

遥感影像的拓扑查询研究及实现

魏丹丹^{1,2},余劲松弟^{1,2*}

(1.福州大学 福建省空间信息工程研究中心 数据挖掘与信息共享教育部重点实验室,福建 福州 350003;
2.数字中国研究院(福州),福建 福州 350003)

摘 要:目前遥感影像的查询方式多是基于关键字的形式,基于拓扑的查询方式较基于关键字的查询方式具有空间几何不变的性质,在保障查询结果的查全率和查准率方面具有重要意义。本文研究了 DE-9IM 及其拓扑接口 JTS 拓扑套件(JTS Topology Suite,JTS),利用 JTS 提供的拓扑判断功能,基于 Solr 开源搜索引擎实现了基于对地观测的 3 交模型(Earth Observation 3 Intersection Model,EO-3IM)的拓扑查询功能。实验证明,JTS 可很好地适用于 EO-3IM。同时也证明了遥感影像拓扑查询的可执行性。

关键词:拓扑查询;JTS 拓扑套件;EO-3IM;对地观测

中图分类号:TP79 **文献标识码:**A

随着对地观测(Earth Observation,EO)系统的快速发展,卫星、机载和地面遥感系统提供了一系列多源异构的影像数据^[1]。目前关于遥感影像的查询研究成果大部分是基于关键字的查询形式^[2-3]。各大遥感数据中心,如 NASA、ESA、地理空间数据云等均提供了基于关键字的查询服务,包括空间范围、时间范围、传感器类型、质量信息、空间分辨率等。而遥感影像与普通数据不同的是,它具有空间特征^[4]。普通的文本方式查询无法充分利用查询几何条件与影像之间的空间关系,不能保证查询结果的完整性和准确性。

顾及拓扑关系的遥感影像查询方式弥补了文本查询方式的缺陷,保证了结果的查全率与查准率。同时也可根据空间对象之间的拓扑交集、补集、全集返回影像,实现个性化查询。

JTS^[5]是一个用于创建和处理矢量几何对象的 Java 库。它实现了 OpenGIS Consortium Simple Feature Specification for SQL^[6]中定义的几何对象、空间拓扑核心操作算法^[7]以及 DE-9IM^[8]的所有二维拓扑关系的判断功能。Solr(<http://lucene.apache.org/solr/>)是基于 Apache Lucene 构建的开源搜索引擎,它对外提供 Web Service 的 API 接口^[9],查询

结果可以以 XML、JSON 等多种格式返回;同时也是全文搜索引擎^[10],仅支持关键字的查询方式。作为一种数据共享的有效手段^[11],OpenSearch^[12]是网站和搜索引擎以标准格式发布搜索结果,允许用户用统一的方式访问不同的符合 OpenSearch 规范的网站和搜索引擎。利用 OpenSearch 为拓扑查询提供标准化的查询接口,用户可以根据网站或搜索引擎的查询响应结果进行 RSS 源订阅^[13],无需重复输入查询条件进行遥感影像的查询,极大地缩短了用户获取数据的时间,且能使用户及时掌握数据库中最新更新的数据。

本研究通过 JTS 插件定制拓扑模型实现,将订阅查询结果以拓扑语义关键字的形式存储在数据库中;再利用 Solr 这一开源搜索引擎实现拓扑查询,并根据查询结果更新相应的索引库;然后借助 OpenSearch 规范,对 Solr 查询结果进行 atom 转换,可以实现 atom 格式响应,即生成 RSS 源,实现遥感影像拓扑查询结果的订阅,进一步实现遥感影像的自动发现功能。

1 EO 数据模型

影像搜索实质上是对影像的描述数据进行搜索。OGC 提出了用地理可扩展标记语言(Geo-

graphic Markup Language, GML) 来表示 Coverage 的方法,目的是用统一的格式集成所有形式的地理信息^[14]。GML Coverage 具有平台无关性^[15]、自我描述性等特点,可以轻松通过 HTTP 协议在分布异构的环境中进行传输,以便交换与共享^[16]。本文的 EO 数据模型是根据 EO Coverage 模型^[14]建立。该模型以遥感影像为数据源,对其进行 XML 编码,从而将遥感影像映射为 Coverage 数据,使多源异构遥感影像具有统一的编码格式,为后续影像搜索提供数据基础。Coverage 定义主要包括 4 部分:索引域 Domain set、范围类型 Range type、链接 xlink,以及元数据 metadata。如图 1 所示。

Domain set 记录了 Coverage 的时空范围信息, Range type 描述了 Coverage 的类型,xlink 表示的是 Coverage 即影像文件所在的位置,metadata 展示了 Coverage 的元数据信息。

2 DE-9IM 与 EO-3IM 的映射

2.1 模型比较

DE-9IM 即二维扩展的 9 交模型。该模型共定义了 9 种拓扑关系谓词,包括:相等(Equals)、相离

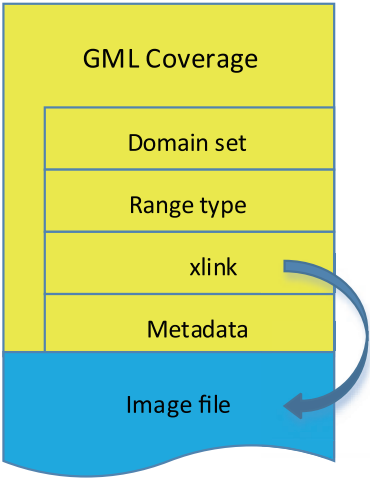


图 1 GML Coverage 模型

Fig.1 GML Coverage model

(Disjoint)、相接(Touches)、包含(Contains)、覆盖(Covers)、相交(Intersects)、被包含(Within)、被覆盖(CoveredBy)、穿过(Crosses)、重叠(Overlaps)。该模型利用一个几何对象的内部、边界、外部与另一个几何对象的内部、边界、外部的相交部分的维度来描述空间拓扑关系,定义如下:

DE-9IM(a,b) = [dim(I(a) ∩ I(b)) dim(I(a) ∩ B(b)) dim(I(a) ∩ E(b))
dim(B(a) ∩ I(b)) dim(B(a) ∩ B(b)) dim(B(a) ∩ E(b))
dim(E(a) ∩ I(b)) dim(E(a) ∩ B(b)) dim(E(a) ∩ E(b))]。

对于 DE-9IM 的 9 交矩阵的每个元素,共有 5 种取值方式{-, 0, 1, 2, 3},理论上共有 4⁹ = 262144 种可能性。其中,-表示相交部分为空集,0 表示相交于点,1 表示相交于线,2 表示相交于面,3 表示相交于体。该模型应用于遥感影像的拓扑查询时比较复杂。因此有必要提出一种适合于影像查询的拓扑模型。

EO-3IM 即对地观测扩展的 3 交模型,仅用以区别点/面、线/面、面/面、体/面 4 种类型。该模型定义了 8 种拓扑关系谓词:包括相等(Equals)、相离(Disjoint)、相接(Touches)、包含(Contains)、相交(Intersects)、被包含(Within)、被覆盖(CoveredBy)、穿过(Crosses)、重叠(Overlaps)。该模型使用用户的查询条件(点、线、面、体)与遥感影像(面)的内部、边界、外部相交部分的维度来描述空间拓扑关系,定义如下:

EO-3IM(a,b) = [dim(a ∩ I(b)) dim(a ∩ B(b))
dim(a ∩ E(b))]。

对于 EO-3IM 的 3 交矩阵的每个元素,共有 5 种取值方式{∅, 0, 1, 2, 3}。较于 DE-9IM 拓扑关系的冗余度高达 99.97%,EO-3IM 拓扑关系的冗余度仅为 79.5%,大大降低了拓扑模型的复杂度。然而,EO-3IM 的拓扑关系区分能力却不及 DE-9IM。为了弥补这一缺陷,EO-3IM 将 1 分成两个部分:1(┐)表示线且有两段端点;1(□)表示线且为闭环,即起点与终点是同一点,无端点,以区分更多的拓扑关系,满足后续遥感影像拓扑查询的需要。其中,1(□)的取值情形有且仅当 dim(a ∩ I(b)) = 2 时,dim(a ∩ B(b))才有取 1(□)的可能。1(□)共有 5 种可能拓扑情形,其他的共有 5³ = 125 种拓扑情形。理论上共有 130 种可能性。其中:∅代表相交于空集,0 表示相交于点,1(┐)表示相交于线且有两个端点,1(□)表示相交于线且为闭环,区别于 1(┐),2 表示相交于面,3 表示相交于体。EO-3IM 能区分的具体拓扑关系如图 2 所示。

	相离	相接(点)	相接(线)	相接(面)	被包含	穿过(点)	穿过(线)	穿过(面)	重叠	包含	相等	被覆盖(点)	被覆盖(线)
点-面	 [0,0,0]	 [0,0,0]			 [0,0,0]								
线-面	 [0,0,1(-)]	 [0,0,1(-)]	 [0,1(-),0]	 [0,1(-),1(-)]	 [1(-),0,0]	 [0,0,1(-)]	 [1(-),0,1(-)]					 [1(-),0,0]	
面-面	 [0,0,2]	 [0,0,2]	 [0,1(-),2]		 [2,0,0]	 [1(-),0,2]	 [1(-),0,2]		 [2,1(-),2]	 [2,1(□),2]	 [2,1(□),0]	 [2,0,0]	 [2,1(-),0]
体-面	 [0,0,3]	 [0,0,3]	 [0,1(-),3]			 [2,0,3]	 [2,1(-),3]	 [2,0,3]		 [2,1(□),3]			

图 2 拓扑关系语义规则
Fig.2 Semantic rules for topological relationships

2.2 模型映射

DE-9IM 与 EO-3IM 的拓扑编码是两种不同的编码模式。以重叠关系为例,如图 3 所示。DE-9IM 的拓扑编码为 212101212, EO-3IM 的拓扑编码为 21(¬)2。而这两种编码模式表达的拓扑关系是一样的。本研究将根据拓扑谓词进行 DE-9IM 到 EO-3IM 的拓扑关系的映射。具体映射规则如表 1 所示。

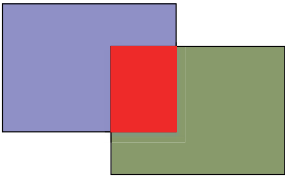


图 3 重叠关系
Fig.3 Overlapping relationship

表 1 拓扑关系映射规则
Tab.1 Mapping rules for topological relationships

拓扑关系谓词	DE-9IM	EO-3IM
相等	$\begin{bmatrix} T & * & F \\ * & * & F \\ F & F & * \end{bmatrix}$	[2 1(□) F]
包含(覆盖)	$\begin{bmatrix} T & * & * \\ * & * & * \\ F & F & * \end{bmatrix} \begin{bmatrix} * & T & * \\ * & * & * \\ F & F & * \end{bmatrix} \begin{bmatrix} * & * & * \\ T & * & * \\ F & F & * \end{bmatrix} \begin{bmatrix} * & * & * \\ * & T & * \\ F & F & * \end{bmatrix}$	[2 1(□) T]
被覆盖	$\begin{bmatrix} T & * & F \\ * & * & F \\ * & * & * \end{bmatrix} \begin{bmatrix} * & T & F \\ * & * & F \\ * & * & * \end{bmatrix} \begin{bmatrix} * & * & F \\ T & * & F \\ * & * & * \end{bmatrix} \begin{bmatrix} * & * & F \\ * & T & F \\ * & * & * \end{bmatrix}$	[T T F]
穿过	$\begin{bmatrix} T & * & T \\ * & * & * \\ * & * & * \end{bmatrix} \begin{bmatrix} T & * & * \\ * & * & * \\ T & * & * \end{bmatrix} \begin{bmatrix} 0 & * & * \\ * & * & * \\ * & * & * \end{bmatrix}$	[T F T] [T 0 T]
相离	$\begin{bmatrix} F & F & * \\ F & F & * \\ * & * & * \end{bmatrix}$	[F F T]
相接	$\begin{bmatrix} F & T & * \\ * & * & * \\ * & * & * \end{bmatrix} \begin{bmatrix} F & * & * \\ T & * & * \\ * & * & * \end{bmatrix} \begin{bmatrix} F & * & * \\ * & T & * \\ * & * & * \end{bmatrix}$	[F T *]
重叠	$\begin{bmatrix} T & * & T \\ * & * & * \\ T & * & * \end{bmatrix} \begin{bmatrix} 1 & * & T \\ * & * & * \\ T & * & * \end{bmatrix}$	[2 1(¬) 2]
被包含	$\begin{bmatrix} T & * & F \\ * & * & F \\ * & * & * \end{bmatrix}$	[T F F]
相交	$\begin{bmatrix} T & * & * \\ * & * & * \\ * & * & * \end{bmatrix} \begin{bmatrix} * & T & * \\ * & * & * \\ * & * & * \end{bmatrix} \begin{bmatrix} * & * & * \\ T & * & * \\ * & * & * \end{bmatrix} \begin{bmatrix} * & * & * \\ * & T & * \\ * & * & * \end{bmatrix}$	[T * *] [* T *]

其中： F 表示取 $\emptyset$ ； T 代表取 0、1、2； $*$ 代表取 $\emptyset$ 、0、1、2。

3 基于 JTS 的拓扑查询实现

JTS 目前实现的功能是针对二维空间中的简单几何对象,而 EO-3IM 定义的是三维空间中的拓扑模型。要使 JTS 与 EO-3IM 无缝衔接,需要将 EO-3IM 的三维拓扑空间中的拓扑对象转换成二维拓扑空间中的拓扑对象。由于 EO-3IM 仅定义两个方向的拓扑关系:水平方向以及铅锤方向,水平方向直接按照表 1 映射规则进行映射即可;铅锤方向的线可转换成多点、面可转换成多线、体可转换成多面。如图 4 所示。再按照表 1 定义的映射规则,与 DE-9IM 进行一一映射。

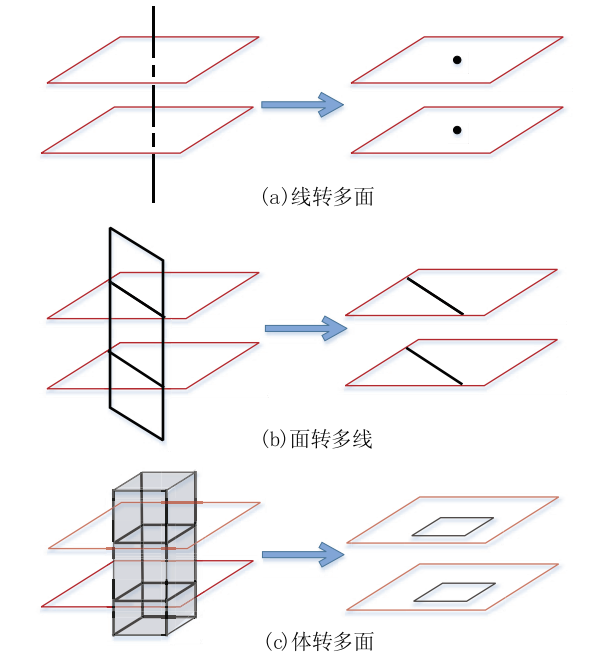


图 4 降维处理

Fig.4 Dimensionality reduction

Solr 是基于 Apache Lucene 构建的高并发、高效率的企业级开源搜索引擎^[17]。它附带了一个基于 HTTP 的管理界面^[18]。用户可以通过 HTTP GET 操作,向 Solr 提交查询请求,得到 XML、JSON、CSV 等格式的返回结果。通过 xml 脚本配置,可以与数据库连接,并对其进行增删改查操作。Solr 可扩展性极强,通过关键字“Search”,支持一个或多个关键字查询。OpenSearch 作为一个开源搜索协议,仅定义一个查询字段“SearchTerm”,其他的查询字段均可自定义扩展,与 Solr 查询机制相似。因此 Solr 与 OpenSearch 可无缝融合使用。

本研究通过 JTS 插件定制拓扑模型实现,将订

阅查询结果存储在数据库中,Solr 更新相应的索引库,再以拓扑语义关键字的形式返回订阅结果。结果通过 OpenSearch 进行转换为 atom 格式,生成 RSS 源。如图 5 所示。

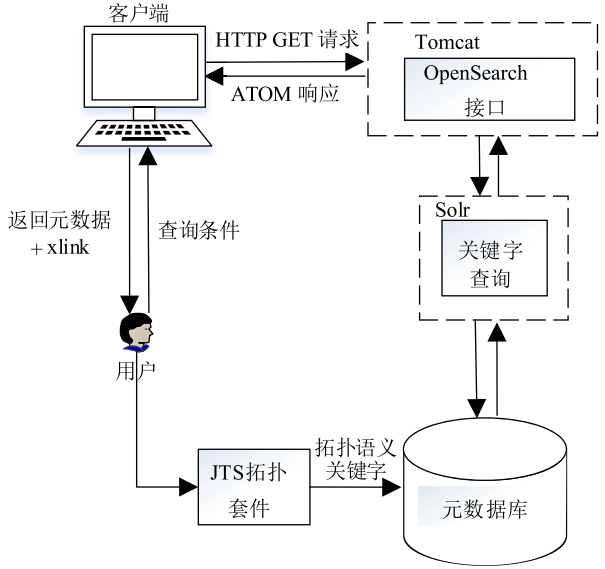


图 5 原型架构图

Fig.5 Diagram of prototype architecture

4 实验结果

4.1 数据集

实验选取两个数据源进行集成,分别是地理空间数据云 (<http://www.gscloud.cn/>) 中的 landsat 系列以及国家综合地球观测数据共享平台 (<http://www.chinageoss.org>) 的 GF-1。共 93911 元数据利用爬虫技术获得,对所获取的 csv 格式的元数据,基于 ISO 19163-1^[19] 的 XML Schema 进行 XML 映射,再对 XML 映射后的影像元数据进行一致性检验,只将通过验证的元数据存储到数据库中。如图 6 所示。

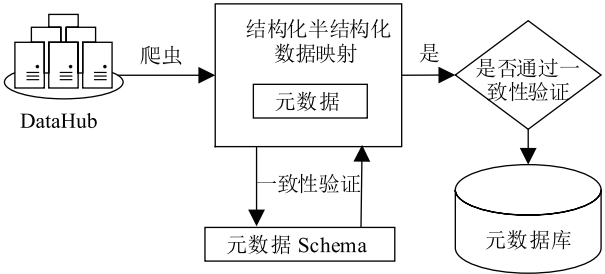


图 6 元数据一致性验证流程

Fig.6 Process of metadata consistency verification

4.2 数据查询案例

福州市位于中国东南沿海,地处闽江三角洲,其主城区地理范围为北纬 28°30'~29°30',东经

118°30′~119°30′,是“一带一路”的重要门户枢纽。随着城镇化的快速发展,在经济进步的同时,也导致了生态失衡、资源分布不均等问题。在开展城镇化的进程中,必会引起土地利用/覆盖变化。因此,研究福州主城区的土地利用变化,对减缓该地区的城镇化进程与生态环境之间的矛盾具有重要意义。为了分析福州主城区的土地利用变化,需要获取包含它的所有影像数据。根据 OpenSearch 描述文档 (OpenSearch Description Document, OSDD) 的模板定制 OpenSearch 查询请求:http://localhost:8983/solr/OpenSearch/select?fq=fuzhoucontain:yes&q=*:*&wt=xslt&tr=fuzhou.xsl, OpenSearch 查询响应返回数据库中所有符合查询条件的共 201 条数据。如下示例所示。用户根据该响应结果即 RSS 源订阅自己的搜索引擎,一旦数据库中有符合该条件的数据,更新便会及时推送给用户。

```
示例:
<feed
xmlns: opensearch = " http://a9. com/-/spec/
opensearch/1.1/"
xmlns=" http://www.w3.org/2005/Atom">
<title>EO Data for fuzhou</title>
<subtitle>201 results found in 0 seconds.</subtitle>
<logo>http://localhost:8983/solr/admin/solr\_small.
png</logo>
<updated/>
<author>
<name>Dan</name>
</author>
<id>
http://localhost:8983/solr/OpenSearch/select?
fq=fuzhoucontain:yes&q=\*:\*&wt=xslt&tr=fuzhou.xsl
</id>
<opensearch:totalResults>201</opensearch:totalRe-
sults>
<opensearch:startIndex>0</opensearch:startIndex>
<opensearch:itemsPerPage>20</opensearch:item-
sPerPage>
<opensearch:Query role=" request" searchTerms =
"*:*" startPage="1"/>
<link rel="self" type="application/atom+xml" href
```

```
= "select? q= * : * &wt=xslt&tr=fujian.xsl&start=
0&rows=20"/>
<link rel="first" type="application/atom+xml" href
="select? q= * : * &wt=xslt&tr=fujian.xsl&start=
0&rows=20"/>
<link rel="next" type="application/atom+xml" href
="select? q= * : * &wt=xslt&tr=fujian.xsl&start=
20&rows=20"/>
<link rel="last" type="application/atom+xml" href
="select? q= * : * &wt=xslt&tr=fujian.xsl&start=
200&rows=20"/>
<link rel=" search" type=" application/opensearch-
description+xml" href=" opensearch_Fujian.xml"/>
<entry>
<title>LC81190422013280LGN00</title>
<link href=" http://localhost:8983/solr/meta-
data/select? q=id:LC81190422013280LGN00"/>
<metadata>
<id>LC81190422013280LGN00</id>
<platformInfo>LANDSAT8</platformInfo>
<sensorInfo>OLI_TIRS</sensorInfo>
<path>119</path>
<row>42</row>
<acqisitionTime>2013-10-07T00:00:00Z
</acqisitionTime>
<beginPosition>2013/10/7 2:34</begin-
Position>
<endPosition>2013/10/7 2:34</endPosi-
tion>
<cloudCoveragePercentage>77.69</cloud-
CoveragePercentage>
<RUlat>26.6671</RUlat>
<RUlon>120.051</RUlon>
<RDlat>24.9321</RDlat>
<RDlon>119.609</RDlon>
<LDlat>25.3074</LDlat>
<LDlon>117.767</LDlon>
<LUlat>27.0415</LUlat>
<LUlon>118.182</LUlon>
<illuminationElevationAngle>53.42021221
</illuminationElevationAngle>
<illuminationAzimuthAngle>146.299694
```



```
</illuminationAzimuthAngle>
</metadata>
</entry>
<entry>...</entry>
...
<entry>...</entry>
```

5 结语

本研究为遥感影像提供了一个基于拓扑查询的订阅系统。首先,通过 JTS 拓扑套件实现了 EO-3IM 的拓扑关系判断功能。其次,借助 OpenSearch 规范,对 Solr 查询结果进行 atom 转换,可以实现 atom 格式响应,即生成 RSS 源,实现遥感影像拓扑查询结果的订阅,进一步实现了遥感影像的自动发现功能。然而,本研究的拓扑查询借助了 JTS 拓扑套件进行功能实现,需要将 EO-3IM 与 DE-9IM 进行映射才可应用。因此,如何开发 EO-3IM 的拓扑查询算法是今后的研究重点。

参考文献:

[1]JI Z X, CHAO W Y, QING Q L. Building a spatiotemporal index for earth observation big data[J]. International Journal of Applied Earth Observation and Geoinformation, 2018(73):245-252.

[2]秦昆,陈一祥,甘顺子,等. 高分辨率遥感影像空间结构特征建模方法综述[J]. 中国图象图形学报, 2013, 18(9):1055-1064.

[3]邵奇峰,李枫. 一种基于 HBase 的空间关键字查询算法[J]. 计算机工程与科学, 2015, 37(11):2084-2090.

[4]鲁学军,王钦敏,明冬萍,等. 空间特征在遥感影像分析中的应用[J]. 中国图象图形学报, 2004, 9(6):737-743.

[5]MARTIN D. JTS Topology Suite[EB/OL]. [2019-11-1]. <https://github.com/locationtech/jts>.

[6]OGC 99-049. OpenGIS simple feature specification for SQL revision 1.1[S]. Massachusetts: Open Geospatial Consortium, 1999.

[7]余靖毅,邬伦,高勇. 基于 Storm 的地理编码引擎[J]. 地球信

息科学学报, 2015, 17(12):1431-1441.

[8]CLEMENTINI E, FELICE P D. A comparison of methods for representing topologic relationships [J]. Information Science, 1995 (3):159-178.

[9]MUSTAFA A, TUNCAY A, MUHAMMER I. An analysis on the comparison of the performance and configuration feature of big data tools Solr and Elasticsearch [J]. International Journal of Intelligent Systems and Applications in Engineering, 2016, 4(Special Issue): 8-12.

[10]KERRY K. Advanced search techniques with Hadoop, Lucene, and Solr[J]. Pro Hadoop Data Analytics, 2016(30):91-136.

[11]HAN W G, DI L P, YU G N, et al. Investigating metrics of geospatial web services: the case of a CEOS federated catalog service for earth observation data[J]. Computer & Geoscience, 2016, 92:1-8.

[12]DEWITT C. OpenSearch 1.1 draft 6 [EB/OL]. (2019-12-9) [2020-3-31]. <https://github.com/dewitt/opensearch/blob/master/opensearch-1-1-draft-6.md>.

[13]DEVJI K, XITIJI S. Fossick: an implementation of federated search engine [J]. International Journal of Computer Science Engineering and Information Technology Research, 2016, 6(1):69-78.

[14]OGC 09-146r2. Implementation schema for coverage[S]. Massachusetts: Open Geospatial Consortium, 2012.

[15]余劲松弟,陈远杰,佟瑞菊. 基于 Web 互操作的遥感生态指数计算及应用[J]. 贵州大学学报(自然科学版), 2018, 35(5):41-47.

[16]余劲松弟,吴升. 基于多标准的 Web 互操作时序格网建模研究[J]. 福州大学学报(自然科学版), 2014, 42(6):859-863.

[17]李戴维,李宁. 基于 Solr 的分布式全文检索系统的研究与实现[J]. 计算机与现代化, 2012(11):171-176.

[18]魏涛,孟方圆,袁平,等. 开源搜索引擎 Elasticsearch 和 Solr 对比和分析[J]. 现代计算机, 2018, 6(2):58-61.

[19]ISO/TS 19163-1. Geographic information——Content components and encoding rules for imagery and gridded data——Part 1: content model[S]. Switzerland: International Organization for Standardization, 2016.

(责任编辑:曾 晶)

(下转第 92 页)