

存算分离的原位传感器观测接入方法^①

袁 赛, 陈能成, 肖长江, 杜文英, 王 凯

(武汉大学 测绘遥感信息工程国家重点实验室, 武汉 430070)

摘 要: 原位传感器是智慧城市建设的重要数据来源, 其在城市资源协调、灾害预警、动态监测分析等领域发挥决定性作用. 当前传感器观测接入方法未考虑传感器数据的流式特征, 无统一的接入模型, 导致无法统一过滤特定时空场景下的观测结果, 传感器接入组件可复用性差. 本文提出一种存算分离的原位传感器观测接入方法, 以站点为中心的传感器统一接入模型为基础, 将流式处理框架下的原位传感器观测接入过程分为数据获取、观测过滤与观测存储三个部分. 实验证明该方法能够基于传感器统一接入模型有效接入网络结构异构的原位传感器站点, 并实现对多个原位传感器观测结果特定时间、空间场景下的属性过滤.

关键词: 原位传感器; 观测接入; 存算分离; 流式处理; 过滤

引用格式: 袁赛, 陈能成, 肖长江, 杜文英, 王凯. 存算分离的原位传感器观测接入方法. 计算机系统应用, 2017, 26(7): 90-96. <http://www.c-s-a.org.cn/1003-3254/5874.html>

Storage-Computation Separation Method for In-Situ Sensor Observation Access

YUAN Sai, CHEN Neng-Cheng, XIAO Chang-Jiang, DU Wen-Ying, WANG Kai

(State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China)

Abstract: In-situ sensor is an important data source of intelligent urban construction, which plays a decisive role in urban resource coordination, disaster early warning, dynamic monitoring analysis and other fields. The present sensor observation access method does not consider the flow characteristics of sensor data with no uniform access model resulting in a failure of unified filtering of time and space under the observation and poor reusability of the sensor access components. In this paper, an in-situ sensor observation access method based on storage and computation separation is proposed. Upon the station-based observation access uniform model, the in-situ sensor access process is divided into three parts: data acquisition, observation filtering and observation storage. The experimental results show that the method can effectively access the heterogeneous in-situ sensor station based on the sensor observation access uniform model and realize the attribute filtering in the specific time and space scenes of the multiple in-situ sensor observation results.

Key words: in-situ sensor; observation access; storage-computation separation; stream processing; filter

随着地理空间传感网^[1]的不断发展, 以度量值为观测结果的标量观测原位传感器不断增多, 包括RFID传感器、温度、风力、风向传感器等. 作为城市资源协调、灾害预警、动态监测分析的保障, 传感器观测数据的信息有效性是决策分析的关键. 因而研究原位传感器观测接入对动态获取原位传感器观测结果, 提高

特定时空场景下的观测质量, 实时发布传感器观测服务具有重要作用.

开放地理空间联盟(Open Geospatial Consortium, OGC)在传感网络整合框架(Sensor Web Enablement, SWE)中提出了传感器观测服务(Sensor Observation Service, SOS)^[2], 以传感器建模语言(Sensor Model

① 基金项目: 湖北省自然科学基金创新群体(2016CFA003); 国家自然科学基金(41301441)

收稿时间: 2016-11-21; 收到修改稿时间: 2017-01-04

Language, SensorML)^[3]和观测与测量(Observation and Measurement, O&M)^[4]标准实现异构传感器规则观测数据的存储与发布,而没有规范底层传感器数据流到SWE框架接入过程.针对SWE框架下的传感器观测接入问题,当前的主要解决思路是通过创建映射规则进行协议转换^[5,6],将底层协议转换成SWE协议. Broering等人^[7]在此基础上进行了扩展,提出了传感器总线概念. Karl等人^[8]提出一种GSN方式的传感器接入中间件与传感网连接方法.在传感器观测质量提升方面,张明哲等人^[9]将嵌入式RFID中间件数据过滤模型分为属性、冗余、行为和规则四个部分;易叶青等人^[10]研究了传感器虚假数据过滤问题.上述传感器观测接入研究集中于SWE框架和非SWE框架两个方面的扩展,主要以适配器的方式接入传感器单元,但在接入过程中有三个问题未进行考虑:1)对于异构的传感器接入单元,未能提出统一的传感器接入模型;2)传感器观测数据是一种典型的时间序列数据,具有流式数据实时性、无边性和瞬时性^[11],上述研究未考虑如何在流式数据处理框架下接入传感器观测数据;3)当前传感器观测过滤以观测值为过滤对象,对时间、空间过滤条件涉及较少,未形成一套完整的过滤方法与机制.

综上所述,针对原位传感器观测接入,考虑传感器的异构性与观测数据的流式特征,本文提出一种存算分离的原位传感器观测接入方法,在流式数据处理框架下,以原位传感器站点为单元,将原位传感器接入过程分为数据接收、观测过滤和观测存储三个部分,实现原位传感器网络单元观测数据的体系化过滤接入及实时存储发布.

1 原位传感器观测接入设计原理

原位传感器具有网络结构异构的特征,即相同通信协议下的传感器网络组成的传感器数量和类型不同.同时,传统的传感器观测过滤,忽略时间与空间属性,仅对传感器观测值过滤,因而不同地理场景下的事件监测、预警难以展开.考虑传感器观测数据的异构特征以及过滤计算需求,本文在传感器接入阶段将传感器观测数据结构统一化,将传感器数据接收、观测过滤和观测存储相分离,三者相互独立,依靠传感器观测数据模型衔接,进而解决异构传感器观测接入问题和传感器统一过滤.原位传感器观测接入框架如图1所示,传感器观测接入过程涉及传感器网络层、接入拓扑层

及传感器观测服务层.传感器网络层代表原位传感器实体网络结构,通过原位传感器平台实现与接入拓扑层的数据通信.同时,需要对传感器网络层进行建模并注册到传感器观测服务中;接入拓扑层的接入拓扑单元分为数据接收、观测过滤和观测存储三个部分,它们由传感器流式接入单元配置参数构建成一个传感器接入拓扑单元;原位观测服务层支持传感器网络模型的注册以及拓扑层接入的传感器观测结果的实时存储与发布.接入拓扑层将传感器网络层与观测服务层衔接起来,进而实现原位传感器观测数据的实时接入与发布.存算分离的原位传感器观测接入方法,提高了传感器观测计算和存储部分的复用性.异构的传感器观测接入,只需要重构数据获取组件,直接复用传感器观测过滤与存储组件.

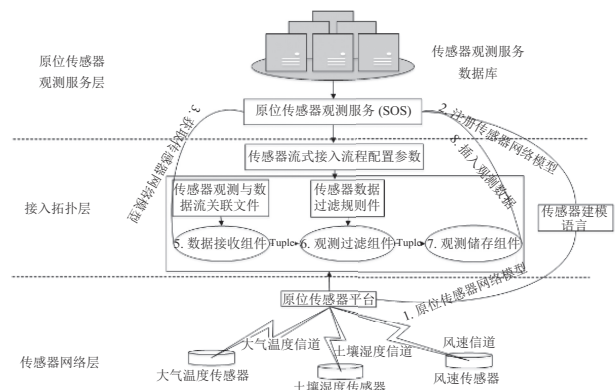


图1 原位传感器观测接入框架

2 存算分离的观测接入方法

基于原位传感器观测接入中的存算分离思想,传感器接入拓扑单元依据传感器流式接入流程配置参数,将数据接收组件、观测过滤组件与观测存储组件有机地结合,形成一个原位传感器观测接入单元.

2.1 数据接收组件

数据接收组件处理流程主要有构建原位传感器统一接入模型、传感器数据流获取与解析以及传感器观测数据发送.

原位传感器统一接入模型以传感器站点为接入单元,将传感器地理空间位置、观测时间和观测属性等信息组织在传感器站点下.如图2所示,传感器站点标识以StationID表示,其下可包含任意传感器单元;传感器单元由传感器标识(SensorID)、传感器位置(position)、传感器观测时间(ObservationTime)以及任

意见观测属性组成; 传感器观测属性则包含传感器观测属性标识(PropertyID)、传感器观测属性名称(PropertyName)、传感器观测属性单位(PropertyUnit)以及传感器观测值(Value)。原位传感器站点及其包含的传感器使用传感器建模语言建模, 并将该模型注册到传感器观测服务中, 数据接收组件在明确接入传感器站点标识后, 解析构造原位传感器统一接入模型, 为原位传感器数据匹配做准备。

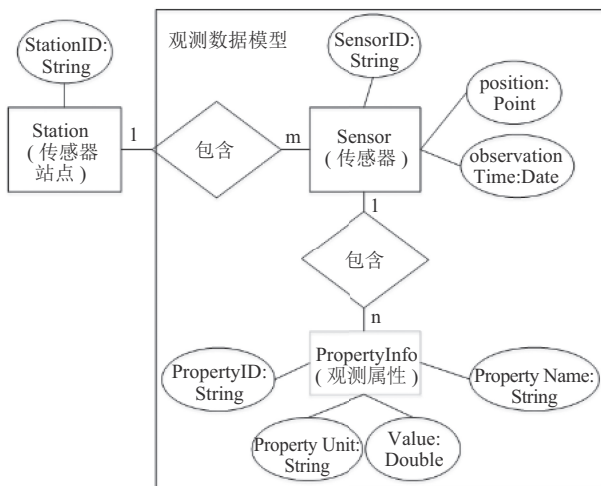


图2 原位传感器统一接入模型数据结构

为实现观测过程, 数据获取组件部分定义了传感器的数据传输协议与传感器数据表达等有关数据接入的元数据信息, 用于解析观测数据流。本文以传感器数据协议规则件描述上述内容, 定义传感器传输协议实现类, 由于数据表达是针对数据流与传感器观测属性的映射关系的描述, 所以需要定义一组数组。传感器的观测属性由传感器标识SensorID和传感器观测属性标识(PropertyID)唯一确定, 对于观测属性固定的二进制观测数据流, 数组结构为[SensorID, PropertyID, DataPosition, DataLength, DataRatio], 其中DataPosition指当前观测属性在数据流中的位置, DataLength指观测属性在数据流中的长度, DataRatio是观测数据转化为标准单位的比率。

原位传感器统一接入模型, 在匹配获得各观测属性观测值后, 转化为传感器接入单元内部传输的观测数据模型。如图2所示, 该模型由传感器自身属性、观测属性和观测结果三部分共同构成。

2.2 观测过滤组件

传感器观测过滤组件负责过滤已经接入的传感器

观测数据。

传感器观测是空间和时间相关的地理现象表达。时间、空间与观测属性是过滤的三个重要内容, 由于传感器观测数据存在同质性与异质性, 既存在观测相同地理对象的传感器, 又存在观测的地理对象类型不同传感器。传感器观测过滤须定义一致的过滤对象数据结构。考虑传感器观测是以空间位置为特征的观测, 时间及观测属性都可以看作其属性, 该数据组织方式为空间数据的Feature数据结构, 因而构建一种Feature特征下的传感器观测数据结构作为其统一观测过滤对象。如图3所示, 统一过滤对象中position为观测位置、sensorID为传感器标识、observationTime为观测时间, 其中传感器标识与观测属性标识唯一确定一组观测对象。

position (x, y)	Properties					
	sensor ID	observation Time	Property ID1	Property ID2	...	Property IDm
	ID	Time	Value1	Value2	...	Valuem

图3 传感器观测Feature型数据结构

传感器观测过滤是对特定时空场景下的观测属性过滤。过滤规则件以时间、空间过滤操作描述当前目标时空场景, 当接入非目标观测属性时, 观测数据不过滤; 而接入观测数据中含有目标观测属性, 且目标观测属性的观测时间、空间包含于过滤规则件规定的时空场景时, 观测数据被过滤筛选; 当目标观测属性不在规定的时空场景范围时, 观测数据无需过滤。图4为观测过滤流程图, 具体如下:

Step 1. 输入观测数据。

Step 2. 将观测数据转化为图3所示的Feature型数据结构;

Step 3. 输入观测过滤规则件, 并判断当前观测过滤规则件是否包含属性过滤操作, 若包含, 则输入该过滤规则件; 否则, 判定为错误规则件, 过滤退出。

Step 4. 判断当前观测数据观测属性是否包含在过滤规则件属性过滤中, 若是, 则转Step 5; 否则, 不过滤, 直接发送当前观测数据。

Step 5. 判断当前观测数据时间是否满足时间过滤操作, 若是, 则转Step 6; 否则, 判定为非目标空间场景数据, 并发送当前观测数据。

Step 6. 判断当前观测数据空间位置是否满足空间过滤操作, 若是, 则转Step 7; 否则, 判定为非目标空间场景数据, 并发送当前观测数据。

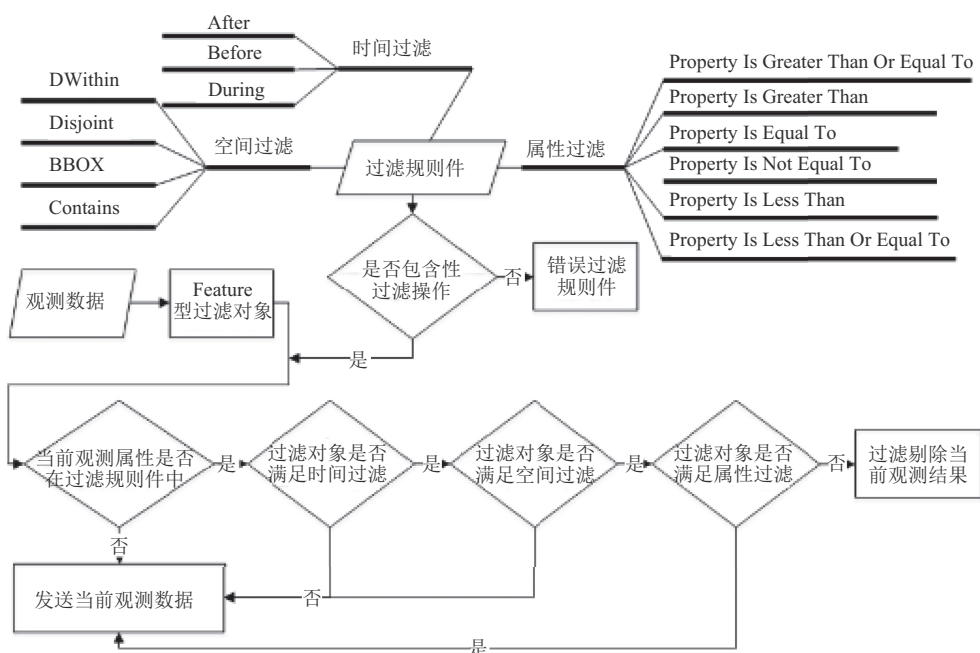


图4 原位传感器观测过滤方法

Step 7. 判断当前观测数据观测属性值是否满足属性过滤操作, 若是, 判定为目标场景需要的观测数据, 并发送当前观测数据; 否则, 判定为目标场景冗余数据, 剔除该观测数据, 不发送。

过滤操作主要有空间过滤、时间过滤和属性过滤三个方面, 三者之间的逻辑操作当前支持AND操作, 若同时对多个观测属性过滤, 观测属性过滤操作之间以逻辑操作OR衔接。时间过滤是针对传感器观测时刻的过滤, 因而只需包含对时刻点的过滤方法, 有在某个时刻之后、在某个时刻之前以及在某个时间段内; 空间过滤是针对原位传感器观测属性的空间采样位置的过滤, 因而过滤对象是空间点, 主要有距离范围内、包含、不包含以及在矩形查询; 属性过滤部分是对观测属性值的过滤, 支持的过滤方法有小于、大于、等于、不等于、小于或等于以及大于或等于。

2.3 观测存储组件

传感器存储组件将异构传感器观测数据以传感器观测数据编码方式统一描述, 调用发布传感器观测服务操作, 将传感器观测数据实时发布到传感器观测服务中。

2.4 传感器接入单元配置与构建

传感器数据接收、观测过滤及观测存储组件是相互独立的, 为构建一个有机的接入拓扑单元, 需明确各

组件的输入参数以及具体实现类。由于传感器观测的异构性特征明显, 传感器接入方法不可一成不变, 所以使用配置文档为依据的接入拓扑单元构建方法, 在传感器接入单元配置参数中声明传感器数据接收、观测过滤以及观测存储的实现类。原位传感器观测单元众多, 本文以站点为中心接入传感器观测数据, 为区分传感器观测接入拓扑单元, 在传感器接入配置参数中以传感器站点标识(stationID)作为接入拓扑标识以及配置文件标识, 各组件部分配置包括组件标识, 组件实现类名称, 组件任务数以及组件参数文档位置。基于此, 生成传感器观测接入拓扑单元。

3 实例分析

3.1 实验场景

实验包含原位传感器建模与观测和传感器接入拓扑两部分。在传感器建模与观测服务方面, 本文在SWE传感网整合框架下, 以SensorML1.0传感器建模语言描述原位传感器站点及传感器单元, 并以SOS 1.0.1中的RegisterSensor操作注册传感器站点单元信息。以Filter 2.0过滤标准实现传感器观测过滤。以O&M 1.0观测预测标准描述传感器观测数据并以SOS中的InsertObservation操作插入到传感器观测服务中。在传感器接入拓扑方面, 原位传感器接入拓扑单元使用Storm 1.0.0框架, 以java语言实现传感器接入各部分组

件,并以XML接入单元配置文档组合各组件,生成接入拓扑单元,并上传到Storm运行平台中。

本文实验选取湖北省武汉市豹澥地区三个原位传感器站点,水文监测站、气象土壤监测站1和气象土壤监测站2。如表1所示,气象土壤监测站1和气象土壤监测站2为通类型观测站点,但地理位置不同,包含风速、风向、雨量、土壤湿度、土壤温度、大气湿度、大气温度、大气压力8种观测对象;水文监测站包含水位、水温、水PH值3种观测对象。气象观测站与水文观测站传感器数量和类型不同,两者组成的传感器网络

单元异构。图5为三个传感器站点的空间分布。

实验中,三个原位传感器站点在接入过程使用相同的观测过滤条件。时间段在以格林威治时间转到本地时间下的2016-09-21T01:00:00.00+0800到2016-09-21T05:00:00.00+0800之间空间范围在(30.46970, 30.46967, 114.52677, 114.52675)矩形范围内,即图9中矩形框P所示范围,空间内只包含气象土壤观测站1;过滤观测对象是属性标识为“urn:ogc:def:property:OGC:1.0:rainfall”的降雨量,设置只保留在上述时空范围内降雨量大于0 mm的观测结果。

表1 原位传感器站点信息表

传感器站点	空间位置	数据表达				
		传感器	观测属性	位置	长度	比率
气象土壤监测站1	(30.46969, 114.52676)	风速传感器	风速	3	2	0.1
		风向传感器	风向	5	2	1
		雨量传感器	雨量	7	2	0.1
		土壤湿度传感器	土壤湿度	19	2	0.1
气象土壤监测站2	(30.46966, 114.52574)	土壤温度传感器	土壤温度	21	2	0.1
		大气温度传感器	大气温度	25	2	0.1
		大气湿度传感器	大气湿度	27	2	0.1
		大气压力传感器	大气压力	31	2	1
水文监测站	(30.46913, 114.52529)	水位传感器	水位	33	2	0.01
		水温传感器	水温	23	2	0.001
		PH值传感器	水PH值	29	2	0.1

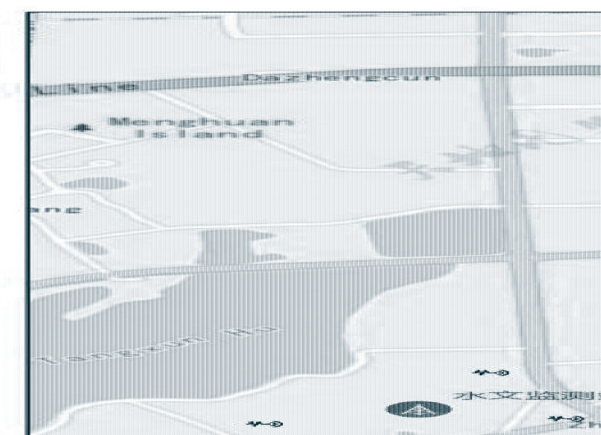


图5 原位传感器站点空间分布示意图

3.2 实验结果分析

3.2.1 接入网络结构异构的原位传感器站点

如图6所示,图6(a)为水文监测站中水PH传感器的大气PH观测结果,图6(b)为气象土壤监测站1的大气温度传感器的大气温度观测结果,因而该方法下使用传感器

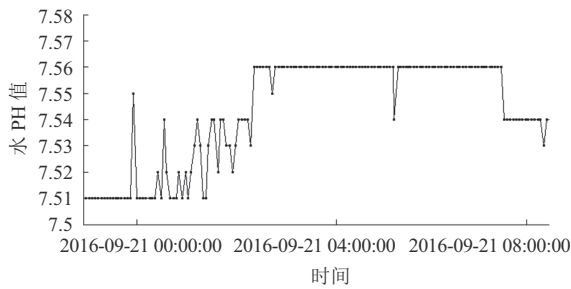
统一接入模型能够有效接入网络结构异构的原位传感器站点。

3.2.2 实现特定空间场景过滤

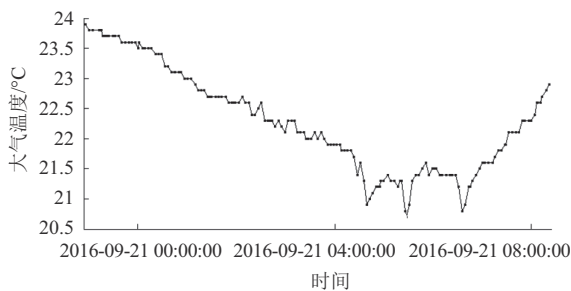
有效的空间场景过滤。与图6(c)气象土壤监测站1的降雨量观测结果不同,图6(b)气象土壤监测站1的大气温度观测在2016-09-21T01:00:00.00+0800到2016-09-21T05:00:00.00+0800时间范围内存在观测值,因而该过滤方法能够有效过滤同一观测站点下的不同观测属性。

有效的空间过滤。图6(e)为气象土壤监测站2的降雨量观测接入结果,其在2016-09-21T01:00:00.00+0800到2016-09-21T05:00:00.00+0800时间范围内存在降雨观测值。尽管该传感器站点与气象土壤监测站1具有相同的降雨量观测属性,但由于其空间范围不在过矩形区P范围内,所以该站点的降雨量观测未进行过滤。

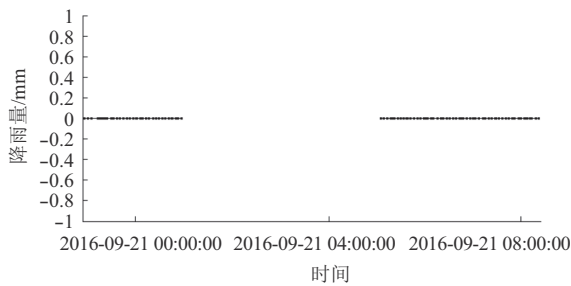
同一站点或不同站点下的不同传感器观测过滤互不影响。图6(c)本身来看,不在过滤时间范围内的时刻具有观测属性,图6(a)(b)(d)在过滤时间段内的不同观测属性未受影响。



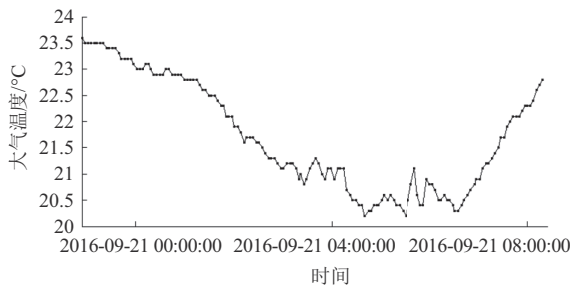
(a) 水文监测站水 PH 值观测接入图



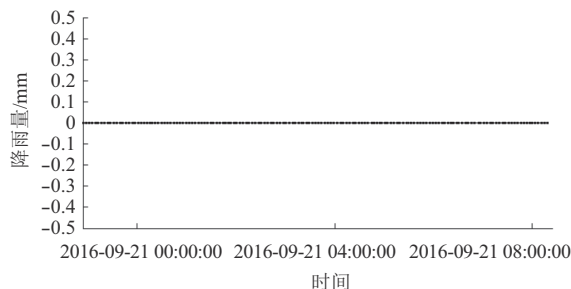
(b) 气象监测站 1 大气温度观测接入图



(c) 气象监测站 1 降雨量观测接入图



(d) 气象监测站 2 大气温度观测接入图



(e) 气象监测站 2 降雨量观测接入图

图6 原位传感器站点部分传感器观测过滤接入结果图

3.2.3 系统负载测试

本文采用模拟传感器站点方式测试该方法系统负载情况,其中每个传感器站点包含十个传感器单元,观测数据观测时间间隔为30秒,传感器对象以SOS服务方式发布.系统负载测试环境为浪潮NF5270M3 台式电脑,处理器为英特尔Xeon E5-2650 0@2 GHz八核,内存DDR3 1333 HZ 32 G.

图7为传感器观测接入数量不断增加系统负载情况,包含内存和CPU占用百分比.初始系统CPU使用率在0%左右,内存使用率在51%左右,当系统不断增加传感器数量时,内存与CPU使用率在不断上升,直到加入传感器观测数量达到500个时,CPU使用率接近100%,内存使用在之后再继续进行传感器接入,则传感器接入单元CPU可用资源变少,传感器观测接入可运行但系统超负载.因而,本方法的传感器观测接入在当前情况下,传感器单元接入时间间隔为30秒时,单机接入在400到500个传感器,能最大利用CPU与内存资源.

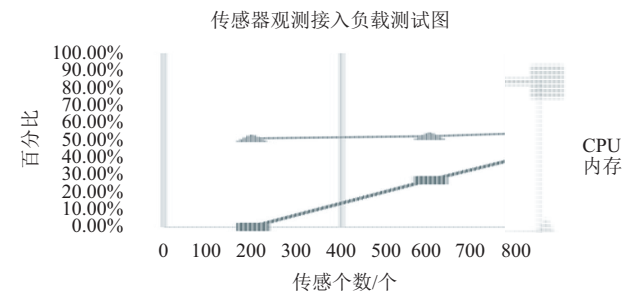


图7 传感器观测接入系统负载测试图

综上,该观测过滤方法,能够实现不同传感器观测接入拓扑单元的相同过滤条件过滤,通过修改接入单元配置文件也可以实现各传感器观测单元不同过滤条件过滤.该过滤方法可有效过滤掉特定时间、空间场景下的特定观测属性,而不影响其他时空场景下的观测数据的接入过程.此外,在单机模式下有效支持多个传感器的观测接入,但由于计算资源有限,单机观测接入数量不宜过多,本文情况下控制在400-500个之间.

4 结果与展望

从实验结果来看,针对观测属性位置固定的原位传感器,以本文提出的存算分离的原位传感器观测接入方法,在并以SWE 1.0传感网整合框架与Storm流式处理框架实验环境下,以传感器统一接入模型为基础,

复用观测过滤及存储组件,可实现异构的传感器网络单元观测接入.除考虑传感器观测属性值的过滤外,提出传感器观测时间、空间的过滤规则,扩展传统的原位传感器观测过滤内容,明确不同时空场景下的传感器观测过滤方法.与当前的传感器观测接入方法相比,本文提出的存算分离的原位传感器观测方法采用流式数据处理框架,传感器观测接入各模块相分离,异构传感器接入单元可复用过滤与存储组件.与当前仅考虑观测值过滤的方法不同,本文定义了传感器观测接入特定时空场景下观测过滤流程,解决了异构传感器的同条件过滤问题,实现了传感器的特定时空场景观测接入.

综合考虑传感器观测的接入问题,当前主要是针对原位传感器的异构性与过滤问题进行设计考虑,而随着大数据时代的来临,当前的传感器数量巨大,多传感器的高并发量,这会引发服务器接入压力问题.在接下来的研究中考虑从多传感器问题的角度出发,考虑服务器负载均衡问题,解决传感器接入的并发问题,这对构建物联网平台具有重要作用.

参考文献

- 1 陈能成, 杨训亮, 王晓蕾. 地理空间传感网信息公共服务平台的设计与实现. 地球信息科学学报, 2013, 15(6): 887-894, 917.
- 2 Bröring A, Stasch C, Echterhoff J. OGC OpenGIS sensor observation service interface standard. OpenGIS Implementation Specification OGC, 2007.
- 3 Botts M, Robin A. OpenGIS sensor model language (SensorML) implementation specification. OpenGIS Implementation Specification OGC, 2007.
- 4 Cox SJD. OpenGIS observation and measurements. OpenGIS Implementation Specification OGC, 2007.
- 5 蔡永娟, 沈春山, 吴仲城, 等. 基于SensorML和IEEE 1451的新型传感器网络标准研究. 仪表技术, 2010, (1): 1-3, 6.
- 6 Hu P Z, Robinson R, Indulka J. Sensor standards: Overview and experiences. Proc. 3rd International Conference on Intelligent Sensors, Sensor Networks and Information. Melbourne, QLD, Australia. 2007. 485-490.
- 7 Broering A, Foerster T, Jirka S, *et al.* Sensor bus: An intermediary layer for linking geosensors and the sensor web. Proc. 1st International Conference and Exhibition on Computing for Geospatial Research & Application. Washington, DC, USA. 2010. 1-8.
- 8 Aberer K, Hauswirth M, Salehi A. The global sensor networks middleware for efficient and flexible deployment and interconnection of sensor networks. Epfl, 2006, (12): 1-21.
- 9 张明哲, 张强, 袁巍, 等. 嵌入式RFID中间件数据过滤模型研究. 计算机工程与设计, 2010, 31(17): 3743-3746.
- 10 易叶青, 林亚平, 彭舸, 等. 无线传感器网络中不依赖MAC认证的虚假数据过滤算法. 通信学报, 2009, 30(6): 53-63.
- 11 丁维龙, 赵卓峰, 韩燕波. Storm: 大数据流式计算及应用实践. 北京: 电子工业出版社, 2015: 15-16.