

Topical Crawler Technology Based on Reinforcement Learning

WANG Youzeng¹, WANG Jinbao²

1. Basic Science Department, Shandong Water Polytechnic, Rizhao, China

2. Rizhao senior Technical school, Rizhao, China,

e-mail: wangyouzeng123@126.com, wangyouzeng@sohu.com

Abstract: With the rapid development of Internet, Web mining technology has been an hotspot. How to effectively control the crawling strategy is the most important successful factor for Web mining. In this paper, we advance a new crawler's crawling strategy combining Reinforcement learning. The experimental results show that this strategy can increase the harvest rate of inquiries.

Keywords: reinforcement learning; topical crawler; q-learning; crawling strategy

基于强化学习的主题爬虫技术

王有增¹, 王金宝²

1. 山东水利职业学院基础科学部, 日照, 中国, 276826

2. 日照高级技工学校, 日照, 中国, 276826

e-mail: wangyouzeng123@126.com, wangyouzeng@sohu.com

【摘要】随着网络的飞速发展, Web 挖掘技术已成为一个研究热点, 而如何有效地控制爬虫爬行策略是影响 Web 挖掘是否成功的最重要因素之一。本文结合强化学习理论, 提出一种新的主题爬虫爬行策略。实验证明该策略能有效的提高查询的收获率。

【关键字】强化学习; 主题爬虫; Q 学习; 爬行策略

1 引言

随着数字化信息时代的到来,网络日渐成为人们获得信息的重要途径。然而网络中信息量巨大且分散无序, Web 用户经常发现难以找到其所需的信息, 造成“信息过载, 知识匮乏”[1]的现状。通用搜索引擎给人们提供了信息检索的方法, 但也存在查准率不高、查全率不能保证等问题。Web 挖掘[2]技术正是应这一需求而出现的一种新技术。

Web 爬虫[3]是 Web 挖掘的核心技术,它是一类可以通过下载 Web 页面、分析页面内容、并跟随其中的链接来递归地遍历 Web 的计算机程序。主题 Web 爬虫比普通的 Web 爬虫的爬行效率更高,它以智能的方式在线爬行 Web 页面,收集更多主题相关页面。如何有效地控制主题爬虫爬行策略是影响 Web 挖掘是否成功的最重要因素之一。本文结合强化学习理论,提出一种新的主题爬虫爬行策略。实验证明该策略能有效的提高查询的收获率。

2 主题爬虫技术

主题爬虫[4,5] (Topical Crawler), 又称为聚焦爬虫 (Focused Crawler) 或主题驱动的爬虫 (Topic-driven Crawler)。它的爬行过程是目标主题驱动的,能够根据用户或信息检索定义的目标主题, 以智能的方式在线爬行 Web 页面,判断页面是否主题相关并进行相应排序,使最有可能是主题相关的页面得以优先爬行。

主题爬虫主要有三大模块构成: 网页爬行器、超文本分类器和超链接评价器。网页爬行器可从 URL 队列中逐个取出最高爬行优先级的 URL, 然后通过 HTTP 协议下载相应的 Web 页面。超文本分类器通过对目标主题信息的学习而获得分类模型, 然后用分类模型对已爬行的页面进行在线分类。超链接评价器先从页面中抽取出候选 URL 并插入到 URL 队列中, 然后通过算法对候选 URL 进行综合评价, 并计算出爬行优先级, 最后根据爬行优先级对 URL 队列进行排序。

主题爬虫能在有限的时间和空间资源条件下, 收

集到尽可能多的主题相关页面，而同时又避免不相关的页面被爬行。但由于主题爬虫缺乏增量学习能力，初始样本页面就成为决定超文本分类器和超链接评价器性能的主要因素，这一问题制约了主题爬虫的发展。本文结合强化学习理论，提出了一种新的主题爬虫爬行策略。

3 强化学习理论

强化学习[6,7]是一种不同于监督学习和非监督学习方法的在线学习理论。它的基本原理是：如果智能体的某个行为策略导致环境对智能体正的奖(reward)，则智能体以后采取这个行为策略的趋势就会加强。

强化学习把学习看作是一个“试探——评价”的过程，首先学习系统感知环境状态，采取某一个动作作用于环境，环境接受该动作后状态发生变化，同时给出一个回报（奖励或惩罚）反馈给强化学习系统，强化学习系统根据强化信号和环境的当前状态再选择下一个动作，选择的原则是使受到奖励的概率增大。选择的动作不仅影响立即强化值，而且影响环境下一刻的状态及最终的强化值。

4 基于强化学习的主题爬虫技术

当我们把 Web 页面看作是不同的状态，遍历一个超链接看作是执行一个动作，把得到主题相关的页面作为获得立即收益，如此爬虫的爬行问题就可转变为学习的过程。但要把强化学习直接应用到 Web 环境下的主题爬虫时，有两个问题必须解决：巨大的状态空间和行动空间。研究人员对基于强化学习的主题爬虫提出了两个假设：（1）假设当前状态与已经爬行到的主题爬虫相关页面无关；（2）不同动作之间的差异可以通过超链接的上下文来区分。根据这两个假设，我们可以忽略状态，并且把 Q 值公式看作从超链接上下文到延迟收益之和的一个映射。

4.1 在线 Q 值计算

我们首先使用通用搜索引擎（如 google 等）搜索相关主题，把搜索的前几项作为样本集。通用搜索引擎搜索结果中的前几项通常都有着较高的主题相似度，我们把它作为初始 URL 来爬行网络。爬行过程采用带折扣收益的 Q 学习方法，各个超链接的 Q 值信息是基于已爬行页面的信息的，并随着爬行过程不断的更新，是一种信息反馈的过程。候选 URL 的 Q 值计算公式如下：

$$Q(u_i) = R(u_i) + \gamma Q(d_i); \quad Q(d_i) = \frac{1}{n} \sum_{j=1, u_j \in d_i}^n Q(u_j)$$

其中， u_i 是已爬行过的超链接， d_i 是爬行超链接 u_i 所获得的页面， γ 是折扣因子。 $R(u_i)$ 是爬行超链接 u_i 所得到的立即收益， $\gamma Q(d_i)$ 表示折算了的长远回报， γ 决定了长远回报相对立即回报的重要性。 u_j 是 d_i 的子 URL，并且是一个已经爬行过的 URL， n 表示 d_i 的已爬行子 URL 的总数。

Q 值公式中的立即收益，即爬行一个超链接所获得的回报，我们采用样本评估和向量相似度相结合的方法。公式如下：

$$R(u_i) = S_a(u_i) + \lambda_1 (S^+(u_i))^{n+}$$

其中， $S_a(u_i)$ 是超链接锚文本与主题关键词的余弦相似度。 $(S^+(u_i))^{n+}$ 是 URL 的样本集相似度得分， λ_1 是折扣因子。

超链接锚文本[8]是超链接的载体，能够很好的反映链接的语义信息。因此用查询条件与锚文本相匹配可以取得好的效果。我们采用向量空间模型（VSM），把用户提交的关键词和锚文本用向量的形式表示出来。 tf_k 分别为主题关键词 t_k 在超链接文本中出现的次数， L 是超链接文本的长度， m 为特征值的个数。公式如下：

$$w_n = \sum_{k=1}^m \frac{tf_k}{L}; \quad S_a(u_i) = \sum_{n=1}^l w_n$$

其中， $S_a(u_i)$ 为网页 i 的相关度值，1 为匹配的超链接条数。

同理，我们采用余弦公式，计算出爬行页面向量与样本集的相似度得分。

$$(S^+(u_i))^{n+} = \frac{\sum tf_m w_m^+}{\sqrt{\sum (tf_m)^2 * \sum (w_m^+)^2}}$$

4.2 超链接上下文到 Q 值的映射

我们采用朴素贝叶斯文本分类器来建立相应的映射关系，把训练数据的收益分离到箱中。同时采用向量空间表示法，把每个特征词用向量地形式表示出来。为了计算方便，把 Q 值归一化在 [0, 1] 区间内，然后将该区间分为 k 个子区间，分别为 $C_1, C_2, \dots, C_i, \dots$ 。我们把每个子区间看作一个类，该类代表 Q 值的平均值。在 Q 值映射问题中，对于每个候选 URL，分类器都要根据该 URL 的上下文内容来判断它属于哪一个

类，并把与该类相对应的 Q 值赋给该候选 URL 以表示他的爬行优先级。由于贝叶斯分类器是一个基于统计特征独立假设的概率分类器，我们假设任何一个特征词在页面中出现的概率与其它特征是否出现没有内在的任何关系。在这种条件下，我们可以利用朴素贝叶斯分类方法来解决这个问题，

公式如下：

$$C_j = \max P(C_j | d_i) = \max P(C_j) P(d_i | C_j) = P(C_j) \prod_{k=1}^{|d_i|} P(w_{ij} | C_j)$$

其中， $|d_i|$ 是文档 d_i 的个数。

根据公式可知，我们把问题转化为求先验概率和后验概率。计算先验概率比较简单，它等于类 C 包含的样本数量占整个样本集的百分比。后验概率的计算可由下来公式得出。

$$P(w_{ij} | C_j) = \frac{1 + \sum_{k=1}^{|C_j|} w_{ik}}{|v| + \sum_{k=1}^{|C_j|} w_k}$$

$$P(C_j) = \frac{N_j}{N_d}$$

其中 $\sum_{k=1}^{|C_j|} w_{ik}$ 是出现在所有类中特征词 w_{ik} 的权重和， $\sum_{k=1}^{|C_j|} w_k$ 是所有类中特征词的权重和。 $|v|$ 是特征词的数量，而 $|C_j|$ 则是类的数量。 N_j 是属于类 C_j 的文档数量， N_d 是所有训练文本的数量。

4.3 基于强化学习的主题爬虫爬行策略

基于强化学习的主题爬虫爬行策略的核心思想就是在主题爬虫爬行过程中加入强化学习理论，使其具备学习能力，从而收集更多、更准确的主题相关页面。主要步骤如下：

- (1) 用通用搜索引擎对用户提交的关键词进行查询，把返回页面的前几项作为种子页面并加入样本集。
- (2) 抽取出页面中的超链接，计算出内容相关度和样本集相关度，按计算的结果进行排序。
- (3) 爬虫按超链接评价价值进行收集页面，对收集到的页面进行页面分析，抽取出主题相关页面中的 URL 进行下一部的爬行。
- (4) 如果收集到的页面满足用户设定的最大下载量，则停止下载；如果还没到，且还有未爬行过的 URL，则返回步骤 (2)，继续爬行。

5 实验

主题爬虫系统最重要的性能指标是收获率[9]和召回率[9]。

我们设计了一组实验，设定查询关键词为“职业学校”，比较宽度优先策略的爬虫、采用深度优先策略的爬虫和基于强化学习的主题爬虫的各自爬行结果。实验结果如下：

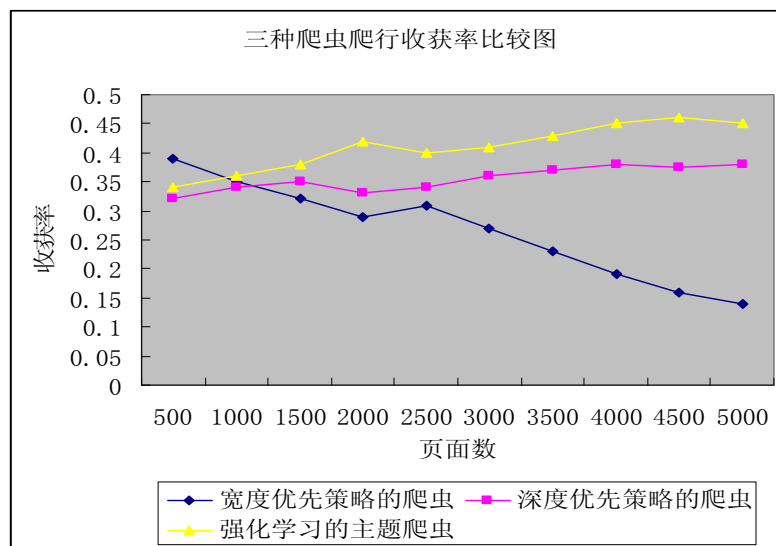


Figure1. Comparative chart on three crawlers' crawling harvest rate
图 1. 三种爬虫爬行收获率比较

实验结果表明，基于强化学习的主题爬虫的收获率虽然最初没有其它两种爬虫的收获率高，但随着爬

行的进一步深入，其收获率则明显高于其它两种爬虫的收获率。

6 总结

随着网络信息资源的急剧增长，人们面对这个海量的信息库常常找不到自己所需的信息。通用搜索引擎给人们提供了信息检索的方法，但也存在查准率不高、查全率不能保证等问题。Web 挖掘为解决这一问题提供了新的研究思路。

Web 爬虫是 Web 挖掘的核心技术，而如何有效地控制爬虫爬行策略是影响 Web 挖掘是否成功的最重要因素之一。本文结合强化学习理论，提出一种新的主题爬虫爬行策略。实验证明该策略能有效的提高查询的收获率。下一步，我们将加入更多机器学习理论，进一步加强主题爬虫的学习能力，使其能在更短的时间内挖掘出更多主题相关的页面。

致谢

本文在完成的过程中得到了崔维群教授的悉心帮助，崔老师在工作繁忙之中，给我提供了许多有参考价值的资料，提出不少建设性意见，在此表示诚挚的感谢。

References (参考文献)

[1] Raymond Kosala, and Hendrik Blockeel. Web Mining Research:

A Survey. SKGKDD Explorations, July 2000.

- [2] Han Jianwei, Meng Xiaofeng, Wang Jing, etc. Web Mining Research. Computer Research and Development. 2001, 38(4): 405-410.
- [3] 韩家炜, 孟小峰, 王静等. Web 挖掘研究. 计算机研究与发展, 2001, 38(4): 405-410.
- [4] J. Cho. Crawling the web: Discovery and Maintenance of large-scale Web Data. PHD Thesis. 2001.
- [5] Li Xueyong, Ouyang Liubo, Li Guohui, Zhong minjuan. Comparative Study on Search Strategies of Web Spider. Computer. Engineering and Applications. 2004: 128-130.
- [6] 李学勇, 欧阳柳波, 李国徽, 钟敏娟. 网络蜘蛛搜索策略比较研究. 计算机工程与应用, 2004: 128-130.
- [7] A. Rungsawang, N. Angkawattanawit. Learnable topic-specific web crawler. Computer Applications, 28 (2005) 97-114.
- [8] Richard S. Sutton, Andrew G. Barto. Reinforcement learning: An Introduction[M]. MIT Press, Cambridge, MA, 1998.
- [9] Gao Yang, Chen Shifu, Lu Xin. An Overview of Reinforcement Learning[J]. Journal of Automatic Chemical. 2004, 30(1): 86-101.
- [10] 高阳, 陈世福, 陆鑫. 强化学习研究综述 [J]. 自动化学报. 2004, 30(1): 86-101.
- [11] Zhang Min, Gao Jianfeng, Ma Shaoping. Web Information Search Based on Linked Descriptive Text and Its Context [J]. Journal of Computer Research and Development. 2004; (1): 221-226.
- [12] 张敏, 高剑峰, 马少平. 基于链接描述文本及其上下文的 Web 信息检索[J]. 计算机研究与发展, 2004; (1): 221-226.
- [13] M. M. Sufyan Beg*. A subjective measure of web search quality. Information Sciences, 2005(169): 365-381.