognitive process in modern search environment. Then, we propose three new click models to improve existing click models in the following aspects: modeling heterogeneous SERP, modeling user's non-sequential search behavior, and combining user behavior information with search result content information.

As for the serach engine user's cognitive process research, we design an experimental search engine to collect both the user's feedback on their examinations and the eye-tracking/click-through data. To our surprise, a large proportion (45.8%) of the results fixated by users are not recognized as being "read". Looking into the tracking data, we found that before the user actually "reads" the result, there is often a "skimming" step in which the user quickly looks at the result without reading it. We thus propose a two-stage examination model which composes of a first "from skimming to reading "stage (Stage 1) and a second "from reading to clicking" stage (Stage 2). We found that the biases (e.g. position bias, domain bias, attractiveness bias) considered in many studies impact in different ways in Stage 1 and Stage 2, which suggests that users make judgments according to different signals in different stages. We also show that the two-stage examination behaviors can be predicted with mouse movement behavior, which can-be collected at large scale. Relevance estimation with the two-stage examination model also outperforms that with a single-stage examination model. This study shows that the user's examination of search results is a complex cognitive process that needs to be investigated in greater depth and this may have a significant impact on Web search.

As for the heterogeneous SERP research, we collect a large scale log data set which contains behavior information on both vertical and ordinary results. We also perform eye-tracking analysis to study user's real-world examining behavior. According these analysis, we find that different result appearances may cause different behavior biases both for vertical results (local effect) and for the whole result lists (global effect). These biases include: examine bias for vertical results (especially those with multimedia components), trust bias for result lists with vertical results, and a higher probability of result revisitation for vertical results. Based on these findings, a novel click model considering these biases besides position bias is constructed to describe interaction with SERPs containing verticals. Experimental results show that the new Vertical-aware Click Model (VCM) is better at interpreting user click behavior on federated searches in terms of both log-likelihood and perplexity than existing models.

As for the user's non-sequential search behavior research, we investigate the problem of properly incorporating non-sequential behavior into click models. We firstly carry out a laboratory eye-tracking study to analyze user's non-sequential examination behavior and then propose a novel click model named Partially Sequential Click Model (PSCM) that captures the practical behavior of users. We compare PSCM with a number of existing click models using two real-world search engine logs. Experimental results show that PSCM outperforms other click models in terms of both predicting click behavior (perplex-ity) and estimating result relevance (NDCG and user preference test). We also publicize the implementations of PSCM and related datasets for possible future comparison studies.

As for the combination of user behavior information and search result content in- formation research, we propose a novel click model framework based on convolutional-neural network architecture to make it more suitable for the ever-changing complex search environment. Compared with traditional probabilistic graphic models, our proposed framework not only uses user behavior

information as input signals, but also adopts the result content information and take the relationships among different search results (the context information of results) into consideration. It properly processes content, context and user behavior information in different neural network layers to make sure that high-level user behavior features will not be “buried” by large amount of content/context features. The proposed model also adopts parameters from existing click models as constraints for variables in the hidden layer, which guarantees the effective estimation of the examination probability and the user perceived relevance parameters. The framework can be adopted to reconstruct most existing click models and experimental results based on large scale practical user behavior data show promising results. State-of-the-art click models such as UBM and PSCM can gain significant improvement after reconstruction with the framework in terms of both click perplexity and NDCG.

Parts of our research (VCM and PSCM) have been implemented in a Chinese com-mercial search engine and show promising improvement in this practical system.

Key words: Click Model; User Behavior Analysis; Eye-tracking; Vertical Search; Conventional Neural Network