

水利元数据动态分面搜索引擎的设计与实现

孔盛球,冯 钧,杜丙帅

(河海大学 计算机与信息学院,江苏 南京 211100)

摘 要:针对大量水利元数据共享需要构建搜索引擎的问题,普通用户对水利元数据领域知识的认知存在缺陷,需要引入一种探索式的访问技术准确地表达出检索请求,以实现元数据检索功能。分面搜索是一种探索式的检索方式,根据物体的多维属性,对搜索结果进行聚类,所以用户可以选择分面值对搜索结果筛选过滤。随着水利元数据的增加及水利元数据异构化程度的提高,分面的数量也不断增加。如果把所有的分面都显示给用户,容易给用户选择分面带来困难。为了将探索式的检索方式运用于水利元数据搜索领域,针对水利元数据分面过多的问题,提出了一种基于保持率的分面推荐算法,设计和实现了水利元数据的动态分面搜索引擎。实验结果表明,所提出的算法能够有效地提高用户的检索效率。

关键词:水利元数据;分面搜索;保持率;分面推荐

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2017)10-0151-05

doi:10.3969/j.issn.1673-629X.2017.10.032

Design and Implementation of Dynamic Faceted Search Engine for Water Conservancy Metadata

KONG Sheng-qiu, FENG Jun, DU Bing-shuai

(College of Computer and Information, Hohai University, Nanjing 211100, China)

Abstract: Aiming at the problem that sharing of lots of water conservancy metadata needs to build a search engine, since the defects of knowledge in the field of water conservancy metadata for ordinary users, it is necessary to introduce an exploratory access technology for users to express retrieval requests exactly to realize the function of metadata retrieval. Faceted search is an exploratory way of retrieval. According to the multi-dimensional attributes of the objects, the system clusters the search results, therefore users can choose facet values to filter them. With the increase of water conservancy metadata and the isomerization of the metadata, the number of facets is also increasing. If all the facets are displayed to users, it is difficult for them to select facets. In order to use exploratory ways of retrieval in the field of water conservancy metadata searching, aiming at the problem of too many facets of water conservancy metadata, a faceted recommendation algorithm based on retention rate is proposed, and the dynamic faceted search engine of water conservancy metadata is designed and implemented. Experimental results show that it can efficiently improve the retrieval efficiency of users.

Key words: water conservancy metadata; faceted search; retention rate; faceted recommendation

0 引 言

随着水利信息化的发展,各级部门都积累了大量的水利信息数据。为了更好地共享水利信息数据,以元数据的形式对其进行描述。通过对水利信息数据的整合,构建元数据信息的发布、发现的目录服务,实现水利信息资源的高效共享^[1]。

传统的数据检索以关键字检索为主,但是在水利元数据检索领域,引入一种探索式的检索方式有助于

用户表达正确的检索请求。分面检索也被称为引导的导航式搜索,是一种流行的和直观的交互模式,通过多维的数据让用户理解、分析和导航以发现和挖掘应用^[2]。分面是指事物的维度,一个物体是多维的,比如一本书有主题、价格、作者等维度,从不同的维度看一个物体将会得到不同的结果。用户通过输入关键字得到初步搜索结果集,系统从不同的维度对搜索结果进行聚类,并将聚类结果以分面术语的形式展示给用户。

收稿日期:2016-11-18

修回日期:2017-03-09

网络出版时间:2017-07-19

基金项目:国家自然科学基金面上项目(61370091);国家科技支撑计划课题(2015BAB07B01);水资源高效开发利用重点专项经费资助项目(2016YFC0402710)

作者简介:孔盛球(1993-),男,硕士研究生,研究方向为信息检索;冯 钧,博士,教授,研究方向为时空数据管理、智能数据处理与数据挖掘、水利信息化。

网络出版地址:http://cnki.net/kcms/detail/61.1450.TP.20170719.1112.074.html

分面搜索能够将被搜索对象的关键属性(分面术语)返回给用户,引导用户选择分面,过滤搜索结果。

水利行业各级部门都产生了大量的水利业务数据,数据是宝贵的资源,如何对日益增长的水利数据资源进行高效的管理和利用已成为水利信息化必须解决的问题^[3]。水利元数据动态分面搜索引擎能够很好地实现水利信息资源的共享。对各级部门采集的水利元数据进行汇总,将水利元数据以对象的形式存储于倒排索引,通过分面对水利元数据进行划分,能够有效地提高用户的检索效率。传统的水利信息资源检索以关键字检索为主,但是这种方式的分类效果不明显。当用户输入某个关键字时,系统会将包含这个关键字的所有检索结果都返回给用户,而用户查询目的可能很明确,仅仅只需要查询某个“负责单位”下包含这个关键字的水利元数据,因此关键字检索需要一种辅助的检索手段来提高检索效率。分面检索以关键字检索为基础,能够很好地引导用户对搜索结果进行筛选,让用户根据自己的检索意图更好地向系统表达检索请求,提升用户的检索体验。

文中将导航式的搜索方式-分面搜索引入水利元数据检索领域,针对水利元数据异构程度大而引发的分面过多的问题,提出一种基于保持率的分面推荐算法,并通过实验验证该方法的有效性。

1 分面检索技术

分面检索(Faceted Search,分面搜索)最初是一种在图书馆管理领域中常用的将多维信息空间进行正交划分的分类体系^[4-6],后逐渐发展为在结构化数据集上的探索式检索技术^[7]。用户通过发送检索请求使系统产生初步搜索结果集,从不同的维度对初步搜索结果集进行聚类产生分面与分面值,通过分面推荐算法为用户推荐最理想的分面术语。分面术语能够引导用户表达正确的检索请求,用户通过选择分面来找到预期的检索结果。由于它的便捷性,在电子商务^[8]、图书馆、音乐、电影等众多领域应用广泛。例如,马蒂·赫斯特研究的 Flamenco(弗拉明戈)项目^[9]成功地运用了层次分面的技术,具有浏览和检索相结合的界面。文献[10]介绍了分面搜索在软件开发领域的应用。还有像 RELATION BROWSER、Freebase Parallax、mSpace、Dynacet^[11]等都是成功运用分面检索的例子。

分面检索中有两个重要的概念:分面和分面值,分面和分面值统称为分面术语。分面指对象的维度,比如在水利元数据中有单位和联系人两个维度,这两个维度是水利元数据的关键属性,可以将它们作为水利元数据的分面。分面能够通过物体的重要属性对物体进行分类。分面下具体的值称为分面值,如联系人

这个分面下有个值叫张三,那么张三就叫分面“联系人”下的分面值。分面搜索就是指对初步搜索结果集聚类,生成分面和分面值,用户通过选定分面和分面值,或者去除已选分面和分面值来缩小或扩大搜索结果范围,找到用户想要的搜索结果。

2 分面推荐算法

为了给用户推荐合适的分面,需要使用合理的分面推荐算法。结合国内外相关文献,分面推荐算法主要有以下几种:选取覆盖率高的分面进行推荐、选取信息熵大的分面进行推荐、选取相关性高的分面进行推荐。

2.1 基于覆盖率的分面推荐算法

基于覆盖率的分面推荐算法在分面检索系统中比较常用。当用户输入关键字或其他查询请求时,系统将返回初步搜索结果集。将初步搜索结果集聚类到多个分面中,每个分面下都有搜索结果的数量。如果一个分面包含的搜索结果数量越高,用户所需要的搜索结果在这个分面下的概率也将会越高。由于初步检索结果集是由用户提交的检索请求得到的,如果一个分面下覆盖的初步检索结果集越大,这个分面可能与用户提交的初步检索请求关联越紧密;并且将搜索结果集大的分面推荐给用户可以避免用户选择分面后搜索结果为空的现象。综上,将覆盖搜索结果集大的分面推荐给用户是比较合理的。可以通过覆盖率来衡量分面下覆盖的初步检索结果集的大小,公式如下:

$$f(B) = \frac{\text{hits}(B)}{\text{hits}} \quad (1)$$

其中, $f(B)$ 表示分面 B 的覆盖率; $\text{hits}(B)$ 表示分面 B 下覆盖初步检索结果集的数量; hits 表示总的初步检索结果集的数量。

通过式(1)计算出每个分面的覆盖率,然后将覆盖率高的前几个分面推荐给用户。

2.2 基于信息熵的分面推荐算法

Cubranic D等开发了 Polestar 分面检索系统,提出了一种基于统计的分面导航模型^[13],在这个模型中提到了所推荐的分面能对检索空间进行有效划分。能够对检索空间进行有效划分,要求检索结果均匀地分布在分面的每个分面值当中,可以通过信息熵来衡量:

$$H_c = \sum p(c_i) \log p(c_i) \quad (2)$$

其中, H_c 表示分面 C 的信息熵; $p(c_i)$ 表示分面值 c_i 的覆盖率,指分面值 c_i 下的搜索结果总数占分面 C 下搜索结果总数的比率。

2.3 基于相关性的分面推荐算法

在用户输入关键字后,系统将返回初步搜索结果集和推荐的分面。用户通过选择分面缩小检索结果集

的范围,同时也要刷新被推荐的分面。当用户选择某个分面后,下一步被推荐的分面应该是与用户所选择的分面最相关的前几个分面。文献[9]认为,分面与分面之间是存在相关性的,而且这个相关性可以度量。对于半结构化文件 XML,节点以树状形式呈现,主节点与父节点存在一定的关系,同样父节点与子节点也存在一定的关联关系。分面其实与 XML 文件中的节点相似,因此可以类推出分面与分面之间也存在一定的关联关系。文献[9]认为分面之间的相关性可以用式(3)度量:

$$xsd = \frac{X_i Y_j}{X_i + Y_j - X_i Y_j} \quad (3)$$

其中, xsd 表示分面 i 与分面 j 的相关性; X_i 表示分面 i 下的搜索结果数量; Y_j 表示分面 j 下的搜索结果数量; $X_i Y_j$ 表示同时在分面 i 和分面 j 下的搜索结果数量。

从式(3)可以看出,当 $X_i Y_j$ 越大,计算出的分面相关性就越高,所以式(3)的核心思想是如果两个分面中共有的搜索结果数量越多,那么这两个分面的相关性就越高。

3 基于保持率的分面推荐算法

现有的分面推荐算法基本上是从覆盖率、信息熵或者相关性这几个角度出发。文献[14]描述了一种基于检索树的分面推荐算法;文献[15]通过对用户的检索日志分析来推荐分面。由于水利元数据异构程度大,数据类型复杂,仅使用现有的分面推荐算法不能获得很好的推荐效果。在传统分面推荐算法的基础上提出一种基于保持率的分面推荐算法。在用户分面检索的过程中,当用户选定 A 分面时,会出现两种情况,某些分面下的搜索结果数量将迅速减少,某些分面下的搜索结果数量几乎保持不变。对于第一种情况,认为这类分面相对于分面 A 的保持率较低;第二种情况则认为这类分面相对于分面 A 的保持率较高。

通过保持率来衡量用户所选分面与待推荐分面的关联程度。当用户选择分面 A , 这时待推荐分面中有两个分面,分面 B 和分面 C , 其中分面 B 对于分面 A 的保持率较高,而分面 C 则较低。用户选定分面 A 没有对分面 B 产生很大的影响,由此可见分面 A 与分面 B 这两个筛选条件比较接近,所以相关性较高;相反用户选定分面 A 对分面 C 产生了很大的影响,分面 A 和分面 C 这两个筛选条件存在很大的区别,所以分面 A 与分面 C 的相关性应该较低。在分面检索过程中,当用户选定分面 A 时,计算所有待推荐分面相对于分面 A 的保持率,将保持率高的分面推荐给用户。分面 B 相对于分面 A 的保持率计算如下:

$$S_{BA} = \frac{C(B/A = a_i)}{C(B)} \quad (4)$$

其中, $C(B)$ 表示分面 B 下的搜索结果数量; $C(B/A = a_i)$ 表示当用户选定分面 A 并且选择分面 A 下的分面值 a_i 时分面 B 下的搜索结果数量。

使用保持率推荐分面,能够在大量的异构数据中计算分面之间的相关性。但是分面 A 保持率高,其包含的搜索结果数量不一定大。如果把搜索结果数量少的分面推荐给用户会影响用户的检索体验,因此提出分面推荐算法将保持率与覆盖率相结合:

$$Score(B) = \alpha \frac{hits(B)}{hits} + \beta \frac{C(B/A = a_i)}{C(B)} \quad (5)$$

其中,第一部分表示分面 B 的覆盖率,用 α 表示其权重;第二部分是保持率,用 β 表示其权重; α 与 β 的取值由被搜索的数据特征决定,可以通过实验获得。

基于保持率和覆盖率的分面推荐算法流程如下:

Facets 推荐算法。

输入:用户所选分面值;

输出:被推荐的分面集。

用户输入关键词 k ,产生初步检索结果集 D

按覆盖率推荐第一组分面集 R

IF(用户选择分面 A 下的分面值 a_i)

FOR ($i = 0; i < facets.length; i++$)

计算 $Score(B_i)$;

ENDFOR

对分面 B_i 从高到低排序,推荐前 4 个分面

ENDIF

第一次分面推荐依据分面的覆盖率,后续的分面检索过程中,用户选定分面 A 下的分面值 a_i , 通过上述算法计算所有待推荐分面 B 的 $Score(B_i)$, 并将分值最高的前 4 个分面推荐给用户。

4 面向水利元数据动态分面搜索引擎系统设计

一个分面检索系统为一个物体分配了多个分类模式,并且用多种方式来表现这个物体,而不是对一个物体用预定义和简单的方式来组织^[16]。有些分面检索系统将分面固定为常用的几个,分面是不会发生变化的,随着用户的选择,显示给用户的分面将会越来越少。动态分面检索系统是针对被搜索对象存在大量分面而设计的,用户每选定一个分面都会重新推荐最合适的四个分面。相对于静态分面检索系统而言,动态分面检索系统对异构数据资源有很好的处理效果。

图 1 为面向水利元数据动态分面搜索引擎系统的结构框架图,划分为索引模块、检索模块和结果显示模块。以 Lucene 为开源搜索框架,对水利元数据进行索引,从索引中取出分面术语,通过分面推荐算法为用户

推荐分面。在用户检索过程中,使用向量空间模型对搜索结果进行排序。

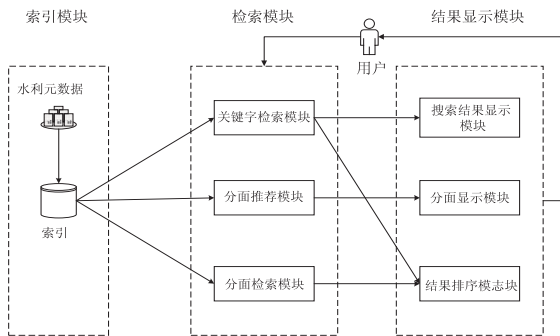


图 1 分面检索系统结构框架

4.1 水利元数据

元数据是描述数据的数据,水利元数据指描述水利对象的数据。水利信息资源元数据包含了对水利信息资源描述的全集,对水利信息行业各单位的水利信息资源进行汇总,将水利信息资源以水利元数据的形式描述,使用半结构化数据 XML 作为水利元数据的载体。使用水利元数据描述水利信息资源具有结构统一、易于存储等优势。

水利元数据是对水利信息资源的抽象提取,使用半结构化数据 XML 作为水利元数据的载体,整个 XML 文档可以抽象为一棵节点树,每个节点带有一定的语义关系。水利元数据的属性值存放在叶子节点,通过对应的路径能够查找到对应的水利元数据的属性值。在分面检索系统中,通过对应的路径把元数据的相关属性提取出来,封装成一个对象,建立到索引中。

4.2 分面推荐模块

动态分面检索系统对异构数据资源有很好的处理效果,并且能够给用户提供良好的体验。静态分面检索将分面固定于检索页面,显示给用户的分面是固定的,并且随着用户对分面的选择,固定于检索页面上的分面数量也会不断减少。动态分面检索系统中,前台显示给用户的分面是不固定的,每次给用户显示 N 个分面,随着用户的选择,只要系统后台分面个数足够,前台显示给用户的分面个数还是 N 个。针对异构数据资源产生分面过多的现象,可以使用动态分面技术为用户推荐最合适的前几个分面,而其他分面则隐藏在后台不予显示。

动态分面搜索引擎关键在于如何实现分面推荐。每次系统为用户显示搜索结果时,都会对搜索结果进行聚类,聚类后可能会产生多个分面。而动态分面搜索引擎要求每次搜索结果的刷新都需要对分面也进行刷新,而且被推荐的分面是最适合用户的前几个分面。分面推荐模块的流程如图 2 所示。

使用 Lucene 作为检索框架, Lucene 作为 Apache

开源检索框架,提供了全文检索功能。Lucene 支持将文本信息建立成倒排索引,同时能够在前台对索引进行检索。在建立索引时,需要对被设定为分面的元数据属性进行标记。由于动态分面搜索引擎包含的分面数量较大,而且随着水利元数据源源不断地加入到索引,分面数量还会不断增长。虽然动态分面检索系统能够很好地为用户推荐合适的分面,但随着分面数量的过多增加,将导致系统检索效率的降低,影响用户的检索体验。为此,需要对索引中被标记的分面设置一个上限值,并且对所有不同种类水利元数据进行统一分析,抽取有关键属性设置为分面。

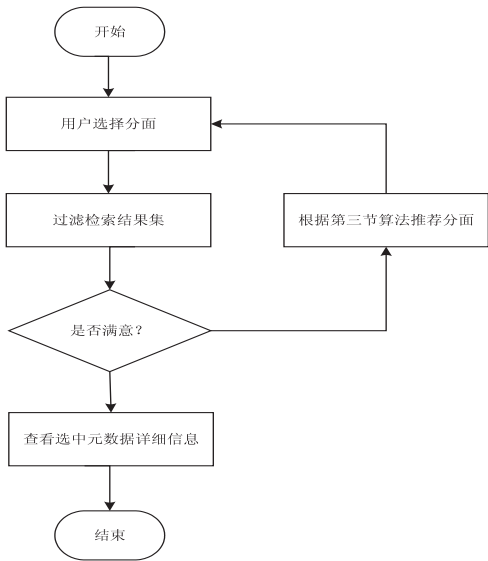


图 2 分面推荐流程

4.3 动态分面搜索引擎系统实现

将文中算法运用于实际项目中,开发了面向水利元数据的动态分面搜索引擎,系统截图如图 3 所示。

从图 3 可以看出,左边为分面检索栏。用户在输入“水库工程”关键字后得到初步检索结果集,并且在分面检索栏显示被推荐的分面。用户可以选择分面检索栏中的具体分面值对检索结果集进行过滤,同时刷新分面检索栏,重新推荐分面,直到用户找到满意的搜索结果为止。

将文中算法使用在面向水利元数据动态分面搜索引擎中。通过实验并且基于用户的反馈,发现动态分面搜索引擎能够有效地提高检索效率。被推荐的分面大部分都是用户所需要的,因此该算法在系统中起到了很大的作用。

5 实验分析

通过实验来验证该算法的有效性。在此,挑选了水利对象的七个属性作为分面,分别是:“卫星名”、“负责单位”、“地址”、“传感器”、“负责人”、“元数据标准名”、“职务”。分面栏只显示四个分面供用户选

择,虽然系统中一共存在七个分面,但是每次只给用户推荐四个。根据实验数据的特征,实验过程中对分面

评分式(5)中的 α 取值为 1, β 取值为 0。



图 3 分面检索页

整个实验过程的描述如下所示：
(1)用户在关键字搜索框中输入“北京”,系统推荐了四个分面显示在分面搜索栏中,这四个分面为:地址、负责单位、传感器、负责人。

(2)用户选择“负责单位”下的“水利部水利信息中心”分面值,系统过滤搜索结果,并重新推荐分面:传感器、负责人、地址、卫星名。

(3)用户选择“传感器”下的“MODIS”分面值,用户在第一条搜索记录中找到所需要的搜索结果。

对上述实验所采集的数据进行分析,发现当用户输入“北京”关键字后,系统检索出初步检索结果集。对检索结果集聚类后一共得出 5 个分面,另外 2 个分面中不包含检索结果。如果根据覆盖率利用式(1)为每个分面计算得分,结果如图 4 所示。

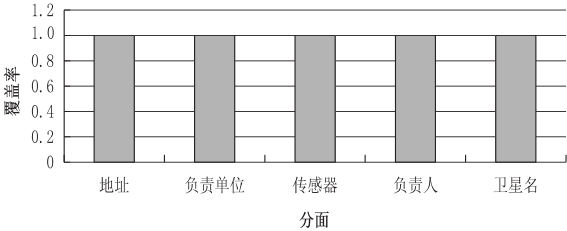


图 4 分面覆盖率条形图

从图中可以看出,如果仅使用覆盖率作为分面推荐的依据,5 个分面得分情况将一致,所以系统将推荐前 4 个分面给用户。

当用户选择分面“负责单位”下的“水利部水利信息中心”分面值时,通过该算法得出各分面分值如图 5 所示。

如图 5 所示,系统将优先推荐分值高的“传感器”与“负责人”两个分面给用户,并且用户在后续的分面检索中选择传感器这个分面时,找到了所需的检索结

果。通过分析,可以得出系统所推荐的分面大部分是用户所需的,因此该算法在实践应用中有很大的作用。

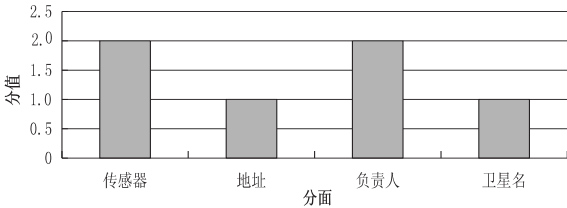


图 5 分面计算分值条形图

6 结束语

针对大量水利元数据,通过构建动态分面搜索引擎实现水利元数据的共享。在现有分面推荐算法的基础上,提出了基于保持率的分面推荐算法。将该算法运用在实际项目中,虽然取得了较好的效果,但仍然存在诸多不足。比如在算法效率上,需要为每个分面计算分值,会浪费大量的时间;在分面推荐因素上缺乏考虑热搜分面的影响。为了提高系统运行的效率,后续工作中可以在用户检索前尝试构建一个分面图,以记载每两个分面之间的保持率,在检索时可以不用每次都计算分面的保持率而直接遍历这个图。同时对于那些被用户经常点击的分面也应该优先推荐,使被推荐的分面更合理。

参考文献：

[1] 冯 钧,唐志贤,黄如春,等. 水利信息资源元数据管理方法研究[J]. 水利信息化,2011(5):1-4.
[2] Liberman S,Lempel R. Approximately optimal facet value selection[J]. Science of Computer Programming,2014,94(1): 18-31.

要求。下一步将在其他硬件平台实现定位功能并在不同噪声环境下提高定位精度。

参考文献:

- [1] Want R, Hopper A, Falcao V, et al. The active badge location system[J]. ACM Transactions on Information Systems, 1992, 10(1):91-102.
 - [2] Harter A, Hopper A. A distributed location system for the active office[J]. Network, 1994, 8(1):62-70.
 - [3] Harter A, Hopper A, Steggles P, et al. The anatomy of a context-aware application[J]. Wireless Networks, 2002, 8(2-3):187-197.
 - [4] Addelese M, Curwen R, Hodges S, et al. Implementing a sentient computing system[J]. IEEE Computer Magazine, 2001, 34(8):50-56.
 - [5] 李元实, 王智, 鲍明, 等. 基于无线声阵列传感器网络的实时多目标跟踪平台设计及实验[J]. 仪器仪表学报, 2012, 33(1):146-154.
 - [6] Luo J A, Feng D, Chen S, et al. Experiments for on-line bearing-only target localization in acoustic array sensor networks[J]. World Congress on Intelligent Control & Automation, 2010, 20(1):1425-1428.
 - [7] 丰宁宁, 王成华. 一种基于分布式协作的 WSN 定位方法[J]. 计算机技术与发展, 2012, 22(5):60-63.
 - [8] 游晓鹏, 鲍金凤, 邱奉美, 等. RSSI 辅助的蒙特卡罗盒定位算法[J]. 计算机技术与发展, 2013, 23(12):11-15.
 - [9] Melodia T, Pompili D, Gungor V C, et al. A distributed coordination framework for wireless sensor and actor networks[C]//Proceedings of the 6th ACM international symposium on mobile ad hoc networking and computing. [s. l.]: ACM, 2005:99-110.
 - [10] Akyildiz I F, Kasimoglu I H. Wireless sensor and actor networks; research challenges[J]. Ad Hoc Networks, 2004, 2(4):351-367.
 - [11] 叶菁, 陈国龙, 吴运兵, 等. 无线传感器网络任务分配的遗传优化算法[J]. 计算机工程与应用, 2011, 46(35):119-122.
 - [12] 田旺兰, 李梦醒, 谭跃. 无线传感器网络节点任务分配模型的建立[J]. 湖南城市学院学报:自然科学版, 2013, 22(1):57-61.
 - [13] 刘梅, 李海昊, 沈毅. 无线传感器网络空中目标跟踪任务分配技术的研究[J]. 宇航学报, 2007, 28(4):960-965.
 - [14] 李戈, 孟祥杰, 王晓华, 等. 国内超声波测距研究应用现状[J]. 测绘科学, 2011, 36(4):60-62.
 - [15] 张健, 李钢. 超声波测距系统的设计与研究[J]. 合肥工业大学学报:自然科学版, 2004, 27(6):640-643.
 - [16] Navarro-Serment L E, Paredis C J J, Khosla P K. A beacon system for the localization of distributed robotic teams[C]//Proceedings of the international conference on field and service robotics. [s. l.]:[s. n.], 1999.
 - [17] 张和生, 宋明耀. 提高超声测距精度的设计[J]. 电子产品世界, 2004, 13:87-89.
- +++++
- (上接第 155 页)
- [3] 成建国, 冯钧, 杨鹏, 等. 水利数据资源目录服务关键技术研究[J]. 水利信息化, 2014(6):18-21.
 - [4] Hai Z, Wilks Y. Faceted search, social networking and interactive semantics[J]. World Wide Web, 2014, 17(4):589-593.
 - [5] Goh Y M, Giess M, McMahon C, et al. From faceted classification to knowledge discovery of semi-structured text records[M]//Foundations of computational intelligence volume 6. Berlin:Springer, 2009:151-169.
 - [6] Wang Q, Ramirez G, Marx M, et al. Overview of the INEX 2011 data-centric track[C]//International workshop of the initiative for the evaluation of XML retrieval. [s. l.]:[s. n.], 2011:118-137.
 - [7] 王莉, 高仲利. 基于分面导航理论的 RDF 数据的持久化研究[J]. 计算机工程与应用, 2010, 46(9):130-133.
 - [8] 刘逸青. 基于用户体验的网站多面搜索导航研究[D]. 上海:上海交通大学, 2010.
 - [9] 郭力洁. XML 分面搜索的关键技术研究[D]. 保定:华北电力大学, 2012.
 - [10] Niu N, Mahmoud A, Yang X. Faceted navigation for software exploration[C]//19th international conference on program comprehension. [s. l.]:IEEE, 2011:193-196.
 - [11] Roy S, B. W. H., Nambiar U, et al. Dynacet: building dynamic faceted search systems over databases[C]//25th international conference on data engineering. [s. l.]:IEEE, 2009:1463-1466.
 - [12] Wang S Y, Zhong L, Jiang D S, et al. Facet description and searching of component resource[C]//International conference on computer science and software engineering. [s. l.]:IEEE, 2008:24-32.
 - [13] Dennis B M, Healey C G. Assisted navigation of complex information spaces[C]//IEEE visualization conference. [s. l.]:IEEE, 2002.
 - [14] 杜丙帅, 李士进, 冯钧, 等. 基于水利对象分类标签的分面推荐方法研究[J]. 计算机与现代化, 2015(12):90-94.
 - [15] Zwol R V, Sigurbjornsson B, Adapala R, et al. Faceted exploration of image search results[C]//Proceedings of the 19th international conference on world wide web. Raleigh, North Carolina, USA:[s. n.], 2010:961-970.
 - [16] Jin C, Hou H, Wu M, et al. Finding facet content on web by position inverted index[C]//Proceedings of the 2012 IEEE 14th international conference on high performance computing and communication & 2012 IEEE 9th international conference on embedded software and systems. [s. l.]:IEEE, 2012:1699-1703.