

互联网信息搜索用户行为模型的探索性研究

朱明泉¹, 张智君^{1*}, 任衍具²

(1. 浙江大学 心理与行为科学系, 浙江 杭州 310028; 2. 中国科学院 心理研究所, 北京 100101)

摘要:探讨互联网信息搜索中用户与网络交互的行为模型. 要求 32 名大学生或研究生被试完成不同类型(开放型与封闭型问题)的信息搜索任务, 记录整个搜索过程, 并提取各类关键事件的发生频次和时间等数据进行分析. 基于路径分析获得的用户搜索行为模型显示: 任务类型因素通过“主题目录链接”与“相关链接”的相关对搜索时间产生非因果效应; 个体因素通过“页面跳转”对搜索时间产生间接因果效应, 通过“相关链接”对绩效产生间接因果效应; 环境因素通过“浏览时间”和“提交请求”对搜索时间产生间接因果效应.

关键词:互联网; 信息搜索; 行为模型; 路径分析

中图分类号: B849; TB18

文献标识码: A

文章编号: 1008-9497(2006)04-475-06

ZHU Ming-quan¹, ZHANG Zhi-jun¹, REN Yan-ju² (1. Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou 310028, China; 2. Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China)

Exploration of the behavioral model of user-web interaction. Journal of Zhejiang University(Science Edition), 2006, 33(4): 475~480

Abstract: The behavioral model of user-web interaction in the process of web information search was examined. Thirty-two university students were asked to search for information related to two open and two closed questions. The search activities were recorded with HYPERCAM and coded as key events, whose frequencies were analyzed later. The user searching behavior model based on path analysis implied that question type had indirect effect on search time through the relation between 'subject directory' and 'related link', user characteristics had indirect effect on search time primly by 'jump', and had indirect effect on performance by 'related link', and environment had indirect effect on search time by browsing time and 'submit request'.

Key words: Internet; information search; behavior model; path analysis

互联网信息搜索已成为最普及的网络活动之一^[1]. 但是, 由于网络信息的大量涌现, 网络环境变得日益复杂, 信息搜索效率因而受到较大影响^[2]. 对此, 许多研究者正努力探索网络信息搜索的特征和规律, 希望设计出更好的搜索工具和更有效的信息组织方式. 该领域的研究目前主要集中在网络搜索行为的影响因素、用户搜索策略以及信息搜索过程中用户与网络的交互模型等方面.

对搜索行为影响因素的研究分析表明: 用户特征、搜索任务类型、系统能力和搜索结果反馈等均对用户的搜索行为和搜索绩效存在重要影响. 其中, 用

户特征涉及用户的背景知识、计算机使用经验、信息需求、在特定领域拥有的知识、认知能力、情感状态和人口统计参数等方面^[2]. 而对任务类型的研究则显示, 搜索问题的开放性对用户的搜索行为有显著影响^[3].

网络搜索策略的研究主要通过行为观测来进行. 众多研究者从不同角度出发对用户的行为策略进行分析. HAWK 和 WANG(2000) 以直接输入 URL、应用搜索引擎和应用超链接等来区分用户的搜索策略^[4]. WHITE 和 IIVONEN(2001) 将搜索策略分为直接地址、主题目录和搜索引擎, 并发现在最

收稿日期: 2005-04-11.

基金项目: 全国教育科学规划资助项目(EBA030410); 浙江省自然科学基金资助项目(301076).

作者简介: 朱明泉(1984-), 女, 硕士研究生, 主要从事工程心理学、认知心理学研究.

* 通讯作者: 张智君, 男, 教授, 主要从事工程心理学、认知心理学研究. E-mail: zjzhang@zju.edu.cn.

初选择策略时使用搜索引擎最多(43%)^[5]. Navarro-Prieto等(1999)提出了自上而下、自下而上和混合模型等3种搜索策略^[6].此外, HAWK和WANG(2000)以纵览、重复检查、探索、跟随链接、后退和前进、寻求捷径、引擎使用、忠实的引擎使用、引擎搜寻和元搜索等作为依据去分析用户的搜索策略^[4]. KIM(2001)则通过嵌入式链接、后退键、主页键、跳转选项和关键字搜索等行为去考察不同导航方式的影响效应^[3].由于搜索引擎是网络搜索中最常用的工具之一,相关研究也较多.与使用搜索引擎有关的研究主要集中在搜索请求方面^[7-10].

一些研究者在充分考虑网络搜索行为影响因素和用户搜索策略的基础上,尝试建立可揭示用户与网络交互过程的行为模型.例如, SPINK和SARACEVIC(1998)提出了信息提取的分层交互模型,该模型包含用户和计算机两个成分,用户领域包含的要素有情境、情感、认知和请求,计算机领域包含的要素有界面、工程、加工和容量^[11]. WANG等(2000)提出了一个多维模型,该模型包含用户、界面和网络空间等3个在搜索过程中相互作用的成分,并认为可将用户与网络的交互活动看作是由界面促进的交流过程^[12].上述两个模型的差异在于,后者将界面看成一个独立的维度,并将各个维度的元素细化. HODKINSON等(2000)则用流程图的方式概括了用户的网络搜索行为^[13],清晰地描述了用户的一般搜索过程.

所有这些模型囊括了影响网络搜索绩效的主要因素,并强调用户与系统交互的重要性.但是,它们均没有考察这些因素影响搜索行为进而影响搜索绩效的方式,也无法确定这些因素在整个交互模型中的权重,不利于相关研究者根据模型提出相应的设计建议.此外,以往对中文网络搜索行为的研究多限于在自编的有限网页上进行,与真实网络环境存在差别.因此,本研究在真实的网络环境下,以搜索行为作为中介变量,探索用户与网络交互的行为模型.即借助行为观察技术,综合定性和定量数据来分析用户的搜索行为,并由此考察各种重要因素在网络信息交互模型内的作用机制,建立网络信息搜索行为的一般模型.

1 研究方法

1.1 实验方法

采用实验室研究的方法.具体步骤为:(1)利用视频软件(Hypercam)实时记录被试的搜索行为;

(2)利用Window Media Player软件回放记录到的视频资料,提取关键事件并对行为进行编码;(3)对编码信息进行统计处理,建立用户信息搜索的行为模型.

先选取男女各5名大学生进行预备实验,根据预备实验的资料和以往研究的成果确定将用户行为区分为“直接输入网址”、“使用引擎”、“提交请求”、“主题目录链接”、“相关链接”、“应用搜索结果”、“页面跳转”、“浏览”、“使用导航”以及“其它”等10项关键行为事件,以进行编码和分析.

1.2 搜索任务

共设计了4项搜索问题,其中2项为开放型问题(或称主题性问题),另2项为封闭型问题(或称事实性问题).开放型问题无特定答案,而封闭型问题则有特定答案.这些问题所涉及的内容可忽略被试专业和背景知识的影响.

1.3 实验程序和被试

要求被试借助IE浏览器以自然的搜索方式尽快而准确地依次找到各个问题的答案,并将答案拷贝到Word文档中保存.搜索和回答每个问题的时间限制为15分钟.在实验开始前,清空IE浏览器的历史记录和收藏夹,以避免受前一位搜索者行为的影响.

32位具有类似网络搜索经验的本科生或研究生(男24名,女8名)参与正式实验.每位被试均需完成2个类型共4项问题.实验采用拉丁方设计以平衡所搜索问题的顺序效应.

2 结果

对视频文件进行回放,记录各个关键事件的发生频次和搜索时间(不包括保存结果和等待页面下载的时间),作为行为分析的指标.此外,研究者需根据被试所搜索到的信息的内容对答案的准确性和圆满性作出主观评价(以百分数表示),该结果称为“绩效评价分数”.以搜索时间除相应的绩效评价分数,得到“综合绩效”.结果如表1左侧所示.

2.1 搜索行为分析

2.1.1 搜索行为的一般分析

从表1可以发现,被试选择“直接输入网址”搜索信息的频率很少,所有被试最初都选择使用搜索引擎.同时,在每个问题上选择引擎的次数是1.42次,即被试很少更换引擎.这些结果与SPINK和OZMUL TU(2002)^[7]以及OZMUTLU等(2004)^[9]的研究结果一致.

表1 不同任务类型和绩效水平下的关键事件和绩效平均值

Table 1 Average frequencies of key events and average scores of performance in different situations

关键事件或绩效类别	总平均值	封闭型问题		开放型问题			
		高绩效	低绩效	高绩效	低绩效		
关键事件	直接输入网址(次)	0.17	0.094	0.344	0.063	0.094	
	使用引擎(次)	1.42	1.500	1.219	1.188	1.434	
	提交请求(次)	3.09	3.563	3.000	2.906	2.906	
	主题目录链接(次)	2.05	3.156	3.375	0.656	0.781	
	相关链接(次)	3.87	2.500	5.750	2.438	4.656	
	应用搜索结果(次)	5.54	4.531	5.969	5.281	6.312	
	页面跳转(次)	16.52	9.125	27.125	9.000	22.031	
	浏览	时间(s)	344.66	313.938	394.781	275.063	394.88
		次数(次)	22.72	20.625	30.625	17.219	24.250
		使用导航(次)	1.45	1.500	1.781	0.719	0.719
其它	1.6	0.938	1.500	1.031	1.625		
绩效	搜索时间(s)	491.80	439.969	555.406	423.375	548.438	
	绩效评价分数	78.79	84.143	65.394	86.980	73.646	
	综合绩效	0.21	0.280	0.133	0.274	0.161	

使用搜索引擎搜索信息时,搜索者对关键词的选择在很大程度上影响到信息搜索的效率.在遇到搜索障碍时,被试总是首先更换搜索关键词,只有在多次更换关键词仍不能获得所需信息时才会考虑更换引擎并重复搜索过程(每个引擎上提交请求的平均次数是2.17次).

在系统返回搜索结果后,可采取“一次只打开一个搜索结果”或“同时打开多个搜索结果”两种不同的策略.同时打开多个页面,在页面跳转时容易出错,而一次只打开一个结果,可能会重复打开相同的内容.从表1“页面跳转”(支持同时打开多个页面)和“使用导航”(支持一次打开一个结果)的数据可以发现,大多数被试在点击链接时倾向于将结果打开在新的窗口内,而不是在同一个窗口将原页面内容替换.

当页面内容较少时,被试只须快速浏览一次即可了解是否包含所需信息.如果页面内容较多且有些与主题有关,则可能须多次重复浏览以确定是否包含所需的信息.另外,被试往往没有耐心等待页面下载.在网速较慢的情况下,如果可能,会同时搜索多个问题,同时打开多个结果页面.在等待下载的过程中,被试经常在页面间进行快速的循环跳转,似乎通过这种方式可提高搜索效率.因此,网络系统特点会影响浏览时间和页面跳转频次.

2.1.2 不同任务类型和不同被试类别对搜索行为的影响

根据综合绩效分数将被试均分为“高绩效”和“低绩效”两个组别.高绩效组的平均综合绩效分数为0.28,低绩效组为0.15,两者差异非常显著($t =$

$5.927, p < 0.01$).不同任务类型和不同绩效水平下各个关键事件和绩效指标的平均值如表1右侧所示.

对各个关键事件指标进行2(任务类型)×2(综合绩效)的方差分析,结果发现“使用引擎”的交互作用显著($F(1, 30) = 4.392, p < 0.05$),即在搜索封闭型问题时高绩效组使用引擎的次数较多,而在搜索开放型问题时低绩效组使用引擎的次数较多.任务类型主效应显著的有“主题目录链接”和“使用导航”.具体为:搜索封闭型问题时使用主题目录链接的次数高于开放型问题($F(1, 30) = 16.693, p < 0.01$),对封闭型问题使用导航的次数多于开放型问题($F(1, 30) = 4.678, p < 0.05$).作者推测,封闭型问题有确定的答案,搜索者在一个相关网站内使用主题目录的可能性更大,而这种链接往往诱发在当前窗口打开新页面的行为,因此搜索者须较多地利用导航键来完成搜索.绩效主效应显著的有“相关链接”、“页面跳转”、“浏览次数”和“浏览时间”.具体为:低绩效组使用相关链接的次数多于高绩效组($F(1, 30) = 11.356, p < 0.01$),低绩效组页面跳转的频次高于高绩效组($F(1, 30) = 18.046, p < 0.01$),低绩效组的浏览次数高于高绩效组($F(1, 30) = 7.602, p < 0.01$),低绩效组的浏览时间多于高绩效组($F(1, 30) = 7.989, p < 0.01$).

根据以上分析,笔者认为:任务类型影响“主题目录链接”,通过“主题目录链接”间接影响“使用导航”,并最终影响到搜索效率;使用引擎的次数对“提交请求”的次数产生影响,并由此影响到搜索时间;影响绩效评价分数的只有“相关链接”,而“页面跳

转”、“浏览次数”和“浏览时间”通过影响搜索时间对搜索绩效产生效应。

2.2 网络信息搜索的行为模型

根据前面的综述,搜索行为和绩效主要受个体、任务类型和环境等因素的影响.这三个因素均通过用户行为而对最终的搜索绩效产生作用,即本研究所涉及的关键事件是影响搜索绩效的中介变量。

因为本研究没有将个体因素和环境因素作为显变量,所以首先仅考虑任务类型对搜索行为的影响.其中,因为“页面跳转”与“浏览次数”存在高相关,所以在模型中去掉“浏览次数”;而前期的探索性分析显示,“应用搜索结果”和“其它”对搜索时间和绩效无显著影响,因此也从模型中去除.假设:相关链接是网络信息组织方式的一大特点,对其利用的程度可影响使用者的搜索绩效,而网页之间的切换频率也会影响搜索时间,浏览则是搜索过程中最主要的认知活动,浏览所用时间必然会影响到搜索时间.根据这一假设所确定的网络信息搜索的行为模型如图 1 所示(虚框以外的部分).具体为:“页面跳转”和“提交请求”的次数影响“浏览时间”,并同时直接对“搜索时间”产生影响;“相关链接”对“绩效评价”产生直接影响,“主题目录链接”影响“使用导航”,“使用引擎”影响“提交请求”.对此模型进行拟合,得到拟合指数如表 2 所示。

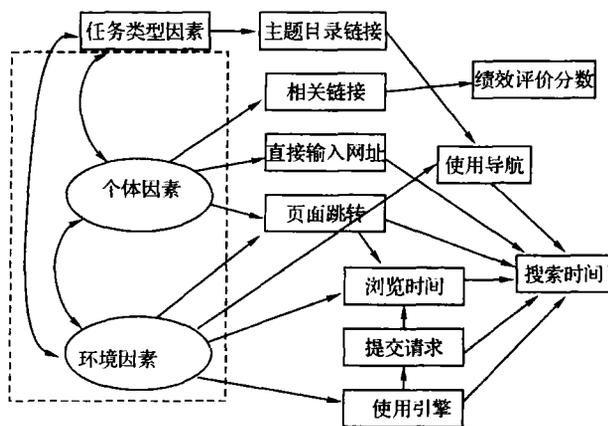


图 1 网络信息搜索的行为模型

Fig.1 Behavior model of Internet information search

表 2 网络信息搜索行为模型和模型的拟合指数

Table 2 Suitability of model

	χ^2	df	P	GFI	NFI	CFI	RMSEA
模型	74.157	33	0.000	0.826	0.692	0.790	0.141
模型	19.525	28	0.881	0.944	0.919	1.000	0.000

从表 2 可以看出,只纳入任务类型因素的模型其拟合情况并不理想.因此,作者考虑将个体因素和环境因素作为潜变量加入模型.根据先前的定性和

定量分析,认为个体的背景知识会影响是否选择“直接输入网址”的策略,人格特征等因素会影响“使用相关链接”和“页面跳转”的倾向性,网页下载等待时间和网页内容设计等网络系统环境特征则会影响到“页面跳转”、“使用导航”和“使用引擎”的频次以及浏览时间,而个体因素、环境特征和任务特征之间存在相关,它们对网络搜索行为存在交互影响.根据这一假设得到模型如图 1 所示(加入虚框部分).但对此模型进行拟合时发生错误,无法拟合数据.对各个路径系数进行检验,发现:任务类型、个体和环境等因素之间不存在显著相关;“使用导航”虽然受其他多个事件的显著影响,但对“搜索绩效”无显著效应,可考虑从模型中剔除;“使用引擎”的选择对“提交请求”的次数没有显著效应,而“直接输入网址”对其产生显著影响.根据上述分析对模型进行了相应的调整,结果如图 2 所示(模型).

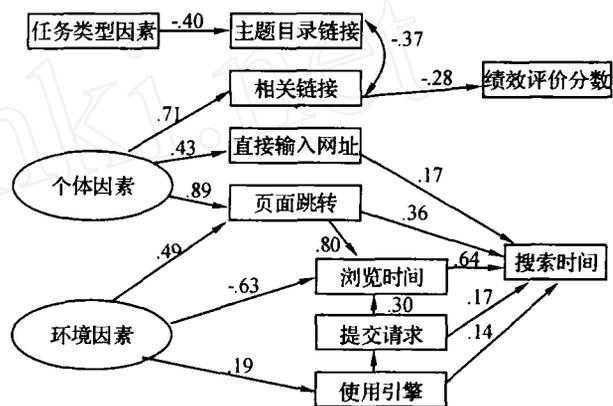


图 2 网络信息搜索的行为模型

Fig.2 Behavior model of Internet information search

模型的拟合指数如表 2 所示.综合比较这些指数,模型显然是拟合最好的.模型的路径系数也如图 2 所示。

将模型中各个变量对搜索时间和绩效的总效应分解为直接效应、间接效应和其它未分解效应.结果如表 3 所示,结合图 2 可以发现:任务类型通过“主题目录链接”与“相关链接”的相关对搜索时间产生非因果效应;个体因素通过“页面跳转”对搜索时间产生间接因果效应,通过“相关链接”对绩效产生间接因果效应;环境因素通过“浏览时间”和“提交请求”对搜索时间产生间接因果效应.个体因素是三个因素中对搜索时间和绩效产生最大因果效应的因素.在各个关键事件中,“页面跳转”、“浏览时间”和“提交请求”对搜索时间产生最大因果效应,“相关链接”是唯一对绩效评价存在因果效应的事件。

表3 网络信息搜索行为模型 的总效应分解

Table 3 Disposed effects of the factors in model

	直接效应	搜索时间间接效应	总效应	直接效应	绩效评价间接效应	总效应
任务类型因素	0	0	0	0	0	-0.040 32
个体因素	0	0.765	0.765	0	-0.193	-0.193 2
环境因素	0	0.06	0.06	0	0	0
主题目录链接	0	0	0	0	0	0.100 8
直接输入网址	-0.13	0	-0.13	0	0	0
页面跳转	0.38	0.485 1	0.865	0	0	0
相关链接	0	0	0	-0.28	0	-0.28
导航	0	0	0	0	0	0
浏览时间	0.63	0	0.63	0	0	0
提交请求	0.17	0.195 3	0.365	0	0	0
使用引擎	0.14	0	0.14	0	0	0

3 讨论

从模型 可以看出,任务类型、个体和环境因素对搜索绩效的影响主要从搜索时间上反映出来。这充分反映了网络信息的特点,即无论采用哪种策略,用户最终都可找到相关的信息,搜索绩效的差异最后只反映在搜索时间上。

任务类型因素对搜索时间和绩效无因果效应,只通过“主题目录链接”与“相关链接”之间的相关对绩效评价产生非因果效应。该结果说明,对难度相当的问题,任务类型对搜索时间和绩效基本无影响,但在不同任务类型下用户选择的策略或者行为模型可能不同。封闭型问题有确定的答案,通过主题分类层次目录,只要所选类别正确,必然可在一定层次找到所需的内容,因此对封闭型问题,“主题目录链接”使用较多;而开放型问题分类不太明确,很难根据主题目录找到所需的信息。

个体因素对搜索时间和绩效的效应最大,并主要通过“页面跳转”和“相关链接”起作用。由于背景知识和认知风格等差异的作用,最初在是否采用“直接输入网址”的策略上存在差异,其后在页面之间切换的倾向性上也存在差异,这均可导致搜索时间产生差异。认知能力和认知风格的差异决定了搜索者对信息的筛选能力、出现偏离操作的次数、关闭已读页面的倾向性等存在差异,从而在“页面跳转”次数上显示出来。个体经验对相关链接质量的判断非常重要,内容高度相关的链接能够提高绩效,减少搜索时间。

系统环境因素主要通过“浏览时间”和“页面跳转”对搜索时间产生效应。网络信息的组织方式和界

面设计都会影响用户的浏览时间,而链接的打开方式等也会影响“页面跳转”频次,从而影响到搜索时间。

从表3可以发现,在各个关键事件对搜索时间的效应中,“浏览时间”、“页面跳转”和“提交请求”的权重最大。其中,浏览是搜索过程中的主要认知活动之一,其耗费时间越长,整个搜索时间必然越长;页面跳转次数越多,跳转时出错的频次也增大,跳转和错误跳转导致的时间损耗均会导致搜索时间延长。提交请求的次数反映了重新搜索的次数,次数越多说明整个搜索过程循环越多,搜索时间越长。对绩效评价存在因果效应的只有“相关链接”,且其负荷为负值,这可能是网络信息组织不恰当造成。

根据链分析的思想,为减少搜索时间,提高绩效,应尽量减少在搜索过程中出现“页面跳转”、“提交请求”的频次和浏览时间,并让“页面跳转”和“提交请求”易于操作,不易出错。界面设计应方便浏览,容易找到主要内容。另外,提高网页中相关链接的质量对搜索绩效也非常重要。

4 结论

(1)任务类型因素通过“主题目录链接”与“相关链接”的相关对搜索时间产生非因果效应;个体因素通过“页面跳转”对搜索时间产生间接因果效应,通过“相关链接”对绩效产生间接因果效应;环境因素通过“浏览时间”和“提交请求”对搜索时间产生间接因果效应;

(2)浏览时间、“页面跳转”和“提交请求”对搜索时间有较大的影响,而“相关链接”影响搜索的准确性。

参考文献(References):

- [1] GORDON M, PATHAK P. Finding information on the World Wide Web: the retrieval effectiveness of search engines[J]. **Information Processing and Management**, 1999, 35(2): 141-180.
- [2] HSIEH-YEE I. Research on Web search behavior[J]. **Library and Information Science Research**, 2001, 23(2): 167-185.
- [3] KIM K S. Information seeking on the Web: Effects of user and task variables[J]. **Library and Information Science Research**, 2001, 23(3): 233-255.
- [4] HAWK W B, WANG P. Users' interaction with the World Wide Web: Problems and problem-solving[C]. In L Woods (Ed.), **Proceedings of the 62nd ASIS Annual Meeting**. Medford, NJ: Information Today, 2000: 256-270.
- [5] WHITE MD, IIVONEN M. Questions as a factor in web search strategy[J]. **Information Processing and Management**, 2001, 37(5): 721-740.
- [6] NAVARRO-PRIETO R, SCAIFE M, ROGERS Y. Cognitive strategies in Web searching[J/OL]. Retrieved March, 2004, 25. <http://zing.ncsl.nist.gov/hfweb/proceedings/navarro-prieto/index.html>.
- [7] SPINK A, OZMUTLU H C. Characteristics of question format web queries: an exploratory study[J]. **Information Processing and Management**, 2002, 38(4): 453-471.
- [8] OZMUTLU S, SPINK A, OZMUTLU H C. Multimedia web searching trends: 1997-2001[J]. **Information Processing and Management**, 2003, 39(4): 611-621.
- [9] SPINK A, OZMUTLU H C, LORENCE D P. Web searching for sexual information: an exploratory study[J]. **Information Processing and Management**, 2004, 40(1): 113-123.
- [10] OZMUTLU S, SPINK A, OZMUTLU H C. A day in the life of Web searching: an exploratory study[J]. **Information Processing and Management**, 2004, 40(2): 319-345.
- [11] SPINK A, SARACEVIC T. Human-computer interaction in information retrieval: Nature and manifestations of feedback[J]. **Interacting with Computers**, 1998, 10(3): 249-267.
- [12] WANG P, WILLIAM B H, CAROL T. Users' interaction with World Wide Web resources: an exploratory study using a holistic approach[J]. **Information Processing and Management**, 2000, 36(2): 229-251.
- [13] HODKINSON C, KIEL G, MCCOLL-KENNEDY R J. Consumer Web search behaviour: diagrammatic illustration of wayfinding on the Web[J]. **Human-Computer Studies**, 2000, 52: 805-830.

(责任编辑 涂红)

(上接第446页)

- [2] LAI H L, ABU KHALIL A, DUNCAN Q M C. The preparation and characterisation of drug-loaded alginate and chitosan sponges[J]. **International Journal of Pharmaceutics**, 2003, 251: 175-181.
- [3] CRINI G. Recent developments in polysaccharide-based materials used as adsorbents in wastewater treatment[J]. **Progress in Polymer Science**, 2005, 30: 38-70.
- [4] 李沙, 李馨儒, 侯新朴. 海藻酸钠-壳聚糖微囊的制备及载药性质的研究[J]. **中华临床医药**, 2002, 14(3): 1-3.
LI Sha, LI Xinru, HOU Xinpu. Preparation and drug-loading capacity of Alginate-Chitosan Microcapsules[J]. **Journal of Chinese Clinical Medicine**, 2002, 14(3): 1-3.
- [5] XIAO Liangyan, KHOR E, LIM L Y. Chitosan-Alginate-CaCl₂ system for membrane coat application[J]. **J Pharm Sci**, 2001, 90: 1134-1142.
- [6] TAKAYAMA K, HIRATA M, MASADA T, et al. Effect of interpolymer complex formation on bioadhesive property and drug release phenomenon of compressed tablet consisting of chitosan and sodium hyaluronate[J]. **Chem Pharm Bull**, 1990, 38: 1993-1997.
- [7] DOMINGUES Z R, CORTES M E, GOMES T A, et al. Bioactive glass as a drug delivery system of tetracycline and tetracycline associated with α -cyclodextrin[J]. **Biomaterials**, 2004, 25: 327-333.

(责任编辑 涂红)