

一种基于事件的大气污染突发事件本体模型

刘菲京,刘 炜,王 旭,刘宗田

(上海大学 计算机工程与科学学院,上海 200444)

摘 要:由于网络存储大气污染突发事件的异构性和不同人理解上的语义差异,正确地理解和描述大气污染突发事件知识对于大气污染突发事件的应急决策至关重要。文中提出一种基于事件的大气污染突发事件本体模型,该模型抽象出大气污染领域中最一般化的事件类以及相关的事件要素,建立事件之间的语义关系。利用本体建模工具 Protégé 构建出该模型,并用一个“工厂废气排放”的例子验证了该模型的可用性。研究结果表明,该模型可以很好地对大气污染突发事件的知识进行表示,支持事件间的语义推理。

关键词:大气污染突发事件;事件;事件本体;描述逻辑

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2015)06-0029-06

doi:10.3969/j.issn.1673-629X.2015.06.007

An Air Pollution Emergency Ontology Model Based on Event

LIU Fei-jing, LIU Wei, WANG Xu, LIU Zong-tian

(School of Computer Engineering and Science, Shanghai University,
Shanghai 200444, China)

Abstract: Because of the heterogeneousness of air pollution emergency stored by Internet and semantic differences of people in understanding, it is important to understand and describe air pollution emergency knowledge correctly in emergency decision-making. An air pollution emergency ontology model based on event is proposed. The model abstracts event classes in the field of air pollution, analyzes the semantic relations between them and establishes six elements of event classes. Air pollution emergency ontology model is constructed with the ontology modeling tool Protégé and a "Factory Emission" case is described to testify the validity of the model. The results show that the model can well represent the air pollution emergency knowledge, supporting semantic reasoning between events.

Key words: air pollution emergency; event; event ontology; description logic

0 引言

近年来,随着经济的快速增长和城镇化的大力推进,环境突发事件频繁发生,尤其是大气污染突发事件(如雾霾、工厂废气排放等)和由大气污染所引发的社会事件(如市民抢盐和疯抢口罩等),严重影响到人们的生活和社会的稳定。而互联网上有关这些事件的报道也越来越多,信息泛滥,数据异构,缺乏合理的知识表示和互操作性。因此,通过对互联网中大气污染突发事件信息进行有效地收集和处理,建立基于语义的大气污染突发事件知识库,可实现对互联网中大气污染突发事件信息的自动侦测、语义分析和智能检索,进而帮助环境监管机构和应急响应部门快速地做出应急决策。

本体技术的发展和应用于大气污染突发事件的语义分析与预警提供了一种新的途径。传统的本体技术可通过对大气污染突发事件对象进行分析,确定大气污染突发事件领域的概念以及概念之间的关系,进而实现概念之间的推理。但是,存在很多不足之处:

(1)概念离散问题,突发事件概念和突发事件的参与者、地点和时间不是作为一个有机整体存在,因此在构建本体过程中存在大量的概念分类工作;

(2)传统本体往往容易忽视突发事件的动态性,比如很难描述突发事件随时间变化所呈现出的不同状态;

(3)传统本体中,概念的关系单一,因此不能对事件之间的语义关系进行充分描述,造成了推理能力的

收稿日期:2014-08-10

修回日期:2014-11-11

网络出版时间:2015-05-06

基金项目:国家自然科学基金资助项目(61305053,61273328);上海市自然科学基金(12ZR1410900)

作者简介:刘菲京(1989-),女,硕士研究生,研究方向为事件本体形式化;刘 炜,副研究员,研究方向为语义本体、知识表示;刘宗田,教授,博士生导师,研究方向为人工智能、软件工程等。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20150506.1645.038.html>

不足。

所以文中提出了一种基于事件的大气污染突发事件本体模型,该模型充分考虑到了大气污染突发事件描述中所包含的动作、地点、对象、时间、事件发生的状态变化和文本语言的表现形式以及事件之间的语义关系。

1 相关工作

目前,对大气污染的研究多数是局限于对某个特定城市的大气污染的空气质量进行数据分析,很少研究大气污染突发事件以及这些突发事件所引发的相关事件,并且利用本体解决大气污染突发事件问题的研究也不多见。Ling 等^[1]以吉隆坡为例,分析了该城市的空气质量等级和居民健康状况,研究结果表明城市土地使用和交通极大地影响了吉隆坡的空气质量 and 居民的身体健康。Jaime Reyes 等^[2]利用基于遗传算法的聚类技术分析了墨西哥的空气质量数据,对聚类数据的分析可以得到一些关于空气质量的未知的信息,但是需要采集大量的数据。陈俏等^[3]将支持向量机应用于大气污染物浓度预测,提出了一种大气污染物浓度预测模型。Russo A 等^[4]提出了大气污染预测模型,该模型旨在识别和预测大气污染事件,对数据的可用性要求很高。景跃军等^[5]通过分析郑州、北京两地大气污染突发事件应急预案启动经验,得到大气污染突发事件驱动因素及突发事件演化模型。魏振钢等^[6]基于高斯扩散模式建立 1 个大气污染模型。王玉平等^[7]将特征-块中心差分法应用于一维大气污染模型中。邵超峰等^[8]在综合分析突发性大气污染事件的人体健康与应急管理研究的基础上,提出了突发性大气污染事件的环境风险评估与管理的研究方法。Chen Wenjun 等^[9]为了能够更好地处理大气污染突发事件,提出了基于地理信息系统(GIS)的技术解决方案,将 GIS 技术和大气污染扩散数学模型相结合,利用 GIS 技术来定位大气污染突发事件的位置,查询附近环境的敏感地带。蒋浩等^[10]为了能够智能化处理环境污染突发事件,提出了一套环境污染突发事件应急预案的数字化方法和执行机制。还有基于知识的方法和 MAS(Multi-Agent Systems)方法^[11]也提出了解决大气问题的途径,并且大大减小了构建领域知识的复杂性。Claudine Métral^[12]通过本体将空气质量模型(Air quality models)和 3D 城市模型(3D city models)集成起来,通过创建空气质量模型本体以及扩展 OUPP(Ontology of Urban Planning Process)中介本体,模拟城市系统,根据大气流动和化学方程式去评估城市的空气质量,帮助城市规划者更好地处理空气污染问题。Mihaela M. Oprea 等^[13]提出了一个空气污染本

体(AIR_POLLUTION_Onto),该本体在一定程度上分析和控制了空气污染,但是该本体更多的是一个概念集合,对于关系的描述也都是概念之间的,而且关系也很有限,缺乏动态性。

2 事件本体的相关概念

事件本体是共享的客观存在的事件类系统模型的明确的形式化规范说明^[14],由事件类集合(ECS)、事件类之间的关系(R)和规则(Rules)构成。文中将以刘宗田等^[14]提出的事件本体模型为基础,来构建大气污染突发事件本体模型。下面对文献[14]中的事件、事件类以及事件间的关系等概念做简要介绍。

2.1 事件和事件类的定义

定义 1 事件(Event):指在某个特定的时间和地点下发生的,由若干角色参与的,表现出若干动作特征的一件事情。形式上用 e 来表示,被定义为一个六元组结构:

$$e::=_{\text{def}}(A,O,T,V,P,L)$$

其中, A 表示动作; O 表示对象,包括参与事件的所有角色; T 为时间,表示事件发生的时间段,从事件发生的起点到事件结束的终点,当起点等于终点时,该时间段即为一个时间点; V 表示环境; P 为断言,表示事件发生过程中对象状态的集合,由事件发生的前置状态和后置状态组成; L 为语言表现,表示事件的语言规律,包括核心词集合、核心词搭配等。

定义 2 事件类(Event Class):指具有共同特征的事件集合,用 EC 表示,如下:

$$EC=(E,C_1,C_2,\dots,C_6)$$

$$C_i=\{c_{i1},c_{i2},\dots,c_{im},\dots\} \quad (1 \leq i \leq 6, m \geq 0)$$

其中, E 为事件的集合,称为事件类的外延; C_i 为事件类的内涵,表示每个事件在第 i 个要素上具有的共同特性集合, c_{im} 是事件类中每个事件在第 i 个要素上具有的一个共同特性。

事件是事件类的实例化,如“顺德凤翔工业区工厂废气排放”即为“工厂废气排放”的一个实例。

2.2 事件类关系

事件类之间的关系分为两大类:分类关系和非分类关系。分类关系又称为上下位关系或父子关系,用 isKindOf 表示。非分类关系主要研究以下四种:因果关系(cause)、组成关系(isComposedOf)、跟随关系(follow)和并发关系(concur)。其中 isKindOf 关系和 isComposedOf 比较容易混淆,这里用两个例子做个简单地解释:如“大气污染”isKindOf“环境污染”;“大气污染”事件类由“污染物排放”、“检测污染源”等事件类组成。cause 和 isComposedOf 是从逻辑上来判断的, follow 和 concur 主要是根据时间来判断的,时间上有

先后但不能间隔太久并且没有重叠的即为跟随关系(如“起床”和“刷牙”),时间上有重叠但不完全重叠的为并发关系(如“打雷”和“闪电”)。这些关系既存在于事件实例之间,也存在于事件类之间。

3 基于事件的大气污染突发事件本体模型的构建和应用

目前国内外关于基于事件的大气污染突发事件本体模型的研究并不多见,大部分都是着重于定义大气污染的静态概念以及概念之间的简单关系,很少关注大气污染中发生的事件以及事件之间的语义关系。文中提出了一种基于事件的大气污染突发事件本体模型,该模型能够正确理解和描述大气污染突发事件和事件发生的过程以及事件之间的语义关系,这对大气污染突发事件的应急决策至关重要。文中利用 Protégé 构建基于事件的大气污染突发事件本体模型,以 OWL DL 为研究基础,用该模型描述了一个“工厂

废气排放”的例子,并对其进行形式化表示。

3.1 基于事件的大气污染突发事件本体模型的构建

文中以刘宗田等^[14]提出的事件本体模型(Event Ontology Model, EOM)为上层事件本体模型,在此基础上构建大气污染突发事件本体模型,简称为 APECOM (Air Pollution Emergency Case Ontology Model)。形式上用 APECOM 表示,被定义成一个四元组结构:

$$\text{APECOM} ::= \text{def}(\text{AP_EventClasses}, \text{AP_Relations}, \text{AP_Individuals}, \text{AP_Rules})$$

其中, AP_EventClasses 表示大气污染突发事件的事件类集合; AP_Relations 表示大气污染突发事件里所有关系的集合; AP_Individuals 表示大气污染突发事件的实例集合; AP_Rules 表示大气污染突发事件规则的集合,用于事件间的推理。

文中通过本体构建工具 Protégé 来构建基于事件的大气污染突发事件本体模型部分截图,如图 1 所示。

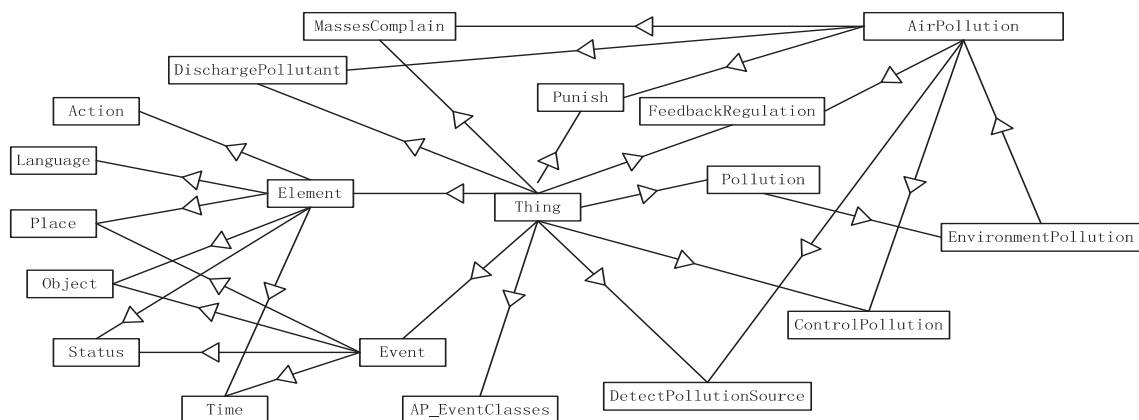


图1 基于事件的大气污染突发事件本体模型

1) 大气污染突发事件的事件类 AP_EventClasses。

大气污染突发事件的事件类又可分为动态事件类 AP_DynamicEventClasses 和静态概念类 AP_StaticConcepts,形式化定义如下:

$$\text{AP_EventClasses} ::=$$

$$\text{def}(\text{AP_DynamicEventClasses}, \text{AP_StaticConcepts})$$

动态事件类是基于事件的大气污染突发事件本体模型的核心,相比一般大气污染本体只注重对大气污染的分类,文中通过研究发现,大气污染中所发生的一系列事件,其事件与事件之间都有着紧密的语义关系,通过分析事件之间的语义关系可以体现出大气污染突发事件的发生过程和处理过程,这对大气污染突发事件的应急决策至关重要。通过六要素对事件的描述以及分析事件之间关系,应急决策人员可以了解大气污染突发事件的整个发生过程,并做出正确的处理。

动态事件类包含的核心事件类有:污染物排放、居

民生活环境恶化、提出污染预防方案、相关部门调查、检测污染源、监管污染情况、反馈监管情况、立案处罚、污染治理、检查污染治理情况和群众投诉等等。

静态概念类是大气污染突发事件中的相关概念的集合,核心的概念类如下:

(1) 大气污染突发事件的一般对象类 (AP_GeneralObj): 它是事件类六要素的对象要素类的子类,是大气污染突发事件的一般对象的集合,是除了事件参与者之外的对象,如工厂、废气等。

(2) 大气污染突发事件的参与者 (AP_Actor): 它是事件类六要素的对象要素类的子类,是大气污染突发事件的参与者,如工人、居民、相关部门组织等。

(3) 大气污染突发事件的时间 (AP_Time): 它可以分为瞬时时间 (Instant) 和段时间 (Interval),瞬时时间即为开始时间等于结束时间,它表示一个时间点,可以描述一个大气污染突发事件的任何一个时间点,如发生和结束。段时间即结束时间大于开始时间,它表示

一个时间段,描述了一个大气污染突发事件的发生过程。

(4)大气污染突发事件的地点 (AP_Place):它可以分为虚地点 (VirtualPlace) 和实地点 (RealPlace)。有些非传统意义上的地点是实地点表示不了的,这里将它定义为虚地点,如“某居民在当地论坛投诉大气污染情况”,“当地论坛”即为虚地点。

(5)大气污染的分类 (AirPollutions):根据研究和治理的需要,也有很多不同的分类,文中按照大气污染存在的状态分为颗粒污染物和气态污染物。颗粒污染物包括粉尘、烟、飞灰等等,气态污染物包括硫氧化物、氮氧化物、碳氧化物等等。

(6)污染源的分类 (PollutionSource):根据污染源的自发性划分,可以分为人类污染源和自然污染源。对人类污染源又可以进行划分,可以分为固定污染源(如化工厂)和移动污染源(如汽车和摩托车);自然污染源,如火山喷发等。

2)大气污染突发事件里所有关系的集合 AP_Relations。

大气污染突发事件里的关系包括事件与事件之间的关系,事件与其要素之间的关系,定义如下:

$$\text{AP_Relations} = \{ \text{AP_R}(c_1, c_2) \mid c_1, c_2 \in \text{AP_EventClasses} \text{ or } c_1 \in \text{AP_EventClasses} \text{ and } c_2 \in \text{Element} \text{ or } c_1, c_2 \in \text{AP_StaticConcepts} \}$$

(1)事件与事件之间的语义关系:主要存在分类关系 (isKindOf) 和非分类关系 (因果 (cause)、跟随 (follow)、并发 (concur) 和组成 (isComposedOf) 关系),通过对大气污染突发事件中事件之间关系的描述,可以体现大气污染突发事件中事件的变化过程。其中, isKindOf 和 isComposedOf 也可以出现在静态概念之间。

(2)事件与要素之间的关系:事件与其六要素之间也有着不同的所属关系, eom:hasGeneralObj, eom:hasActor, eom:atTime, eom:hasIntervalTime, eom:atRealPlace, eom:atVirtualPlace, eom:hasPreStatus, eom:hasPostStatus, eom:hasLanguage, eom:hasAction。

(3)文中使用 Protégé 构建该模型。为了便于构建这些关系,特将它们作为对象属性 (ObjectProperty) 进行构建。除了对象属性,该模型里还有一些数据类型 (DataTypes),如 xsd:float, xsd:int, xsd:string, 等等,那么要素和数据之间还有一种关系即为 hasValue。

3)大气污染突发事件的实例集合 AP_Individuals。大气污染突发事件的实例包括事件类的实例和要素类实例,定义如下:

$$\text{AP_Individuals} = \{ \text{Individual} \mid \text{Individual} \in \text{AP_EventClasses} \text{ or } \text{Individual} \in \text{Element} \}$$

4)大气污染突发事件的规则集合 AP_Rules。

AP_Rules 表示大气污染突发事件规则的集合,由逻辑语言表示,主要用于事件间的推理以及形式化表示,定义如下:

$$\text{AP_Rules} = \{ R; C_1 \times C_2 \times \dots \times C_n \rightarrow Q \mid C_1, C_2, \dots, C_n, Q \in \text{AP_EventClasses} \cup \text{AP_Relations} \}$$

其中, C_i 表示事件发生的条件; Q 表示事件发生的结果。

例 1:北京近日的 PM2.5 指数已超过 200,等级达到 5 级,属重度污染。

$$\exists x(\text{pollutant}(x) \wedge \text{moreThan200}(x) \wedge \text{hasLevel5}(x) \rightarrow \text{SeverePollution}(x))$$

文中利用 Protégé 构建大气污染突发事件本体模型之后会生成一个 owl 的文件,该文件对大气污染突发事件本体模型里的动态事件类、静态概念、关系和实例都有清晰的定义,使得该模型具有很好的扩展性,可根据应用需求再进行扩充,并且支持事件间的推理和数据的查询。下面对构建该模型的几个问题用部分核心代码进行解释说明:

(1)事件是文中模型的核心之一,由于事件是由六要素组成,而 Web 上的数据资源都是通过 RDF 三元组存储,为了能够表示出事件六元组,可以用六个三元组表示一个六元组,部分 owl 代码如下:

```
<!-- http://www.semanticweb.org/ellen/ontologies/2014/7/APECOM#Event -->
<owl:Class rdf:about="http://www.semanticweb.org/ellen/ontologies/2014/7/APECOM#Event">
  <owl:equivalentClass>
    <owl:Class>
      <owl:intersectionOf rdf:parseType="Collection">
        //对象要素
        <owl:Restriction>
          <owl:onProperty rdf:resource="http://www.semanticweb.org/ellen/ontologies/2014/7/APECOM#eom:hasObject"/>
          <owl:onClass rdf:resource="http://www.semanticweb.org/ellen/ontologies/2014/7/APECOM#Object"/>
          <owl:minQualifiedCardinality rdf:datatype="xsd:nonNegativeInteger">1</owl:minQualifiedCardinality>
        </owl:Restriction>
        .....
      </owl:intersectionOf>
    </owl:Class>
  </owl:equivalentClass>
</owl:Class>
```

(2)为了便于构建大气污染突发事件中的各种关系,特将关系作为对象属性 (ObjectProperty) 来构建,由于事件之间的关系、事件与要素之间的关系是不同的,所以通过 rdfs:domain 和 rdfs:range 对这两类关系进行

更一般的形式化表示,提高应急决策的处理效率。

表 1 “工厂废气排放”实例的形式化表示

形式化	解释
a. $\text{FactoryEmission} \subseteq \text{AirPollutionEmergency}$	工厂废气排放事件是大气污染突发事件的一个实例
b. $\text{Factory}(x) \wedge \text{WasteGases}(y) \wedge \text{Discharge}(x,y) \wedge \text{Residents}(z) \wedge \text{Inhale}(z,y) \rightarrow \text{hasAdverseReactions}(z)$	工厂排放废气,居民吸入废气,出现不良反应
c. $\text{Residents}(x) \wedge \text{hasDizzy}(x) \wedge \text{hasNausea}(x) \wedge \text{hasCough}(x) \wedge \text{Factory}(y) \rightarrow \text{Complain}(x,y)$	居民有头晕、恶心、咳嗽等不良反应,对工厂进行投诉
d. $\text{EPA}(x) \wedge \text{Local}(y) \wedge \text{TestAQ}(x,y) \rightarrow \text{Results}(z)$	环保部门对当地的空气质量进行检测,得出检测结果
e. $\text{AQI}(x) \wedge \text{moreThan200}(x) \wedge \text{Standard}(y) \wedge \text{Exceed}(x,y) \rightarrow \text{SeverePollution}(z)$	由 d 的结果可知,AQI 超过 200,超过标准值,属重度污染
f. $\text{LocalGov}(x) \wedge \text{Factory}(y) \wedge \text{SeverePollution}(z) \rightarrow \text{Investigate}(x,y)$	LocalGov(x) 表示当地政府,由 b,c,d,e 可知,当地政府要对该工厂进行调查

注:EPA,the Environmental Protection Department,环保部门;AQ,Air Quality,空气质量;AQI,Air Quality Index,空气质量指数

4 结束语

文中采用本体构建工具 Protégé 构建了一种基于事件的大气污染突发事件本体模型,该模型有完整和一致的基于事件的概念和清晰的语义,可用于大气污染突发事件的知识表示和推理,并用一个“工厂废气排放”的实例证明了该模型的有效性。

在未来的工作当中,会继续丰富该模型,并且利用同样的方法去构建其他的环境突发事件本体模型,如水污染突发事件本体模型、土壤污染突发事件本体模型等等。不同类型的环境突发事件本体在事件类、事件要素以及事件关系的构建上有很强的相似性,通过对这些事件类、事件要素以及相关的概念进行抽象,构建具有一般性的环境突发事件本体模式。

参考文献:

[1] Ling H L O,Ting K H,Shaharuddin A,et al. Air quality and human health in urban settlement:case study of Kuala Lumpur city[C]//Proc of international conference on science and social research. Kuala Lumpur:IEEE,2010:510-515.

[2] Reyes J,Sánchez A. Analysis of air quality data in Mexico city with clustering techniques based on genetic algorithms[C]//Proc of international conference on electronics,communications and computing. Cholulu:IEEE,2013:27-31.

[3] 陈 俏,曹根牛,陈 柳. 支持向量机应用于大气污染物浓度预测[J]. 计算机技术与发展,2010,20(1):250-252.

[4] Russo A,Soares A O. Hybrid model for urban air pollution forecasting:a stochastic spatio-temporal approach[J]. Mathe-

matical Geosciences,2014,46(1):75-93.

[5] 景跃军,张 昀,李 元. 大气污染突发事件演化过程及应急管理探讨[J]. 环境保护,2014,42(11):52-53.

[6] 魏振钢,郭遵强,张 琳,等. 基于高斯模式的大气污染模型的应用[J]. 中国海洋大学学报:自然科学版,2008,38(2):327-330.

[7] 王玉平,张志跃. 一维大气污染模型的特征-块中心差分法[J]. 计算力学学报,2011,28(5):693-698.

[8] 邵超峰,鞠美庭,张裕芬,等. 突发性大气污染事件的环境风险评估与管理[J]. 环境科学与技术,2009,32(6):200-205.

[9] Chen W J,Chen S Z. Application of GIS technology in the emergency monitoring of sudden air pollution accident[C]//Proc of international conference on information science and engineering. [s. l.]:[s. n.],2010:3550-3555.

[10] 蒋 浩,荆 玲,武港山. 突发环境污染事件应急预案数字化方法研究[J]. 计算机技术与发展,2013,23(5):6-9.

[11] Oprea M. A case study of knowledge modelling in an air pollution control decision support system[J]. AI Communications,2005,18(4):293-303.

[12] Métal C,Falquet G,Karatsas K. Ontologies for the integration of air quality models and 3D city models[C]//Proc of 2nd workshop on conceptual models for urban practitioners. Bologna,Italy:Società Editrice Esculapio,2008:27-42.

[13] Oprea M M. AIR_POLLUTION_Onto:an ontology for air pollution analysis and control[M]//Artificial intelligence applications and innovations III. US:Springer,2009:135-143.

[14] 刘宗田,黄美丽,周 文,等. 面向事件的本体研究[J]. 计算机科学,2009,36(11):189-192.