

基于 SUMO 的交通网格决策本体层次模型设计 *

王煜¹, 陈学广¹, 洪流¹, 龚静雯^{1,2}

(1. 华中科技大学 系统工程研究所, 武汉 430074; 2. 武汉理工大学 湖北省数字制造重点实验室, 武汉 430070)

摘要: 交通信息资源的合理集成是提高智能交通系统性能的有效手段。为集成交通网格上异构信息与不同领域本体, 利用网格作为智能交通系统的信息资源存储与协同工作平台, 利用本体作为智能交通系统资源表示和信息描述的基础, 提出了基于 SUMO 的交通网格决策本体层次结构, 论述了异构信息获取机理。实例证明此模型有利于交通信息资源的整合利用, 并为基于网格的智能交通系统的决策支持及与网格上其他系统之间的开放式协作奠定了语义基础。

关键词: 智能交通系统; 交通网格; 决策本体; 层次模型; 推荐高层合并本体

中图分类号: TP31; TP393 **文献标志码:** A **文章编号:** 1001-3695(2009)04-1373-04

Design of SUMO-based ontology layered model for transportation grid resource

WANG Yu¹, CHEN Xue-guang¹, HONG Liu¹, GONG Jing-wen^{1,2}

(1. Institute of Systems Engineering, Huazhong University of Science & Technology, Wuhan 430074, China; 2. Hubei Digital Manufacturing Key Laboratory, Wuhan University of Technology, Wuhan 430070, China)

Abstract: Reasonable integration of transportation information resources can effectively improve the performance of ITS. In order to achieve the integration goal of ITS, this paper proposed a novel SUMO-based four-layer ITS decision ontology by using grid as resource storage and cooperation platform, and utilizing ontology as resource expression base. Based on the model, this paper discussed a mechanism of heterogeneous information access and demonstrated it by a traffic accident scenario. The result shows that this model availablely integrates the heterogeneous resources and different domain-ontology by mapping their semantic to SUMO and promotes the decision support and open collaboration capacities of ITS.

Key words: ITS; transportation grid; decision ontology; layered model; SUMO (suggested upper merged ontology)

0 引言

智能交通系统^[1] (ITS) 是由一系列既相对独立又协同工作的子系统组成, 其运行效率很大程度上依赖于这些子系统的资源共享与协同工作能力。由于城市交通信息瞬息万变, 同时不同来源的各类信息其表现形式不尽相同, 使交通信息具有容量的海量性和内容的异构性^[2] 等特点, 这就造成了不同区域、不同部门之间的交通决策信息难以有效地共享与使用, 限制了交通决策支持能力的提供。因此, 要解决 ITS 系统的协同工作与资源调度问题, 必须从资源整合入手, 解决交通信息的海量性存储与基于知识的交通决策信息资源的集成问题。

为解决 ITS 资源共享与信息集成的问题, 很多研究者在交通本体构建及应用上作了尝试。黄珂萍等人^[2] 面向公交系统建立了市内交通本体, 并定义了类内公理及实例的基本规则; 刘卫宁等人^[3] 从实际遗留系统的集成出发, 采用虚拟集成法将异地数据通过元数据及元属性进行整合, 以构建 ITS 虚拟共用信息平台; 李阳等人^[4] 面向中国智能交通系统的语义集成问题, 提出了概念化的三层本体体系结构, 并给出了系统的本体分类方法; 马金素等人^[5] 构造了交通问题求解本体, 并论述了基于本体的类比匹配机制和匹配算法。

以上的方法解决了 ITS 信息集成及交通本体构建的一些问题, 但仍存在着以下不足:

a) 交通领域内本体多为个人自发构建, 缺乏一致的概念基础与表示形式, 易出现系统问题定义不清、本体边界模糊、概念粒度不当、公理规则缺乏针对性等问题, 造成已建本体难以识别与重用。

b) 目前的交通本体构建多以单一应用为导向, 仅仅注重 ITS 某个侧面的应用实际而未提供满足 ITS 框架的本体层次模型。

c) ITS 系统是整个城市信息系统的组成部分, 其本体构建需考虑到开放性与扩展性要求。而目前交通本体构建并未考虑跨领域的信息共享和知识集成, 因此在系统集成时各领域的本体需要相互翻译与理解。由于缺乏一致的建模基础及设计模式, 随着系统的增长和发展, 交通本体通常必须进行重新设计以排除那些导致系统不能适应新需求的底层假设。

d) 目前的交通本体均采用英文词汇库作为本体概念与属性的词汇基础, 难以适应我国以中文为核心的交通信息的分解与重构需求。

本文从 ITS 数据集成、信息集成、应用集成与系统扩展的整体性出发, 利用推荐高层合并本体 SUMO 作为核心本体, 提

收稿日期: 2008-05-30; 修回日期: 2008-07-16 基金项目: 国家自然科学基金资助项目 (60773188); 武汉市重点科技攻关项目 (200710321090-2)

作者简介: 王煜 (1982-), 男, 湖南长沙人, 博士研究生, 主要研究方向为决策支持、语义网络 (amenem1982@yahoo.com.cn); 陈学广 (1947-), 男, 教授, 博导, 主要研究方向为决策支持系统、信息系统工程; 洪流 (1978-), 男, 博士后, 主要研究方向为复杂网络、网格技术、决策支持; 龚静雯 (1981-), 女, 博士研究生, 主要研究方向为网格技术、数据挖掘。

出并论述了交通网格决策本体的层次模型,为基于网络的 ITS 的决策支持及跨系统之间的开放式协作奠定语义基础。

1 基于 SUMO 的交通网格决策本体层次模型

1.1 推荐高层合并本体 SUMO

推荐高层合并本体^[6](SUMO)是通用上层本体的一种,最初由 Lan Niles 和 Adam Pease 开发,后提交给 IEEE Standard Upper Ontology (SUO) Working Group 作为推荐标准,包含 1 770 个概念、7 278 条断言和 1 240 条规则。作为轻量级本体的 SUMO 合并了一系列现有的高层本体,试图通过整合不同上层本体的关系和分类来改善语义 Web 的搜索、交流与互操作。SUMO 包括三个基本层次:核心层由 11 个子集组成,包括结构本体(structural ontology)、基础本体(base ontology)、集合/类型理论(set/class theory)、数学部分(numeric)、基于 Allen 的现实关系构建部分(temporal)、拓扑部分(mereotopology)、图论部分(graph)、计量单位(measure)及局部过程、对象和属性的子结构与公理部分;第二层为中间层本体(mid-level ontology);最下层为基于上两层的领域本体部分。

SUMO 适合于交通网格决策本体的构造,具体表现在:

a) 统一的建模基础。SUMO 整合了目前存在的多数高层本体,其上层知识本体以一般、抽象、哲学的元概念为基础,足够涵盖广阔范围的领域区域。不同于从叙词表或网络信息抓取领域本体的概念关系,SUMO 提供了一个预先定义好的上层通用本体概念基础,适合于领域本体在其上的扩展。SUMO 已经实现了领域本体与其核心层之间的映射,包括政府、财政、运输、地理、军事等领域本体,这使得 SUMO 能提供超过 20 000 个概念与 6 000 条公理,为这些领域内具体应用本体的建立和交互奠定了建模基础,大大提高了相似领域本体的开发效率。

b) 明确的概念化标准。SUMO 改变了基于叙词表的本体构建中存在的语义关系不明确的情况,通过 SUO-KIF 语言进行形式化描述,并定期自动地转换为 OWL、LOOM、DAML、XML 格式。这使得 SUMO 的概念具有明确的语义描述形式和完全的公理化,而其多本体描述语言的相互映射也适用于基于网络的决策支持系统中不同的本体描述语言定义的本体的共享与复用。

c) 一致的本体设计模式。SUMO 所定义的时空、计量、代数概念具有很强的实用性与重用性,符合一般哲学上关于时空、代数的共同认知,能够为更深层次的本体设计与建模一致性提供设计模式。

d) 强大的可扩展性。SUMO 内的信息远远超过字典或术语表,能使计算机处理更多内容细节和结构。首先,SUMO 与目前世界上最大的叙词表 WordNet^[7]建立了映射,而 WordNet 丰富的词汇集能够为基于 SUMO 的各个领域本体的概念与关系描述提供充分的词汇选择;其次,SUMO 具有语言生成模板,可以转换成印度语、中文、意大利语、德语、捷克和斯洛伐克语和英语,也为不同语种的领域建模提供了扩展的基础。

1.2 交通网格决策资源本体层次模型

本文采用 SUMO 作为交通领域决策资源本体的通用上层本体,将交通网格决策本体分为四层结构(图 1),以满足智能交通系统跨领域协作和提供非凡决策支持能力的要求。四层结构从智能交通系统的不同功能层次出发,满足了由语法到语

义、由基础设施到高层应用、由信息检索到系统集成的 ITS 功能需求。以 SUMO 为核心的四层结构改善了以往 ITS 交通本体在开放性和扩展性方面的不足,更利于各系统间的协同工作与统一指挥。

1) 语法层 该层包括 WordNet 与 Sinica BOW^[8]。这一层的作用是作为信息萃取时所需要的自然语言分解与翻译的基础,以及领域本体构建时词汇选择与本体结构的基础。其中,WordNet 构成英语的词汇表,且是下层 SUMO 的词汇映射基础,为交通信息的查找、萃取与组合提供了强大的词汇关系库;Sinica BOW 是台湾中央研究院语言学研究所所以 SUMO 和 WordNet 为基础建立的中研院中英双语知识词网(Academia Sinica Bilingual Ontological WordNet),为中文的本体研究提供基础架构,且是中文自然语言处理和应用知识本体的依据,可以利用其中文词汇与 WordNet 英语词汇及 SUMO 概念的映射进行英中概念转换和本体表示,改善目前大多数本体表示语言与工具对中文推理支持不足的局面。

2) 通用上层本体层 该层本体层即为 SUMO 本体层。这一层主要为领域本体结构的构建与领域本体之间的集成提供一致性的语义基础。通过将各领域本体相关概念对 SUMO 概念的映射或包含连接,可以将不同的小本体组合成具有层次性与相关性的本体集,同时也为与跨领域网络上其余使用 SUMO 为基础的领域本体的集成和扩展提供了基础。

3) 领域本体层 本文通过对 SUMO 概念和属性的继承或关联构建交通领域本体。从主要数据源(GIS 数据和视频数据)、交通领域知识、组织结构出发,本文将交通网格决策资源本体集定义为以下四个领域子本体:

(a) 地理本体(geographic ontology, GO)。它提供与地理信息相关的概念与关系,包括道路、路口、建筑、天气、国家、城市、位置等;此本体提供与 GIS 数据及地图转换层次一致的本体概念集。图 2 显示了地理本体对 SUMO 中基础概念的引用关系及概念、属性扩充。其中:mainStem 表示概念“主干道”,为 SUMO 相应概念 roadway“道路”的子类;maxSpeedAllowed 为增加的道路属性“最高时速”。这些概念及属性的设置为 ITS 提供了必要的 GIS 信息的语义映射基础。

基于 SUMO 的交通网格决策资源本体集片段程序如下:

```
<owl: Class rdf: about = " http://www. owl-ontologies. com/GeographyOntology. owl#MainStem" >
  <rdfs: subclass of rdf: resource = " http://www. ontologyportal. org/translations/SUMO. owl#Roadway" />
</owl: Class >
<owl: Class rdf: about = " http://www. owl-ontologies. com/GeographyOntology. owl#LateralRoadway" >
  <rdfs: subclass of rdf: resource = " http://www. ontologyportal. org/translations/SUMO. owl#Roadway" />
  <owl: disjointwith >
  <owl: Class rdf: ID = " MainStem" />
</owl: disjointWith >
</owl: Class >
<owl: ObjectProperty rdf: ID = " maxSpeedAllowed" >
  <rdfs: domain rdf: resource = " http://www. ontologyportal. ogr/translations/SUMO. owl#Roadway" />
  <rdfs: range rdf: resource = " http://www. ontologyportal. org/translations/WUMO. owl#Measure" />
</owl: ObjectProperty >
```

```

<owl:Class rdf:about = "http://www.owl-ontologies.com/GeographyOntology.owl#Hospital">
  <rdfs:subClassOf rdf:resource = "http://www.ontologyportal.org/translations/SUMO.owl#Building"/>
  <rdfs:comment>A building which provides medical care to patients.
</rdfs:comment>
</owl:Class>
<owl:ObjectProperty rdf:ID = "near">
  <rdfs:domain rdf:resource = "http://www.ontologyportal.org/transla-
tions/SUMO.owl#Building"/>
  <rdfs:range rdf:resource = "http://www.ontologyportal.org/transla-
tions/SUMO.owl#Roadway"/>
</owl:ObjectProperty>

```

(b) 交通本体 (transportation ontology, TO)。该本体提供与交通信息相关的概念与关系,又分为静态本体和事件本体两类。其中静态本体描述固有的、实物性的概念,如车辆、红绿灯、站台、乘客等;事件本体则描述临时的、动作类的概念,如堵车、交通事故、加油等。

(c) 多媒体本体 (multimedia ontology, MO)。交通视频数据是智能交通决策支持系统中,用于交通流量预测、交通事故鉴定、交通路径诱导等高层决策支持的重要数据源。为满足多媒体语义信息和形象特征信息的整合,本文采用多媒体内容描述接口 MPEG-7 的概念扩展——基于 MPEG-7 架构的多媒体本体^[9]。MPEG-7 是 MPEG 组织制定的国际标准,它规定了一组能有效描述各种多媒体信息的描述器以及它们之间的内在联系,通过其定义的描述定义语言 DDL (description definition language) 形成标准化的 XML Schema 描述,提供了一个通用、灵活、可扩展的多媒体内容描述机制 MDS (multimedia description schemes) 用于描述多媒体环境中的视觉信息和音频信息。通过与 MPEG-7 架构的整合,本文利用所定义的多媒体本体对图像、图形、图片等多媒体文件进行符合交通领域信息 (如车祸地点、发生时间、发生车辆等) 的语义文字描述,为基于语义文本的交通监控多媒体文件查找提供语义基础。

(d) 交通组织本体 (organization ontology, OO)。交通网格决策支持系统的一个重要功能是实现各部门协同工作,因此需要定义各个部门的相互关系与数据权限。本文采用根据中国智能交通系统体系框架建立的交通组织本体,为组织之间的协作提供知识表示,包括交通管理中心、旅客运输部门、交通信息服务提供商、紧急事件管理部门、基础设施管理部门、货物运输服务提供者、政府执法部门。

4) 应用本体层 (application ontology) 应用本体是在领域本体的基础上建立的与具体系统应用相关的概念与关系集合,如车辆诊断维护本体、交通事故处理本体等。单一的应用本体均与一个特定的应用相关联,可以采用事先定义的方式存储于本体库中,也可利用本体编辑工具从领域本体中分离概念进行实时编辑。用户通过对应用本体中概念的选择、组合并输入 <属性—值> 对描述,完成对特定应用中异构资源的查询与整合,保证跨子系统应用的实现。

2 交通网格决策本体应用场景

2.1 四层交通决策本体模型的应用范围

基于 SUMO 的四层交通决策本体模型为基于网格的智能交通决策支持系统跨领域、多部门的协作奠定了知识基础,它

的应用能够提供以下传统 ITS 所难以提供的非平凡的决策支持功能。

1) 基于网格的数据集成 网格平台运用网络技术与通信协议为分散在各地、各个 ITS 子系统之间的异构数据库提供了一致的数据访问服务接口。通过网格中间件将远程异地、异构的数据库无缝地组织起来,可以建立起网格层次上的统一虚拟数据库,实现远程大规模异构数据的物理与结构集成,为分布式大规模数据挖掘提供了平台支持。

2) 基于语义的交通信息集成 通过对虚拟数据库上数据资源的本体概念化描述,使用语义查询语句对物理集成的数据资源进行语义查询与组合。同时,对大规模数据挖掘所得到的数据实例进行基于语义的二次集成,建立信息层次上的相互联系,为信息的有效共享提供保证。

3) 基于语义的交通服务集成 利用对用户的交通服务请求进行基于本体的服务查找及匹配处理,通过诱导的方式提供与用户需求相同或相似的交通服务来帮助用户进行服务的选择与交通工作流的定义,促进子系统内的服务集成并建立相应的语义案例库。

4) 基于语义的交通应用集成 基于网格的交通决策支持系统是一个跨领域的分布式决策支持系统。组织内各领域子系统通过 SUMO 这个一致的上层本体基础进行本体集间的映射,建立共同的知识库。利用共同知识库对各系统模块进行服务封装和语义描述,提供应用集成的语义服务配置与管理,实现跨领域系统的决策资源共享和协同调度工作。

2.2 基于语义的信息获取流程

基于语义的交通信息获取是交通网格信息集成的基础。采用符合 OGSA 规范的网格服务作为本体应用的统一操作单元,系统基于本体的信息获取及关联应用所涉及的网格服务流程定义如图 2 所示。

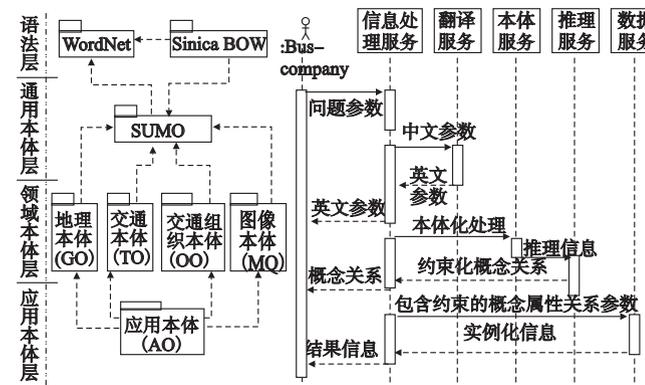


图1 交通网格决策资源本体模型

图2 基于本体的信息获取及关联应用流程

1) 问题提交 用户利用本体层次模型中应用本体所定义的和与系统功能相关联的概念与属性对应的 Sinica BOW 中文词库进行求解问题的描述。其中问题的描述采用 <属性—值> 对的形式进行表示,并提交给信息处理服务进行信息获取。

2) 信息处理服务 接收交通网格用户提供的中文问题参数,并通过对其他服务的调用完成信息获取功能。

3) 翻译服务 借助 Sinica BOW 的中英文概念互译模型对中文决策问题描述信息进行英文化处理。

4) 本体服务 对已英文化的决策问题参数进行相关本体文件的获取及概念的匹配。

5) 推理服务 对本体文件中涉及到问题参数的相关概念

进行基于知识库的推理,获取概念与属性实例间的约束关系。

6) 数据服务 利用本体实例查询语言结合参数选择的约束条件对与决策问题参数相关的属性值进行实例化查询与匹配,通过本体概念属性与各异构数据库字段之间的映射关系获取决策支持所需要的集成信息。

2.3 交通网格决策本体应用场景举例

城市交通应急系统是智能交通系统的重要组成部分,其主要功能是处理城市道路交通突发事件,包括重特大交通事故、严重的交通堵塞、道路及相邻区域的治安事件等。城市交通应急处理的过程通常需要多个部门的协同工作和统一调度完成,如公安交警的事故现场保护和应急救援,公安消防的消防救护、化学物品的应急抢救,卫生部门的医疗抢救等。但由于各道路交通事故应急系统分散于多个部门,各部门根据自身应急救援的特点建立了相对独立的应急系统体系,其信息和资源难以集成与共享,很大程度上限制了城市交通应急系统功效的发挥。利用本文采取的四层本体结构,可以在对遗留系统作很小改动的基础上,通过本体概念及关系的映射,将参与应急的多个子系统的信息进行有效的集成,最终得到所需要的信息以提供交通应急的决策支持。本文模拟了南方某市公交车辆在某路口发生特大交通事故的场景,其涉及到的部门子系统处理事件包括:

- a) 公交运营系统。增开公交车辆,就近选择维修站点进行车辆的拖运及维修。
- b) 122 接处警系统。调用路口监控录像进行交通事故分析及交通车辆疏导。
- c) 119 火警系统。确定是否需要增派消防提供车辆破拆等方面的帮助。
- d) 医疗急救系统。通过医院位置、医护人员、设备要求、当前医疗能力等信息选择就近医院进行事故人员抢救。
- e) 交通诱导系统。通过交通诱导减少车祸发生路段的车堵情况。

图 3 表示了基于本体的公交车祸处理实例中部分信息集成后的相互关系。其中,TO:trafficAccident 通过 TO:happensWhen 与车祸发生时间关联,通过 TO:happensWhere 与车祸发生地点关联;MO:camera 通过 TO:locationAt 与车祸地点关联,以提供交通监控视频信息;GO:hospital 与 GO:garage 则通过 GO:near 与车祸发生地点相关联,以提供最近医院和维修点信息。

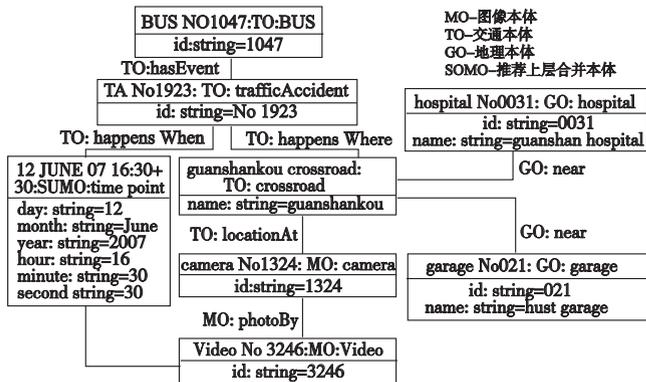


图3 基于本体的公交车祸处理实例信息关系(简化图)

用户采用 2.2 节提出的基于本体的信息获取流程,首先选择交通事故处理本体并利用用户界面进行概念属性的关联,建立起实例的事故处理信息模型,如通过“邻接”属性关联事故

点位置与医院、交警、火警、修理厂等,利用“摄制”属性表示事故与事故图片的关联关系;其次,对构建的事故处理信息模型进行已有信息(属性—值)对的初始化,包括事故点位置、事故损失情况、事故情形、事故环境等,并提交给信息处理服务进行信息处理;信息处理服务调用其他内部服务进行实例信息的获取;最后,参考知识库、规则库中对交通事故处理中相关问题的约束条件,如对事故人员伤亡所需医疗设施、医疗能力等信息的筛选条件,通过权重设置对最终查询所涉及的实例化信息进行排序,选出最优的事故处理方案供用户使用。相应的,也可根据信息对预案库及案例库进行基于语义的匹配与查找得到交通诱导方案支持。

3 结束语

本文阐述了在 ITS 系统中引入基于网格的资源共享与协同调度平台和基于本体的资源表示与描述方法的重要性,分析了目前交通本体构建及应用方法的不足,提出了基于 SUMO 的交通网格决策本体层次模型。通过引入通用上层本体 SUMO,构建了具有中英文对照及丰富词汇基础的交通网格决策资源本体集,解决了传统本体构建方法所造成的本体集成难度高、耗时长、不支持中文推理的问题。该框架具有良好的开放性和适用性,使得交通网格环境下分布式异构交通信息的访问和集成成为可能,也为网格环境下其他领域本体模型的构建提供了方法指导。

下阶段笔者将重心放到交通决策本体集的应用研究上,如基于本体的交通网格服务体系结构设计、基于本体的交通网格的案例表示与查询推理、基于本体的交通网格 workflow 设计等研究课题上,完善基于本体的交通网格系统结构与运作机理。

参考文献:

- [1] “九五”国家重点科技攻关项目《中国智能运输系统体系框架》专题组. 中国智能运输系统体系框架[M]. 北京:人民交通出版社, 2003:10-22.
- [2] 黄珂萍, 蒋昌俊. 基于本体的城市交通的知识分析和推理[J]. 计算机科学, 2007, 34(3):192-196.
- [3] 刘卫宁, 朱怡. ITS 虚拟共用信息平台的数据集成包装器[J]. 计算机工程与设计, 2006, 27(22):4197-4200.
- [4] 李阳, 翟军, 陈燕. 用本体实现智能交通系统的语义集成[J]. 信息技术, 2005, 29(6):10-13.
- [5] 马金素, 陈毛狗, 袁捷. 基于本体的问题求解模型在 CSP 中应用的研究[J]. 计算机应用与软件, 2006, 23(10):60-61, 70.
- [6] PEASE A, NILES I, LI J. The suggested upper merged ontology: a large ontology for the semantic Web and its applications[C]//Proc of Working Notes of the AAAI-2002 Workshop on Ontology and the Semantic Web. Edmonton, Comada: [s. n.], 2002.
- [7] WordNet [EB/OL]. (2004-12-18) [2007-10-20]. <http://www.cogsci.princeton.edu/~wn>.
- [8] HUANG Chu-ren, CHANG Ru-yng, LEE Shang-bin. Sinica BOW (bilingual ontological WordNet): integration of bilingual WordNet and SUMO [C]//Proc of the 4th International Conference on Language Resources and Evaluation(LREC2004). Lisbon Portugal: [s. n.], 2004:26-28.
- [9] TSINARAKI C, POLYOROS P, CHRISTODOULAKIS S. Interoperability support between MPEG-7/21 and OWL in DS-MIRF[J]. IEEE Trans on Knowledge and Data Engineering, 2006, 19(2):219-232.