

DOI:10.13203/j.whugis20150378



文章编号:1671-8860(2018)01-0010-07

顾及信誉的众源时空数据模型

周晓光¹ 赵肄江^{1,2} 李光强¹ 张攀¹

1 中南大学地球科学与信息物理学院,湖南长沙,410083

2 湖南科技大学知识处理与网络化制造湖南省普通高校重点实验室,湖南湘潭,411201

摘要:提出了一种顾及信誉度的众源时空数据模型。在分析众源时空数据中地理要素、目标状态、对象版本、贡献者、信誉度、改变现实空间实体或信息空间对象状态的事件等要素间的相互作用机理的基础上,采用面向对象方法设计了一种顾及信誉度的众源时空数据组织方法,用UML对其进行描述,分析了与信誉度相关操作及其联动关系,得出了8条联动规则。开发了顾及信誉度的众源时空数据管理原型系统,验证了所提模型的有效性。

关键词:众源数据;信誉度;时空数据模型;版本;操作

中图法分类号:P208

文献标志码:A

众源数据或志愿者地理信息(volunteered geographic information, VGI)具有覆盖范围广、免费共享、准实时等特点^[1-2],正迅速成为专业地理信息的重要补充^[1, 3]。开放街道地图(OpenStreetMap, OSM)目前已在基于网络的地理信息服务、应急救援等领域得到广泛应用。

目前,学术界对众源数据的研究工作主要集中于VGI应用、数据质量与可靠性评价、数据存储管理与索引等方面^[1, 4]。与传统时空数据由专业人员按照严格的程序规范进行采集与更新^[5]不同,众源数据由大众志愿者自发标报,其可信度具有诸多不确定性,且贡献者的信誉对众源数据质量有重要影响^[6-7]。文献[6]认为贡献者的非专业性是VGI数据质量问题的重要原因。文献[7]认为贡献者在对地理事物的认知上有很大的区别,有必要建立针对地理信息贡献者的评价模型。文献[8]提出信息信任概念,将信任作为众源数据质量评价的一种手段。文献[9]通过众源数据的版本数、用户数等来计算目标可信度。文献[10]提出了一种基于版本相似度的众源数据贡献者信誉度评价方法。总之,与信誉相关的数据在众源数据质量控制中起着举足轻重的作用,故需建立顾及信誉的众源时空数据模型。

目前在众源数据模型方面的理论研究成果较

少。在众源数据收集平台方面,OSM采用关系数据库存储点、线和关系数据及其属性,当前数据和历史数据分别存储,未对信誉相关数据进行有效组织管理;谷歌地图和WikiMapia均采用社会等级机制对志愿者贡献目标进行质量控制^[2],其用户级别主要根据贡献量确定,但由于贡献量与贡献目标质量间的不对等关系,其数据组织仍不能满足众源数据可信度评价需求。

时空数据模型研究方面已取得丰富的研究成果,按照模型侧重点的不同,可将时空数据模型分为^[11]:①侧重记录实体状态的时空数据模型,如时空快照、时空立方体、非第一范式时空数据模型等^[12-13];②侧重时空目标变化过程和因果关系的模型,如基于事件的时空数据模型、面向事件的版本化模型、基于事件的因果关系模型等^[14-16];③侧重时空目标和时空关系描述,如面向对象的时空数据模型等^[17];④侧重移动目标的实时数据模型,如实时地理信息系统(Geographic Information System, GIS)时空数据模型等^[18-19]。上述时空数据模型各具特点,已经在相应行业中得到应用^[18]。但是,上述模型都是针对专业地理信息建立,由于专业地理信息不存在与信誉相关的要素,都不能满足众源数据信誉计算与服务需求。

总之,现有时空数据模型不能有效表达用户

收稿日期:2015-10-14

项目资助:国家自然科学基金(41371366)。

第一作者:周晓光,博士,教授,研究方向为时空数据建模、更新与服务。zxcgsu@foxmail.com

通讯作者:赵肄江,博士,讲师。zhaoyijiang@163.com

和目标的信誉信息,不能提供根据信誉选择目标版本等与信誉相关的信息查询与服务。因此,本文借鉴传统时空数据模型的建模方法,在分析众源数据形成机制、信誉要素及其相互关系的基础上提出了一种顾及信誉度的众源时空数据模型,设计了信誉相关操作及其联动关系,并通过实验分析了该模型的有效性。

1 众源数据要素及其相互关系分析

1.1 众源空间数据形成机制分析

在专业地理信息中,数据库对象版本与现实世界目标状态一一对应,数据更新的原因就是目标状态发生了变化。但在众源数据环境下,数据更新的原因主要包括目标修正和目标变化两类。前者对应目标状态没有变化的修改,即贡献者在地理目标本身没有变化的情况下,为提升目标质量等而进行修改;后者对应于目标状态发生变化的修改。在这种存在大量修改的细粒度时间框架下,地理目标的同一状态可能包含多个版本,传统时空数据模型难以有效处理这种情况。如图 1 描述了多个志愿者共同编辑某地理目标的同一状态,形成多个目标版本的过程,图 1 中贡献者 c_g 创建了该状态的第一个版本 v_1 ,用户 c_{j-2} 通过编辑版本 v_{i-2} 得到版本 v_{i-1} , c_{j-1} 、 c_j 、 c_k 通过编辑 v_{i-1} 分别得到版本 v_{i0} 、 v_{i1} 、 v_{i2} , c_n 通过编辑 v_{i1} 得到版本 v_m 。值得注意的是,图 1 仅描述了一个状态内的编辑演化过程,在实际情况下,同一地理目标可能包含多个状态,而且每个状态都可能包含多个版本。

文献[20]对 OSM 中地理目标编辑历史的研究表明,目标的最终版本不一定是最好版本,如图 1 所示同一目标状态可能存在 m 个版本, v_m 为最终版本但不一定为最佳版本。因此,如何选择目标的最佳版本是 VGI 的一个重要问题,多位学者认为目标可信度可以作为衡量目标质量的依据[8-9],可对同一目标状态的各个版本进行可信度评价,取可信度最高的版本作为该状态的选用版本。

贡献者信誉对众源数据可信度有着重要影响,因此,在众源时空数据模型中应该有效组织影响贡献者信誉的相应要素。目前在众源数据环境中,用户之间的评价包括显式评价和隐式评价两类。显式评价,即用户采取直接打分方式对其他用户贡献对象直接评分,如淘宝中对商品的评价打分;隐式评价,即用户不直接评分,而是采用其

他相关信息推断出用户的隐含评分。事实上,众源数据的编辑过程中蕴含了用户之间的隐式评价[10],如图 1 中,该地理目标的某一状态包含多个版本,由于同一状态下的目标本身并未发生变化,所以版本 v_{i-2} 与 v_1 间的相似程度隐含了版本 v_{i-2} 的编辑者对版本 v_1 的支持程度,具体表现为相似程度越大,支持度越大,反之亦然。所以,版本 v_{i-2} 与 v_1 间的相似程度隐含了版本 v_{i-2} 的贡献者对 c_g 的信誉评价。同理, c_{j-1} 通过版本 v_{i-1} 隐式评价了 c_g 和 c_{j-2} , c_n 通过其参与贡献版本 v_m 隐式评价了之前的其他版本的贡献者,即 c_g 、 c_{j-2} 、 c_{j-1} 、 c_j 、 c_k 。根据用户之间的隐式评价,可以计算出用户之间的评价价值,然后通过加权计算可得出贡献者的信誉。

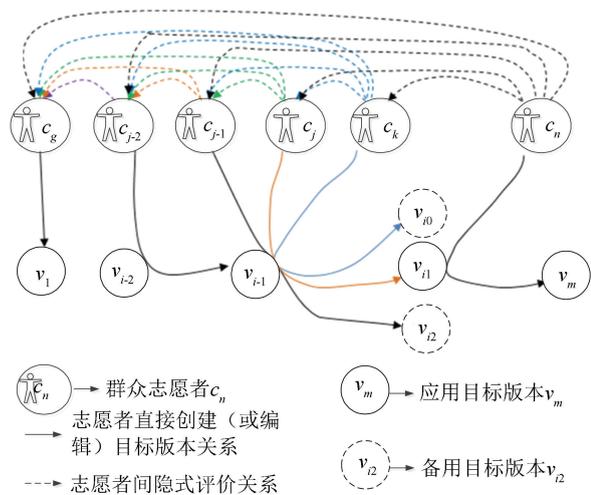


图 1 众源数据形成机制

Fig.1 Formation Mechanism of Crowdsourcing Data

从众源空间数据的形成过程可看出,空间目标、目标状态、地理事件(即导致目标或目标状态变化的原因)是众源空间数据与专业 GIS 系统的共有要素,贡献者、目标状态版本、贡献者信誉、版本信誉、状态版本变化事件则是众源空间数据系统所特有的信誉要素。

1.2 信誉要素及其相互关系分析

众源空间数据由普通大众自发贡献,贡献过程中贡献者信誉和目标信誉在不断演化。分析清楚信誉要素间的相互关系是建立顾及信誉的众源时空数据组织及信誉操作联动机制的基础。

众源时空数据系统要素主要包括空间目标、目标状态、目标版本、地理事件、贡献者、贡献者信誉、版本信誉、信息事件(即状态版本变化事件)。

贡献者主要指大众用户,是众源空间数据系统的重要组成部分。借鉴电子商务等相关众源数据系统方法,在每个用户注册后即可拥有根据注

册信息计算获得的初始信誉;用户在系统中贡献地理对象后,其他用户对他的贡献可进行直接或间接评价,即可形成评价信誉。因此,贡献者信誉包括初始信誉和评价信誉等。

在传统时空数据模型中,事件往往是空间对象状态变化的原因,对象状态、版本与地理变化事件具有一一对应关系。在众源地理信息环境中,由于一个对象状态可能对应多个版本,每个版本对应一个贡献者的一次“新建/修改/删除”事件。因此,众源地理信息环境中包含地理事件和信息事件两类事件。地理事件对应于现实空间中真实存在的事件,这类事件改变其影响的地理要素的真实状态;信息事件对应于信息世界中的“新建/修改/删除”事件,在众源数据环境中,表现为用户通过对目标版本的修改以改变其在信息空间中的状态,但现实世界中空间目标的真实状态并没有改变。众源时空数据系统核心要素及其相互关系如图2所示。

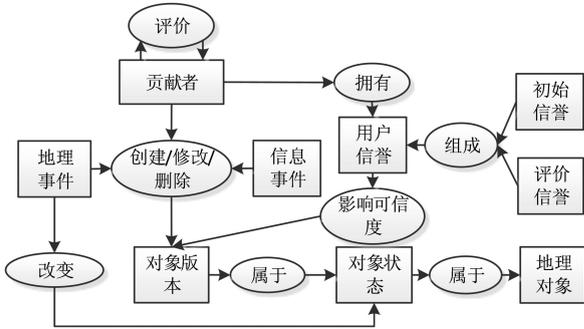


图2 众源数据信誉相关要素及关系
Fig.2 Reputation Related VGI Elements and Their Relationships

2 顾及信誉的众源时空数据组织

统一建模语言(unified modeling language, UML)具有简单、统一的特点,已成为可视化建模语言的工业标准,因此,本文选用UML对模型进行描述。为突出重点,图3仅给出了该模型的UML简图。

地理对象类用于描述地理对象的生命周期可用元组(编号 ID , 地理对象类型编码 L , 开始时间 T_s , 结束时间 T_e 等)表达;对象状态类用于描述对象状态的存在周期,可用元组(编号 ID , 地理对象编号 G , 状态选用版本号 S , 开始时间 T_s , 结束时间等 T_e)表达;对象版本类用于描述地理对象的数据库编辑记录,可用元组(编号 ID , 地理对象编号 G , 状态编号 S , 空间特征 K , 专题特征

A , 贡献者 C , 版本编号 V , 开始时间 T_s , 结束时间 T_e 等)表达;当前状态版本类用于描述所有当前地理对象最新状态信息,可用元组(编号 ID , 地理对象编号 G , 状态编号 S , 版本号 V , 是否最佳版本 V_b 等)表达。

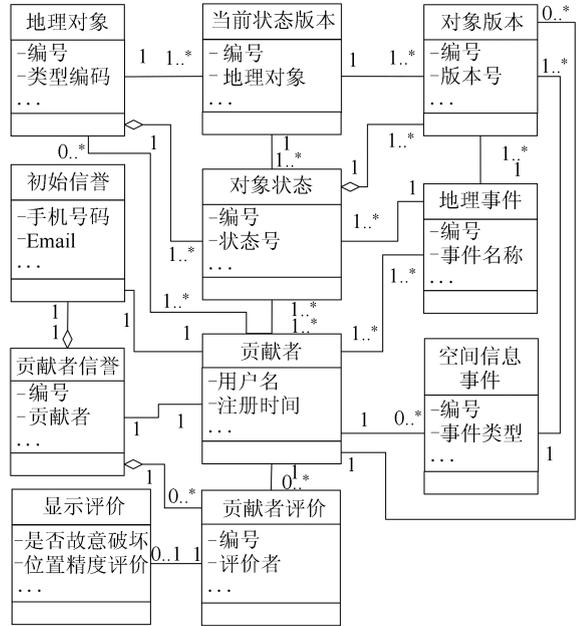


图3 顾及信誉度的众源时空数据UML简图
Fig.3 UML Diagram of VGI Model Considering Trust Degree

贡献者类用于描述贡献者的注册信息,可用元组(编号 ID , 用户名 N , 密码 P , 注册时间 T , 是否注销 X , 是否临时停用 B 等)表达;贡献者评价类用于描述贡献者间的评价,可用元组(编号 ID , 被评价者 C_1 , 评价者 C_2 , 被评价版本 V_1 , 评价版本 V_2 , 目标类型 L_g , 评价类型 L_a , 评价时间 T , 评价值 Z 等)表达;显式评价类用于描述贡献者对众源数据的显式评价,可用元组(编号 ID , 是否为故意破坏 H , 是否为错误对象 E , 位置精度评分 W , 语义精度评分 Y 等)表达;初始信誉类用于描述与贡献者信誉影响要素及初始值,可用元组(编号 ID , 邮箱 M , 电话 D , 真实姓名 N , 身份证号 I , 其他信息 F , 信息验证 J , 初始信誉值 R , 时间 T 等)表达;贡献者综合信誉类用于描述贡献者综合信誉,可用元组(编号 ID , 贡献者 C , 信誉值 R , 开始时间 T_s , 结束时间 T_e 等)表达,信息空间事件类用于描述信息空间事件,可用元组(编号 ID , 事件类型 L_e , 贡献者 C , 目标版本号 V , 时间 T 等)表达;地理事件类用于描述真实地理事件,可用元组(编号 ID , 事件名称 N , 事件类型 L_e , 描述 I , 贡献者 C , 目标版本号 V , 时间 T 等)表达。各类的形式化表达及主要操作见表1。

表 1 类的形式化表达和操作

Tab.1 Formalization and Operations of Class

类名	英文名	形式化表达	主要操作
地理对象类	GeoObject	(ID, L, T_s, T_e)	增加/获取/删除地理对象, 移库
对象状态类	State	(ID, G, S, T_s, T_e)	增加/获取/删除对象状态, 移库
对象版本类	Version	$(ID, G, S, K, A, C, V, T_s, T_e)$	增加/删除/获取对象版本, 获取对象状态, 版本可信度计算, 移库
当前状态版本类	CurrStateVer	(ID, G, S, V, V_b)	增加/更新/删除状态版本
贡献者类	Contributor	(ID, N, P, T, X, B)	增加/查询/注销/阻挡志愿者, 移库
贡献者评价类	UserEvaluate	$(ID, C_1, C_2, V_1, V_2, L_g, L_a, T, Z)$	增加/获取/删除评价, 移库
显式评价类	ExplicitEvaluation	(ID, H, E, W, Y)	增加/修改/获取/删除显式评价, 移库
初始信誉类	InitRep	$(ID, M, D, N, I, F, J, R, T)$	增加/修改/获取初始信誉, 贡献者信息验证, 初始信誉计算, 移库
贡献者综合信誉类	Reputation	(ID, C, R, T_s, T_e)	增加/更新/获取贡献者信誉, 移库
信息空间事件类	InfoEvent	(ID, L_e, C, V, T)	增加/查询事件, 获取事件相关地理对象, 移库
地理事件类	GeoEvent	(ID, N, L_e, I, C, V, T)	增加/查询事件, 获取事件地理对象, 移库

3 信誉度操作及其联动关系

数据操作是数据模型的重要组成部分, 在顾及信誉度的众源时空数据模型中, 对象版本可信度、贡献者信誉度和对象编辑过程之间的联动变化关系复杂, 为了自动维护信誉要素间的一致性, 需要建立信誉度操作间的联动关系。

3.1 信誉度操作

信誉操作包括初始信誉类中的新增初始信誉 $addInitRep()$ 、更新初始信誉 $updateInitRep()$ 、获取初始信誉 $getInitRep()$ 、验证用户信息 $validateUserInfo()$; 贡献者评价信誉类中的增加评价 $addEvaluation()$ 、获取评价 $getEvaluation()$; 贡献者综合信誉类中的增加信誉 $addReputation()$ 、更新信誉 $updateReputation()$ 、获取信誉 $getReputation()$; 对象版本类中的可信度计算 $computeTrustDegree()$ 、更新可信度 $updateTrustDegree()$ 、获取可信度 $getTrustDegree()$ 等操作。

3.2 信誉操作联动规则

为了确保顾及信誉度的时空数据系统顺利运行, 维护顾及信誉的时空数据一致性, 在用户注册、新建目标版本、评价目标版本等操作过程中需要自动驱动相应的联动操作, 为此需要设计信誉操作联动规则。假设目标 A 包含 n 个版本 (n 为大于 0 的整数), 按照生成时间顺序依次为 $v_1, v_2, v_3, \dots, v_n$, 版本 $v_i (1 < i \leq n)$ 的贡献者为用户 c_i 。以用户 c_i 贡献目标 A 的版本 v_i 为例, 信誉度操作间的主要联动关系如图 4 所示。

为提高信誉度联动操作的可靠性和鲁棒性, 将联动关系可形式化成如下规则。

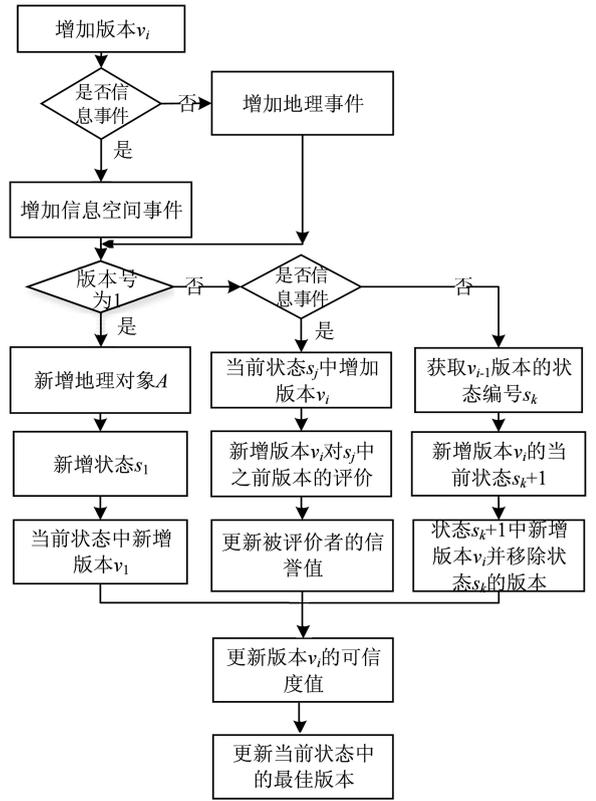


图 4 信誉度主要联动操作

Fig.4 Linkage Operations of Reputation

规则 1 当贡献者 c_i 注册时, 在 $InitRep$ 和 $Reputation$ 中分别增加 c_i 对象。

规则 2 当贡献者 c_i 贡献某目标 A 的版本 v_i , 且对应的事件类型为信息空间事件时, 则添加一个信息空间事件, 否则添加一个地理事件。

规则 3 当贡献者 c_i 新增目标 A 的版本 v_i , 其联动操作包括增加地理对象、增加对象状态、设置

版本可信度和在 CurrStateVer 中增加当前版本。

规则 4 当对地理目标进行修改,且事件类型不为信息空间事件,则对应该目标在现实世界中的状态变化,相应联动操作包括新增目标状态,更新目标版本的可信度,同时在 CurrStateVer 中增加当前版本并删除上一状态的所有版本。

规则 5 当对地理目标版本修改,且事件类型为信息空间事件,则对前述同一状态版本进行评价,也即版本 v_i 对之前版本 v_k 进行评价,然后更新该版本可信度值,最后在当前状态版本中增加版本 v_i 。

规则 6 当产生版本 v_j 对 v_i 的隐式评价或用户 c 对 v_i 的显式评价,则评价并更新版本 v_i 的贡献者 c_i 的信誉。

规则 7 当用户 c_i 的初始信誉信息有变更,则更新该贡献者的信誉。

规则 8 当版本 v_i 的可信度值更新,则当前状态的最佳版本可能产生了变化,需重新设定当前状态的最佳版本。

4 众源时空数据模型实验与分析

为了验证本文顾及信誉度的众源时空数据模型的有效性,作者设计了相应的数据结构,其中众源数据的目标版本存储借鉴了 OSM 的数据存储方式,构建了顾及信誉度的众源时空数据管理原型系统。由于众源数据采集周期长,本文采用 OSM 中的 Berlin 市历史数据集作为实验数据,首先将编辑历史数据(包括历史修改版本)导入系统,然后根据目标的整个编辑过程计算出参与贡献的用户信誉和目标可信度,由于实验数据中没有显式评价,用户评价根据隐式评价确定^[10]。目标版本可信度的计算根据前一版本的可信度、当前版本的用户信誉值、版本之间的修改比例进行计算。在实验系统中,可查询地理对象当前状态和历史状态的快照,可以选择目标需要显示的状态和版本,并可对同一对象的任意两个版本之间进行相似度的比较和计算;在实验系统中,还可以根据某个对象的任意版本查询其贡献者的贡献情况,如图 5 中 Berlin 某游泳池对象(对象 ID: 4839528)的 7 个版本的贡献者信誉和贡献目标数信息;另外还可以查看与设置用户信誉、回溯用户贡献的任意目标的各版本情况,可根据信誉度选择最可信版本作为采用版本等。

从实验结果可以看出,顾及信誉度的众源时空数据模型能够有效地存储和处理用户信誉、目

标可信度等相关信息,并能对用户的贡献情况进行统计分析等。与传统时空数据模型相比,本模型主要增加的部分为贡献者、贡献者初始信誉、贡献者评价、贡献者信誉的存储。存储量的增加主要集中在贡献者评价表。本模型中,贡献者评价表中的字段均采用数值型记录,每条记录的存储占 40 B,存储 1 000 000 条记录约占 38.15 MB;根据 OSM 中柏林市的实际历史数据采用 SQL Server 2008 的存储情况,面目标的版本数共计 489 386 条,占用存储 196.22 MB,也即假设存储 1 000 000 条面目标的版本数据,约需要 400.95 MB 的存储空间。假如目标记录数和贡献者评价记录数据为 1:1,则贡献者评价记录表和面目标记录表存储量占比约为 9.51%,考虑到数据库中还需存储其他数据,增加贡献者评价记录表后数据存储量的增加应小于该比例。由此可知,所增加的存储量在可控范围之内,并通过定期移库操作,确保了模型的实用性。



图 5 地理对象各版本浏览

Fig.5 Versions of a Geographic Object

5 结语

在 VGI 环境中,贡献者信誉对空间目标的质量有着重要影响,需要对贡献者进行信誉建模和评价;另外,空间对象的同一状态对应多个目标版本,多个版本之间需要可信度评价,以获取信誉值最高的目标版本。为了有效存储管理贡献者信誉信息,计算目标版本的可信度,本文提出了一种顾及信誉度的众源时空数据模型,用 UML 图描述了众源地理信息系统中的要素及其相互关系,分析了与信誉度相关的操作及其联动关系;使用“对

象-状态-版本”的分层管理方式,将目标状态与其最可信版本联系起来,能更有效管理众源数据中存在大量对象版本的情况;通过实验验证本文模型的有效性。

应该指出的是,影响众源数据可信度(或质量)的因素很多,包括贡献用户的诚信度、专业水平、用户间的相互评价、目标本身质量及其与周边目标间关系的协调性等。本文实验系统信誉计算与信誉相关操作及其联动作用等均主要考虑了根据用户编辑修改形成的间接评价信誉,尚未考虑根据目标本身质量及其与周边目标间关系的协调性来计算目标信誉的方法及其对于信誉相关操作的影响。后续工作将关注顾及目标本身质量及其与周边目标间关系的协调性的目标信誉度计算方法及其对于信誉联动操作的影响。

参 考 文 献

- [1] Li Deren, Qian Xinlin. A Brief Introduction of Data Management for Volunteered Geographic Information[J]. *Geomatics and Information Science of Wuhan University*, 2010, 35(4): 379-383(李德仁,钱新林. 浅论自发地理信息的数据管理[J]. 武汉大学学报·信息科学版, 2010, 35(4): 379-383)
- [2] Goodchild M F, Li L. Assuring the Quality of Volunteered Geographic Information[J]. *Spatial Statistics*, 2012, 1: 110-120
- [3] Qian Xinlin. Research on the Representation and Management of Geospatial Data from Volunteered Geographic Information[D]. Wuhan: Wuhan University, 2011(钱新林. 面向自发地理信息的空间数据表达与管理方法研究[D]. 武汉: 武汉大学, 2011)
- [4] Wang Ming, Li Qingquan, Hu Qingwu, et al. Quality Analysis on Crowd Sourcing Geographic Data with Open Street Map Data[J]. *Geomatics and Information Science of Wuhan University*, 2013, 38(12): 1 490-1 494(王明,李清泉,胡庆武,等. 面向众源开放街道地图空间数据的质量评价方法[J]. 武汉大学学报·信息科学版, 2013, 38(12): 1 490-1 494)
- [5] Chen Jun, Wang Donghua, Shan Yaoling, et al. Master Design and Technical Development for National 1 : 50 000 Topographic Data-base Updating Engineering in China[J]. *Acta Geodaetica et Cartographica Sinica*, 2010, 39(1): 7-10(陈军,王东华,商瑶玲,等. 国家 1 : 50 000 数据库更新工程总体设计研究与技术创新[J]. 测绘学报, 2010, 39(1): 7-10)
- [6] Shan Jie, Qin Kun, Huang Changqing, et al. Methods of Crowd Sourcing Geographic Data Processing and Analysis[J]. *Geomatics and Information Science of Wuhan University*, 2014, 39(4): 390-396(单杰,秦昆,黄长青,等. 众源地理数据处理与分析方法探讨[J]. 武汉大学学报·信息科学版, 2014, 39(4): 390-396)
- [7] Grira J, Bédard Y, Roche S. Spatial Data Uncertainty in the VGI World: Going from Consumer to Producer[J]. *Geomatica*, 2010, 64(1): 61-72
- [8] Bishr M, Mantelas L. A Trust and Reputation Model for Filtering and Classifying Knowledge About Urban Growth[J]. *GeoJournal*, 2008, 72(3-4): 229-237
- [9] Keßler C, de Groot R T A. Trust as a Proxy Measure for the Quality of Volunteered Geographic Information in the Case of OpenStreetMap[M]//Geographic Information Science at the Heart of Europe. Berlin, Germany: Springer, 2013
- [10] Zhao Yijiang, Zhou Xiaoguang. Version Similarity-based Model for Volunteers' Reputation of Volunteered Geographic Information[J]. *Acta Eodaetica et Cartographica Sinica*, 2015, 44(5): 578-584(赵肄江,周晓光. 地理信息志愿者信誉度评估的版本相似度模型——以面目标为例[J]. 测绘学报, 2015, 44(5): 578-584)
- [11] Zhou Xiaoguang. Incremental Updating of Cadastral Database [M]. Beijing: Surveying and Mapping Press, 2007(周晓光. 地籍数据库增量更新[M]. 北京: 测绘出版社, 2007)
- [12] Langran G. A Review of Temporal Database Research and its Use in GIS Applications[J]. *International Journal of Geographical Information System*, 1989, 3(3): 215-232
- [13] Shu Hong. A Unification of Gail Langran's Spatio-temporal Data Models[J]. *Geomatics and Information Science of Wuhan University*, 2007, 32(8): 723-726(舒红. Gail Langran 时空数据模型的统一[J]. 武汉大学学报·信息科学版, 2007, 32(8): 723-726)
- [14] Jiang Jie, Chen Jun. Event-Based Spatio-Temporal Database Design for Land Subdivision System[J]. *Acta Geodaetica et Cartographica Sinica*, 2000, 29(1): 64-70(蒋捷,陈军. 基于事件的土地划拨时空数据库若干思考[J]. 测绘学报, 2000, 29(1): 64-70)
- [15] Xue Cunjin, Zhou Chenghu, Su Fenzhen, et al. Research on Process-Oriented Spatio-Temporal Data Model[J]. *Acta Geodaetica et Cartographica Sinica*, 2010, 39(1): 95-101(薛存金,周成虎,苏奋振,等. 面向过程的时空数据模型研究[J]. 测绘学报, 2010, 39(1): 95-101)

- [16] Meng Lingkui, Zhao Chunyu, Lin Zhiyong, et al. Research and Implementation of Spatio-Temporal Data Model Based on Time-Varying Sequence of Geographical Events[J]. *Geomatics and Information Science of Wuhan University*, 2003, 28(2): 202-207(孟令奎, 赵春宇, 林志勇, 等. 基于地理事件时变序列的时空数据模型研究与实现[J]. 武汉大学学报·信息科学版, 2003, 28(2): 202-207)
- [17] Lei Qihong, Liu Yaolin, Yin Zhangcai, et al. Spatio-Temporal Data Model in Cadastral Information System Based on Weighted Graph[J]. *Geomatics and Information Science of Wuhan University*, 2006, 31(7): 640-642, 652(雷起宏, 刘耀林, 尹章才, 等. 基于加权图的地籍时空数据描述模型研究[J]. 武汉大学学报·信息科学版, 2006, 31(7): 640-642, 652)
- [18] Gong Jianya, Li Xiaolong, Wu Huayi. Spatio-Temporal Data Model for Real-Time GIS[J]. *Acta Geodaetica et Cartographica Sinica*, 2014, 43(3): 226-232(龚健雅, 李小龙, 吴华意. 实时 GIS 时空数据模型[J]. 测绘学报, 2014, 43(3): 226-232)
- [19] Zhang Hengcai, Lu Feng, Chen Jie. An Integrated Data Model for Spatial-Temporal Trajectories and the Social Relationships of Moving Objects[J]. *Geomatics and Information Science of Wuhan University*, 2014, 39(6): 711-718(张恒才, 陆锋, 陈洁. 移动对象时空轨迹及社交关系一体化数据模型[J]. 武汉大学学报·信息科学版, 2014, 39(6): 711-718)
- [20] Mooney P, Corcoran P. Characteristics of Heavily Edited Objects in OpenStreetMap[J]. *Future Internet*, 2012, 4(1): 285-305

Crowdsourcing Spatio-Temporal Data Model Considering Reputation

ZHOU Xiaoguang¹ ZHAO Yijiang^{1,2} LI Guangqiang¹ ZHANG Pan¹

¹ School of Geosciences and Info-physics, Central South University, Changsha 410083, China

² Key Laboratory of Knowledge Processing and Networked Manufacturing, Hunan University of Science and Technology, Xiangtan 411201, China

Abstract: Crowdsourcing data are contributed by non-professionals incorporating new properties such as the contributor's reputation, and degree of trust in the contributed geographic objects. Furthermore, a crowdsourced geographic object usually has multiple versions when it is modified by several volunteers, so a mechanism for evaluating the reputation of a contributor is an alternative way to select the most creditable version. These new issues cannot be expressed and processed in traditional spatio-temporal data models. Therefore, a new crowdsourcing spatio-temporal data model is proposed in this paper, which takes reputation into consideration. The main elements in crowdsourcing data, e. g., geographic object, object status, object version, contributor, reputation, and evenst that change the state of an entity in the real world or the object in information system and their interaction mechanisms were analyzed. An object-oriented approach was used to design a crowdsourcing data model, and a UML diagram of this crowdsourcing spatio-temporal data model considering degree of trust is presented. The reputation related operations and their linkage relationships were analyzed, and eight reputation linkage operation rules were established. A prototype system using this crowdsourcing spatio-temporal data model was developed to verify the effectiveness of the model.

Key words: crowdsourcing data; trust degree; spatio-temporal data model; version; operation

First author: ZHOU Xiaoguang, PhD, professor, specializes in spatio-temporal data modelling, updating and service. E-mail: zxgcsu@foxmail.com

Corresponding author: ZHAO Yijiang, PhD, lecturer, specializes in volunteered geographic information. E-mail: zhaoyijiang@163.com

Foundation support: The National Natural Science Foundation of China, No. 41371366.