

面向物联网的异构实体间关系服务模型

李冠宇, 安敬民[†]

(大连海事大学 信息科学技术学院, 辽宁 大连 116026)

摘要:物联网中的实体大多以异构形式存在,解决异构问题是推进物联网发展的重要因素。为此,在分析了实体及实体上载到网络中信息的特点后,基于 Agent 提出一种异构实体关系服务模型。该模型具有实体信息数据处理和构建实体服务逻辑两种功能,利用提出的全新的实体物间动态关系计算公式将异构数据转化为物间关系紧密度(值),并依此生成了 Agent 服务序列,结合二分图理论实现物联网环境下的 Agent 服务逻辑;屏蔽了原本因实体和其信息的异构而使得其无法正常交互通信的缺点,实体通过模型中与之对应的 Agent 进行交互,并在已构建的服务逻辑下支配 Agent;最后对该模型进行了仿真实验,结果表明其可满足物联网环境中的用户服务需求,在数据更新和服务准确率及冲突的处理上要优于传统的本体语义技术对异构问题的处理结果。

关键词:物联网;实体异构;Agent;数据处理;服务逻辑

中图分类号:TP18;TP393

文献标志码:A

Internet of Things-oriented Heterogeneous Entities Relation Service Model

LI Guanyu, AN Jingmin[†]

(Faculty of Information Science and Technology, Dalian Maritime University, Dalian 116026, China)

Abstract: Entities mostly exist in heterogeneous forms in Internet of Things (IoT), so solving heterogeneous issue is a vital task for developing IoT. On the basis of analysis on the characteristics of entities and their uploaded information in IoT, an agent-based heterogeneous entities relation service model was proposed. The functions of model include entity information data processing and entity service logic construction, and through using the new inter-objects dynamic relation calculation formula proposed in the paper, the heterogeneous data was transformed into their relation closeness (values), and the agent service sequence was generated based on it. Combining with the bipartite graph theory, the agent service logic was realized in IoT environment, which overcomes the heterogeneity that entities are unable to communicate with each other properly. In the model, entity interactions were completely replaced with agent interactions, and these agents were governed under the established service logic. Finally, the simulation experiment was performed on the model. And the result shows that not only can user's service requirements in IoT environment be met but also it is better than the traditional ontologies and semantics technology to deal with heterogeneous issues in the data updating, service quality and conflict handling.

Key words: internet of things; heterogeneous entities; Agent; data processing; service logic

* 收稿日期:2018-05-23

基金项目:国家自然科学基金资助项目(61371090,61602076), National Natural Science Foundation of China(61371090,61602076)

作者简介:李冠宇(1963—),男,辽宁丹东人,大连海事大学教授,博士生导师,博士

[†] 通讯联系人, E-mail:870457569@qq.com

自物联网(Internet of Things, IoT)这一概念被提出以来,已受到学术界和工业界的广泛关注,并被誉为一代网络。然而,IoT中的实体异构性使得IoT的发展遇到了瓶颈,所以如何更好地处理IoT实体的异构问题是推动IoT发展的重要因素,也是一项具有挑战性且有价值的工作。

IoT环境中,实体异构使通过RFID或传感器上传到网络中的实体信息无法正常交换,致使实体间交互或通信受阻,用户得不到满意的服务。针对这一问题,近年来国内外专家学者做了诸多研究,其中近两年的研究成果主要有:文献[1]提出了一种轻量级IoT语义模型,将IoT中的关键概念进行语义标注,发现传感器中的异构实体数据并行进行互操作。文献[2]提出了一种语义物联网框架,利用本体对IoT中的特定领域进行统一的半自动标注,以实现领域内异构实体的互操作。文献[3]提出了一种IoT中服务定义和发现的语义感知框架,通过API定义语言所定义的本体,支持服务的语义标记,进而实现异构实体在网络空间中的服务功能。文献[4]提出了一种利用IoT中的实体异构关系推荐事物的系统结构,依据人的行为习惯,发掘并计算出IoT中异构实体间隐含的关联关系,且以此关系向用户推荐实体服务。文献[5]提出了一种异构IoT实体的逻辑交互方法,以实现环境辅助生活(Ambient Assisted Living),面向实体并对异构实体数据进行语义标注,构建实体交互的语义逻辑,以判断当前用户需求的服务。

这些对于IoT环境下异构实体交互的方法均有各自的优点,文献[1-2]主要是对IoT中异构实体的数据做处理,并未涉及到服务层面;文献[3]主要涉及的是对IoT中的异构实体服务的处理,对于服务的生成和来源并未提及;文献[4-5]对网络中异构实体数据的处理,生成服务和应用均做了详细的研究说明。但在这些方法中对于异构问题的处理主要使用的方法都是本体语义标注技术,在保证了解决结果或服务提供的准确性的同时,并不能保证异构实体交互的冲突问题,且冲突问题也是决定服务质量的重要因素;而且文献[4-5]中的方法均未涉及到数据和服务更新的处理方法,由于IoT中的数据量庞大,数据更新得较快,对于利用数据生成服务而言,服务更新是必要的。本文对于上述存在的问题,受文献[6-7]启发,提出了一种面向IoT的异构实体关系服务模型,该模型分为三个层级并且有两种功能,分别是异构实体数据处理和服务逻辑构建。物

理实体层上的是IoT环境中异构的实体集,并为每一个实体在网络空间中构建一个与之对应的Agent;系统层是利用上载的异构实体信息数据计算出实体间的动态关系,并随着上传数据量增多,动态地更新实体间关系;Agent及服务层将计算得到的实体间关系转化为带有权值的Agent关系网络结构,以此为服务逻辑,通过Agent间互操作向实体层的实体提供服务。最后,采用IoT环境下的6种场景中的数据,与上述文献中的方法进行对比实验。

1 实体关系网络模型的机理

1.1 实体关系的概念模型

IoT实现的宗旨是以为人提供服务为核心^[8-9]。而在IoT环境中,人(individual, I)在日常生活中身边存在着各式各样的实体物(object, O),且与它们频发交互,这种交互过程是人为习惯的一种写照^[4],同时也逐渐地形成了物间的某种隐含的逻辑关系。所以,以某个人 I 为中心的视角来看, I 与周围的 O 可以构成一个简单的小生境(Niche)。根据 I 所处的IoT环境,本文给出Niche的五元组形式化表示:

$$Niche(I, I_E, B, O, S)$$

I 是物联网环境中的个体人,是Niche构建的核心元素;

I_E 是智能环境(Intelligent Environment),是物联网关于 I 日常生活的缩影;

B 是 I 在Niche中的行为(Behavior, 简记 B),每一个行为 B_i 包含多个 I 与 O 的交互动作(Action, 简记 Act),即 $B_i = \{Act_1, Act_2, \dots, Act_k, \dots, Act_n\}$;

O 是物联网中的异构实体物,是与 I 交互的对象,也是向 I 提供服务的对象;

S 是 I_E 中的传感器等设备,将 I 与 O 的交互情景所产生的上下文context上载到网络中。

I 在Niche环境中与 O 的交互可用下列关系式表示:

$$I \otimes O_1 \rightarrow Act \langle I, O_1 \rangle = Act_1,$$

$$I \otimes O_2 \rightarrow Act \langle I, O_2 \rangle = Act_2,$$

...

$$I \otimes O_n \rightarrow Act \langle I, O_n \rangle = Act_n;$$

从上式中可以推理发现, $\{O_1, O_2, \dots, O_n\}$ 间可能存在着关于 I 的隐含逻辑关系,并得出下式:

$$O_1 \otimes O_2 \rightarrow R \langle O_1, O_2 \rangle = R_{12}, \dots, O_1 \otimes O_n \rightarrow R \langle O_1, O_n \rangle = R_{1n}.$$

综上 Niche 的概念模型和实体物间逻辑关系, 得到图 1 所示的 Niche 中的实体关系概念模型.

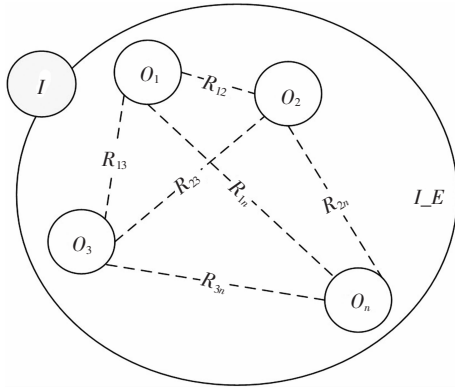


图 1 Niche 中的实体关系概念模型

Fig.1 Conceptual model of entity relations in Niche

1.2 实体关系网络模型 的表示

上文 1.1 节提出了 Niche 中实体关系概念模型, 但由于 IoT 环境中实体的异构性, 给该关系的发掘与应用带来了巨大的挑战. 本文就此问题提出将每一个 Niche 中的实体在网络中建立一个与之对应的虚拟镜像 Agent, 利用其统一并便于管理的特性屏蔽掉实体原本的异构性. 如图 2 所示, 将实体交互信息通过传感器上传到网络中的模型系统层进行计算, 利用物间动态关系计算机理使上下文数据所包含的实体信息转化为物间的关系紧密度(值), 且随着 I 的行为习惯的变化和实体数据的不断更新, 模型将重新计算该关系以代替原值, 保证了物间关系的实时准确性和动态的特点; 将结果与 Agent 结合, 在 Agent 及服务层构建 Agent 逻辑关系网络, 使得 Agent 的交互代替实体交互; 且介于实体与 Agent 的对应关系, 其逻辑关系网络和形成的最终服务逻辑同样具有动态性, 并确保了在不同场景和情况下,

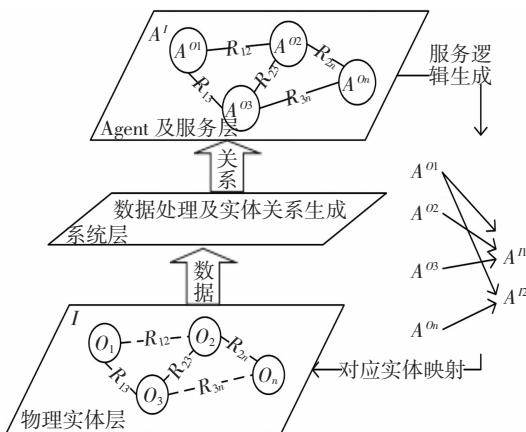


图 2 实体关系及服务逻辑生成原理图

Fig.2 Creating mechanism of entity relations and service logics

向用户提供的服务兼顾了准确率又具有较好的稳定性和高效性.

2 异构实体的 Agent 抽象化表示

2.1 实体的 Agent 构建

本质上, IoT 中实体差异大且异构^[10-11], 这对于研究 IoT 中实体信息交换以及发掘、应用实体间存在的隐含关联关系带来很大难度. 而实体间潜在关系对于 IoT 服务具有重要利用价值. 为屏蔽实体异构性进而实现统一管理, 本文提出将 IoT 或 Niche 中的实体在网络空间中建立一个与之对应的镜像 Agent, 利用 Agent 间交互和通信代替原本异构实体的互操作, 并向 Niche 环境中的用户 I 提供满意的服务.

Agent 是一个独立的智能体, 其可以对接收的上下文进行分析推理并做出相应的动作^[12-15], 所以, 设 Agent 与实体的对应关系表达式为 $A=(Ety, Act)$.

其中, Ety 表示 Niche 中的实体, 且 $Ety=I \cup O$; Act 是 I 与 O 的交互动作. 若令上下文为 Ctx, Act, Ctx 和行为 B 的关系即为: $B_i = \{Act_i(Ctx_i) = Ctx_{i+1}, Act_{i+1}(Ctx_{i+1}) = Ctx_{i+2}, \dots, Act_{i+n}(Ctx_{i+n}) = Ctx_{i+n+1}\}$, 即 B_i 是一个有序的 Agent 动作序列.

在 Niche 中, 按服务与被服务对象及核心与非核心元素来分, 将实体分为 I 与 O; 同理, 对应到 Agent, 将 Agent 分为 Individual Agent (A^I) 和 Object Agent (A^O), $A^I = (I, Act), A^O = (O, Act)$.

2.2 实体与 Agent 的映射关系构建

为了能准确地实现 Agent 间的交互代替异构实体的互操作, 以向用户提供准确高效的服务逻辑, 在网络中建立了实体的 Agent 的同时, 必须构建一种完善的实体 Agent 映射方法, 使得二者一一对应.

将实体 $Ety = I \cup O$ 用 Agent 抽象化表示为 $A = A^I \cup A^O$, 但在映射关系构建过程中并无差别. 文中采用实例名称语义相似度计算方法, 当实体 Ety_i 和智能体 A_i 有 $Sim(Ety_i, A_i) = 1$ 时, 说明实体 Ety_i 与智能体 A_i 名称完全相似, 即为相同, 认为实体 Ety_i 与智能体 A_i 映射匹配.

在已构建的智能体 A_i 中做身份标记 (Identity Tag), 使用文献[16]所提出的实例名称语义相似度计算公式进行身份名称匹配, 如公式(1)和(2)所示:

$$Sim_2(Ety_i, A_i) = \max \frac{\min(c_{Ety_i}, c_{A_i}) - D(c_{Ety_i}, c_{A_i})}{\min(c_{Ety_i}, c_{A_i})} \quad (1)$$

公式(1)是名称的字符串匹配方法,将匹配程度数值化;其中 c_{Ety_i} 和 c_{A_i} 表示向量,分别来自实体及其抽象表示 Agent 的实例概念名称, $D(c_{Ety_i}, c_{A_i})$ 指两个概念间的编辑距离。

$$Sim_w(Ety_i, A_i) = \frac{2 * depth(lso(Ety_i, A_i))}{depth(c_{Ety_i}) + depth(c_{A_i})} \quad (2)$$

公式(2)是基于 WordNet^[17-18]的语义相似度计算方法; $depth(c_{Ety_i})$ 和 $depth(c_{A_i})$ 分别指实体与 Agent 的实例名称在 WordNet 语义树中的深度; $lso(Ety_i, A_i)$ 则表示概念名称 c_{Ety_i} 和 c_{A_i} 的最近共同祖先概念名称。

当 $Sim_z(Ety_i, A_i)$ 与 $SIM_w(Ety_i, A_i)$ 的值均为 1 时,认为实体与 Agent 身份匹配对应成功,否则失败。

2.2.1 身份匹配算法

根据 2.2 小节阐述的实体 Ety 与智能体 A 的匹配方法,以及公式的使用,提出具体的匹配算法 E-A-Matching 如下。

算法名称:E-A-Matching

输入:Ctx;

输出:实体 Ety 与智能体 A 的匹配结果;

1. 从上下文 Ctx 中提取主体 I 与客体 O 的概念名称:

$Ctx | -I.Identity; Ctx | -O.Identity;$

2. For ($i=1; i \leq n; i++$)

```

{
  For ( $j=1; j \leq n; j++$ )
  {
    IF 字符串匹配:  $Sim_z(Ety_i, A_j)=1;$ 
    Visit WordNet Semantic tree 计算  $Sim_w(Ety_i, A_j);$ 
    Else Continue;
    IF  $Sim_w(Ety_i, A_j)=1;$ 
      Return OK; Break;
    Else Continue;
  }
  Return False;
}

```

3 实体物间的隐含关系发现

3.1 实体的信息数据处理

IoT 中传感器和 RFID 等设备将包含实体信息的上下文上传到网络环境,网络中的使用主体对此加以分析推理,并用于预测和推荐服务等方面^[19-20]。对于上载的单个上下文,其可以表示当前实体或者多个实体的某个动作 Act_k ;对于连续上载的上下文,可以体现出某个时间段内实体或实体间的一系列动

作序列 $\{Act_1, Act_2, \dots, Act_k, \dots, Act_n\}$,即某个行为 B_i 。

目前,还未有对于上下文标准及严格的形式化定义^[21-22],本文根据实际需求给出上下文(context)形式化定义:

$$Ctx = (I, Act, O, L, T, F)$$

其中: I 是 Niche 中的实体人; O 表示实体物; Act 是上下文中的谓词(predicate),表示 I 与 O 的交互动作; L 表示 I 与 O 发生交互动作的位置(Location); T 是发生交互的时间区间(time interval),如 $T_s = \langle t_s, t'_s \rangle$, t_s 和 t'_s 分别表示交互的开始和结束时刻; F 指在 T 时 I 使用 O 的频数。

3.2 实体物间的动态关系计算

上下文中的数据被上载到系统层,根据 Ctx 中的 L, T 和 F 计算实体物间的动态关系 $\omega_{R_i}, \omega_{R_j}$ 是一个数值表示,表示 O_i 和 O_j 之间的关系紧密度,并给出公式(3):

$$\omega_{R_i} = \alpha R_{ij}^L + \beta R_{ij}^T + \gamma R_{ij}^F \quad (3)$$

上式中, R_{ij}^L, R_{ij}^T 和 R_{ij}^F 分别表示 O_i 和 O_j 基于 L, T 和 F 因素的关系值; α, β 和 γ 分别是三种因素对应的影响因子,且 $\alpha + \beta + \gamma = 1$ 。

3.2.1 基于位置 L 的实体物间关系计算

物体相对位置间的距离是影响其是否能产生交互关系的重要因素^[4,19]。所以,在计算相对位置关系时,应充分考虑距离影响,结合文献[23]中的空间距离衰减函数,给出公式(4):

$$R_{ij}^L = k p_i p_j d_{ij}^{-\theta} \quad (4)$$

其中, k 是常量系数, p_i 和 p_j 分别是 O_i 和 O_j 的规模大小,在 Niche 环境下可定义为与 O_i 和 O_j 所在同一个位置的实体物数量; d_{ij} 是物体二者间的实际距离, θ 取值为 1.5; 且 $i \neq j$ 。

在计算物体间实际距离时,本文以 L 为单位,计算 d_{ij} ,即为计算 O_i 和 O_j 所在位置 L_i 和 L_j 间的距离,所以采用半正矢公式^[24],如下公式(5)所示:

$$d_{ij} = 2r \cdot \arcsin \left[\sin^2 \left(\frac{\varphi_i - \varphi_j}{2} \right) + \cos(\varphi_i) \cos(\varphi_j) \sin^2 \left(\frac{\lambda_j - \lambda_i}{2} \right) \right]^{\frac{1}{2}} \quad (5)$$

式中: r 表示地球半径, $(\varphi_i, \lambda_i), (\varphi_j, \lambda_j)$ 两点表示 L_i 和 L_j 的纬度和经度。

3.2.2 基于交互时间 T 的实体物间关系计算

Niche 中的 O 是以静态形式存在的,其交互动作 Act 的产生是由 I 决定的; O 间的交互时间 T 即是 I 在某一时间段内 $\langle t_s, t'_s \rangle$ 同时使用 O_i 和 O_j 的时间。

$t(O_i)$ 表示 I 使用 O_i 的时间, $t(O_j)$ 表示 I 使用 O_j

的时间, $t(O_i||O_j)$ 则表示 I 同时使用 O_i 和 O_j 的时间 ($i \neq j$). 对于 Niche 中基于 T 的 O 间关系计算理论可以转化为在使用 O_i 的时间基础上再使用 O_j 的概率计算, 等价于条件概率, 结合大数定律给出公式(6):

$$R_{ij}^t = P(t(O_j)|t(O_i)) = \frac{P(t(O_i||O_j))}{P(t(O_i))} \approx \frac{t(O_i||O_j)}{t(O_i)} \quad (6)$$

3.2.3 基于交互频数 F 的实体物间关系计算

Niche 中的以静态形式存在的 O 间的交互频数是通过 I 同时使用 O_i 和 O_j 的次数求和而来. 对于 O_i 或 O_j 的使用频数的统计, 来源于上文 3.1 小结 Ctx 定义中的 Act 的个数统计.

$f(O_i)$ 表示 I 使用 O_i 的频数, $f(O_j)$ 表示 I 使用 O_j 的频数, $f(O_i||O_j)$ 则表示 I 同时使用 O_i 和 O_j 的频数 ($i \neq j$). 对于 Niche 中基于 F 的 O 间关系计算理论同样可以转化为在使用 O_i 的频数基础上再使用 O_j 的概率计算, 等价于条件概率, 结合大数定律给出公式(7):

$$R_{ij}^f = P(f(O_j)|f(O_i)) = \frac{P(f(O_i||O_j))}{P(f(O_i))} \approx \frac{f(O_i||O_j)}{f(O_i)} \quad (7)$$

3.2.4 影响因子的设定

本文对 Niche 中针对 I 的 O 间动态关系以关系紧密度的形式表现; O 间的关系主要受 L 、 T 和 F 影响, 而且以基于三种因素的 O 间关系和的形式表示关系紧密度. 但因为 I 的行为习惯和方式是不定的, 导致三种因素对关系紧密度的影响权值不固定, 所以人为提前设定影响因子的值是不可取的, 本文提出利用系统现有数据动态计算影响因子的值.

将系统中 d_{ij} 、 $t(O_i)$ 和 $f(O_i)$, 针对同一个 I 进行数据统计并形成集合, O_i 和 O_j 间的距离集 $d_{ij} = \{d_{ij}^1, d_{ij}^2, \dots, d_{ij}^n\}$, 同时使用 O_i 和 O_j 的时间集 $t(O_i||O_j) = \{t_1(O_i||O_j), t_2(O_i||O_j), \dots, t_n(O_i||O_j)\}$, 同时使用 O_i 和 O_j 的频数集 $f(O_i||O_j) = \{f_1(O_i||O_j), f_2(O_i||O_j), \dots, f_n(O_i||O_j)\}$. 由于三个变量单位不统一, 要先进行数据标准化处理^[25], 以达到数据指标间的可比性, \max 和 \min 函数分别表示取集合中最大和最小元素.

$$\begin{aligned} X_{L_{ij}}^* &= \frac{d_{ij}^k - \min(d_{ij})}{\max(d_{ij}) - \min(d_{ij})} \\ X_{T_{ij}}^* &= \frac{t_k(O_i||O_j) - \min(t(O_i||O_j))}{\max(t(O_i||O_j)) - \min(t(O_i||O_j))} \\ X_{F_{ij}}^* &= \frac{f_k(O_i||O_j) - \min(f(O_i||O_j))}{\max(f(O_i||O_j)) - \min(f(O_i||O_j))} \end{aligned} \quad (8)$$

已标准化的三个集合中的元素均在 $[0, 1]$ 内, 并分别取其平均值 $\bar{X}_{L_{ij}}^*$ 、 $\bar{X}_{T_{ij}}^*$ 和 $\bar{X}_{F_{ij}}^*$, 并将其归一化, 计算

α 、 β 和 γ , 如公式(9)所示:

$$\begin{aligned} \alpha &= \frac{\bar{X}_{L_{ij}}^*}{\bar{X}_{L_{ij}}^* + \bar{X}_{T_{ij}}^* + \bar{X}_{F_{ij}}^*} \\ \beta &= \frac{\bar{X}_{T_{ij}}^*}{\bar{X}_{L_{ij}}^* + \bar{X}_{T_{ij}}^* + \bar{X}_{F_{ij}}^*} \\ \gamma &= \frac{\bar{X}_{F_{ij}}^*}{\bar{X}_{L_{ij}}^* + \bar{X}_{T_{ij}}^* + \bar{X}_{F_{ij}}^*} \end{aligned} \quad (9)$$

4 服务逻辑构建

在 Niche 中, 公式(3)~(9)计算出对于 I 的 O 间关系紧密度. 为了更好地应用和管理该计算结果, 提出基于 Agent 构建一种 IoT 中的服务动态逻辑.

4.1 Agent 间的关系网络

第2节完成了实体的 Agent 构建及对应匹配; 现将每一个不同的 A^o 作为图 $G(V, E)$ 的顶点, 则对于 A^i 有 $V(GA^i) = \{A^{o1}, A^{o2}, \dots, A^{on}\}$, G 的边 $E(GA^i) = V(GA^i) \otimes V(GA^i) = \{\langle A^{o1}, A^{o2} \rangle, \langle A^{o1}, A^{o3} \rangle, \langle A^{o2}, A^{on} \rangle, \dots, \langle A^{o1}, A^{on} \rangle\}$, G 是带有权值的无向图, 以 O_i 和 O_j (A^{o1} 和 A^{on}) 的关系紧密度值作为权值. 最终构建如图3所示的关于 A^i 的 A^o 关系网络模型.

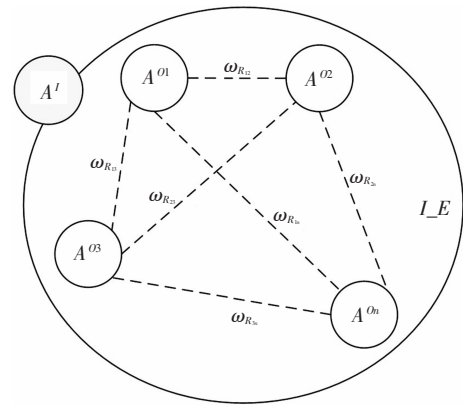


图3 Niche 中关于 A^i 的 A^o 间关系权值图

Fig.3 A^i -based inter- A^o relation weighted graph in Niche

4.2 Agent 间的动态服务逻辑

Niche 中关于 I 的 $O(A^o)$ 间的动态关系的发掘目的在于用其实现 I 的实体物服务. Agent 服务逻辑思想是: 当 I 使用了某个实体物 O_i 时, 对应的 A^{o1} 根据模型系统中与其相关 A^{on} 等 Agent, 按关系紧密度大小排序, 选择当前与 A^{o1} 关系紧密度最大的 A^{on} 等, 与 A^{o1} 共同向 I 提供服务, 以最大程度上满足用户需求. 如表1所示, 由于在关系网络中(图3), 每一个 A^o 都与若干个 Agent 形成关联关系, 但关系紧密度值不同. 所以针对每一个 A^i , 形成一个以关系紧密度值排序的 A^o 序列.

表 1 关于 A^i 的 A^o 间的服务序列样例

Tab.1 A^i based inter- A^o service sequences sample

A^o	Relation Sequence
A^{o1}	$A^{o3}, A^{o2}, \dots, A^{o\alpha}, \dots, A^{o\eta}$
A^{o2}	$A^{o1}, A^{o3}, \dots, A^{o\beta}, \dots, A^{o\eta}$
...	...
$A^{o\alpha}$	$A^{o\beta}, A^{o\alpha}, \dots, A^{o\delta}, \dots, A^{o\eta}$
...	...
$A^{o\eta}$	$A^{o2}, A^{o1}, \dots, A^{o\alpha}, \dots, A^{o\beta}$

根据系统模型中 O 或 A^o 间的关系紧密度和已构建的 A^o 间服务序列,利用二分图理论^[26],生成面向 I (即 A^i)的 A^o 服务逻辑.如图 4 所示.

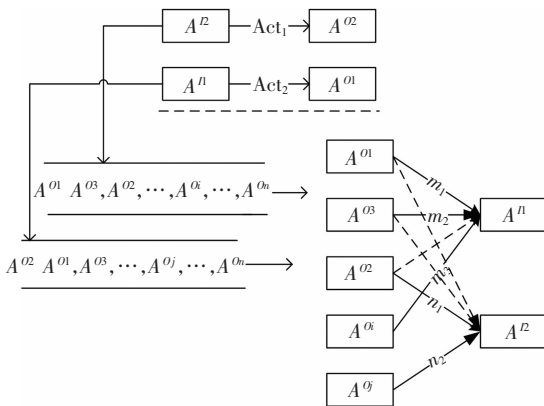


图 4 服务逻辑生成机理图

Fig.4 Service logic generation mechanism diagram

从 Ctx 中获取 I 当前使用 O_k , 在系统中已构建的 A^o 服务序列中抽出 A^{ok} 的 Relation Sequence,并动态生成面向 I 的服务逻辑(按照与 A^{ok} 关系紧密度大小顺序提供服务).如图 4 所示,当两个或多个 Niche 有交集时(有共同的实体物 O),并在对应的服务序列中存在 Agent 的交叉,在生成动态服务逻辑时,依据系统推荐给 I 的次序,决定将交叉的 Agent 服务提供向 A^i1 或 A^i2 .图中虚线表示因 Agent 服务存在交叉而暂时不能向 $A^i(I)$ 提供的 Agent 服务;实线表示目前可以提供的 Agent 服务,并按照服务序列中的顺序,解决了异构实体交互和服务冲突问题.

二分图中的权值表示使用频数 F , 随着 Ctx 不断上传到系统中, L 、 T 和 F 三种要素的统计量不断增加,同时对应的三个影响因子也随之变化,致使 $O(A^o)$ 间的关系紧密度发生变化,最后 A^o 间的服

务序列和服务逻辑改变,实现服务逻辑更新.

5 实验与结果

5.1 实验环境及数据说明

本文实验平台在 Windows10 环境下搭建,是在 Microsoft.NET framework 和 SQLServer 2012database 环境下进行的,运行在 3.80 GHz,4 GB 内存, windows 1 064 位操作系统的台式机上.

为了验证本文提出的系统模型的正确性和有效性,参考了文献[4],收集实体人 I 经常出没的 6 种场景下的 Ctx 数据,模拟 IoT 错综复杂的数据环境;对这 6 种情境下的部分实体物进行标记,收集这些物与个体人交互的 Ctx 数据.这些场景主要是 I 活动的娱乐场所、办公室、厨房、交通工具、医疗场所以及其他的家庭室内应用环境.具体数据详情见表 2,在每一场景下所标记的 O 的数量和收集的 Ctx 数量不同,主要是因为在一时间内,几个实体人 I 在每个场景下的活动频率不同,这样以某一时间段为标准收集数据可以体现出数据的随机性,实验结果更具代表性.

表 2 实验数据集表

Tab.2 Experimental data set

No	情景	O 数量/个	标记数量/个	Ctx 数量/个
1	娱乐场所	28	118	2 897
2	办公室	20	51	5 745
3	厨房	25	103	6 896
4	交通	11	24	2 067
5	医疗	10	18	1 188
6	家庭应用	33	83	7 089

利用 Ctx 作为实验数据集,由于 Ctx 无标准统一的数据格式,在收集到表 2 中的数据时,根据 3.1 小结给出的上下文定义,对其进行人工的信息抽取,以达到可以进行对比实验的要求.抽取后的 Ctx 主要包含 I 和 O 的身份名称,具体的交互动作,所在的地理位置及经纬度,和发生交互的时间以及二者交互的频数.具体数据 Ctx 记录如表 3 所示,文中只列举出厨房和家庭应用的 Ctx 数据实例,其余 4 种场景的数据并未列出,但格式与其完全相同.

表 3 数据 Ctx 实例
Tab.3 Examples of data Ctx

实体人 I (Identity)	动作 Act	实体物 O	位置 L (经度, 纬度)/(°E, °N)	时间 T	频数 F 次
Test I1	Use	Knife	Kitchen(121.535 702, 38.875 473)	11:10:04–11:13:34	1
Test I1	Use	Pot	Kitchen(121.535 703, 38.874 762)	11:11:45–11:24:12	1
Test I1	Use	Knife	Kitchen(121.535 702, 38.875 473)	11:17:14–11:21:54	1
Test I1	Use	Microwave	Kitchen(121.535 709, 38.875 469)	11:17:14–11:21:54	1
...
Test I2	Watch	TV	Bedroom(121.535 711, 38.875 501)	10:54:20–11:21:54	1
Test I2	Use	Massage chair	Bedroom(121.535 707, 38.875 499)	10:55:27–11:21:37	1
...

将收集的所有 Ctx 随机分为两部分,一部分 60% 的数据在实验平台中作为训练集,用于生成模型的服务逻辑规则集,另外一部分 40% 的数据作为测试集,以验证本文提出的系统模型的准确性和有效性。

5.2 实验对比与分析

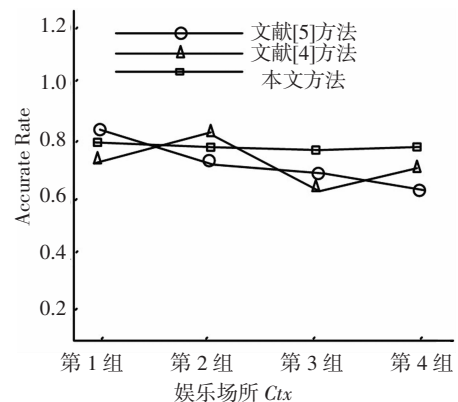
为验证本文方法在数据处理和服务生成这两部分的优势,分别与目前和本文研究相关的最新研究成果,文献[4]和文献[5]进行对比.本文从两个角度进行实验,分为单场景和多场景,详细对比并说明三种方法在对于单个场景和多个场景的数据时,所给出的服务推荐准确率以及服务推荐完成时间。

对于服务推荐的准确率的评估,本文考虑了服务推荐结果出现冲突情况,即出现相同时刻系统向多个 I 推荐同一满足其当前需求的服务的情况,这样会使得部分 I 得不到该推荐的服务,所以属于推荐错误,在计算准确率时应将其计算在内,给出公式(10):

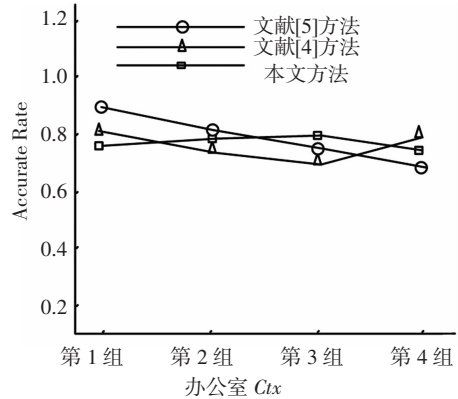
$$AR = \frac{PS - CS}{PS + ES} \quad (10)$$

其中, AR (Accurate Rate) 表示服务推荐的准确率, PS (Proper Services) 指当前系统模型所提供的能最大程度上满足用户需求的服务, ES (Error Services) 指系统模型提供的未能满足用户需求的服务, CS (Conflicting services) 是当前在 PS 中,但在服务过程出现冲突的服务。

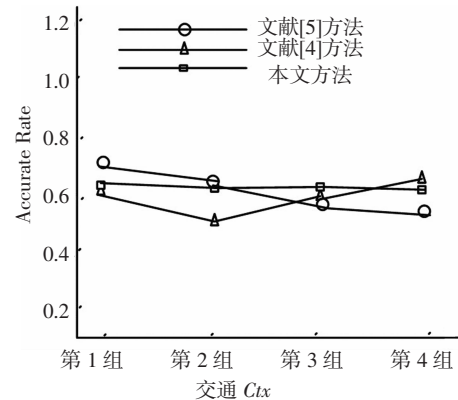
实验 1 在进行利用单场景数据进行实验时,为模拟 IoT 数据实时更新的特性, 本文将实验测试集的数据按照 6 种场景分为对应的 6 组,并在每一组中,将数据均分为 4 小组进行实验,将三种方法在每一小组的数据实验下所得到的结果利用公式(10)计算得到服务推荐准确率进行对比.如下图 5 所示。



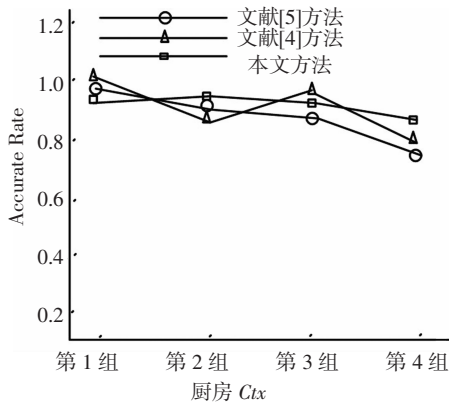
(a) 娱乐场所场景服务准确率对比



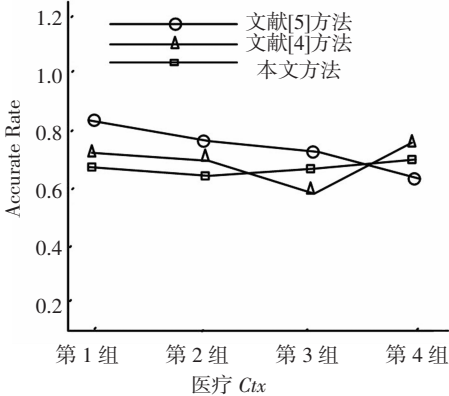
(b) 办公室场景服务准确率对比



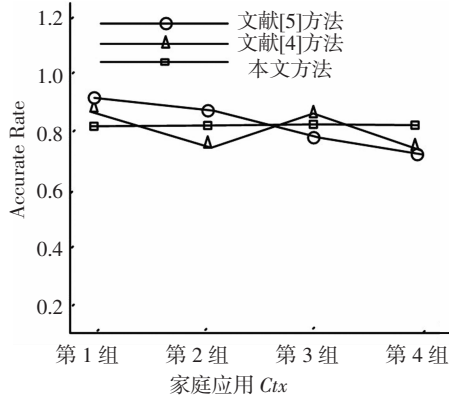
(c) 交通场景服务准确率对比



(d) 厨房场景服务准确率对比



(e) 医疗场景服务准确率对比



(f) 家庭应用场景服务准确率对比

图5 单场景实验的对比

Fig.5 Test comparison on simple scene

从以上分别利用6种场景下的Ctx进行实验的结果可以发现,三种方法在4组数据上得到的服务准确率差距并不是很大,但从4组结果整体来看,文献[4]方法的服务准确率在每组上的结果上下波动较大,没有稳定性,原因是其所设计的关系计算模型中的主要参数是人为主观设定,而未能从实际数据角度考虑;文献[5]方法在6种场景下的结果均呈现出从第1组到第4组递减的情况,原因在于该文中的主要处理技术手段为语义技术,随着数据量

的加大和复杂性的提升,其出现服务冲突的概率增大;而本文的方法考虑了以上两个主要问题,在主要参数设计上采用了一种动态计算机制,得到的参数取值较为适合当前的数据信息,同时利用语义和计算相结合的方式求得当前用户需要的服务,并最终取得如下实验结果:

在6种场景下,每组的结果较为稳定,服务准确率分别稳定在0.79, 0.80, 0.65, 0.87, 0.66, 0.81左右.本文实验考虑了IoT环境的数据不稳定性和动态性,在实验过程中,将数据随机分组并依次实验,本文方法在保证服务准确率同时,稳定性较高.

实验2 在多场景数据下对三种方法进行实验,本文将测试集中6种场景的数据直接随机分为均等的4组,并采用与实验1相同的方式,将三种方法在每组数据的实验结果计算得到服务推荐准确率,如图6所示.

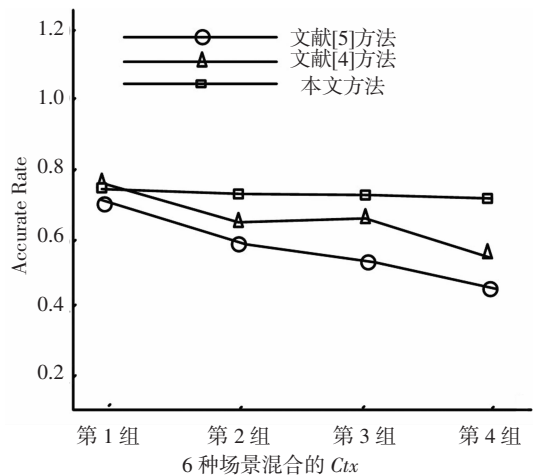


图6 多场景实验对比

Fig.6 Test comparison on multiple scenes

多场景实验是将收集到的6种场景下的Ctx合并随机均等分为4组,以仿真IoT数据环境多样性.该情况下的实验数据相比于实验1,在数据的随机性、复杂性和多样性以及动态性上有了大幅度提升.而从结果可以看出文献[4]和文献[5]在此情况下得到的服务准确率呈波动向下递减趋势,且在实验1单场景的实验结果的均值基础上有明显下降(下降梯度增大),说明其两种方法在应对IoT中数据本身具有的多变和复杂等特点的不适应性,且由实验1到实验2可知,随着数据的上述特性幅度加大,对其方法的影响就越大;本文所提方法的4组实验服务准确率稳定在0.76~0.77附近,说明了本文方法针对该问题的适应性和在处理多环境数据的优势

所在.在错综复杂的IoT数据环境中,能将数据转化为服务,并保持服务准确率的稳定性是目前推进IoT研究重要的一环.

实验3 利用实验2的6种场景混合的数据对模型在单位时间(50 s)内完成服务推荐的效率进行对比,以对本文方法在服务推荐响应时间和模型应用的可扩展性方面进行讨论.使用测试集中的所有数据分别对文献[4]、文献[5]和本文中的模型进行测试,结果如图7所示.

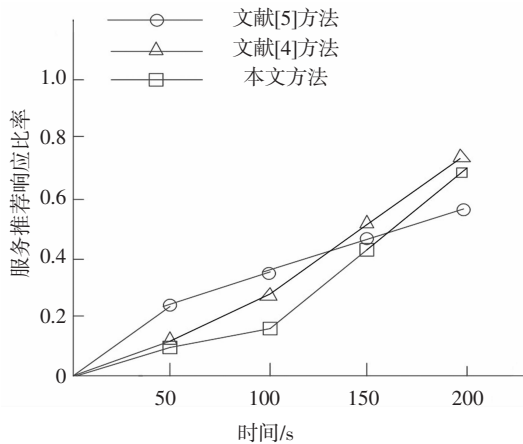


图7 单位时间内服务完成率对比

Fig.7 Comparison of service completion rate in unit time

以50 s为单位对三种方法模型所完成的服务推荐数量进行统计并对比,文献[5]中的方法在前100 s内所响应的服务数要大于文献[4]和本文中的方法;但150 s后,文献[4]和本文方法的优势较为明显,服务响应比更高.文献[5]在实现服务推荐过程中以语义标注和语义逻辑推理为主,而文献[4]与本文主要是通过数学公式计算实现服务推荐过程,所以随着模型运算或推理的数据体量不断增多时,语义复杂性增大,导致文献[5]中方法的效率降低,而文献[4]与本文方法仍具有较好的计算性能和应用扩展性.

综合实验1,2和3以及文献[4-5]分析,文献[4]对IoT异构实体数据处理和服务构建的方法没有考虑对于数据更新的问题的处理,该文中对于计算公式的影响因子是人为设定的,并不能较好地满足当前大量变化的数据处理需求;文献[5]主要对数据采用传统的语义标注处理方式,并构建语义服务逻辑,当数据量庞大且复杂时,易造成语义冲突和服务响应效率下降.本文相比二者,优势在于以下两点:

1)在面向动态变化的数据时,应用含有动态影响因子的计算公式,以解决数据更新的问题,保持

稳定的服务准确性和具有较高的服务响应效率以及良好的应用扩展性.

2)在处理多环境数据时,结合Agent屏蔽了实体的异构性,并将多场景服务虚拟化为以某个I为核心的Niche环境下的服务,应用Niche中动态构建的服务逻辑序列,实现Agent代替实体交互,使得在复杂环境下仍能保持较高且稳定的服务准确性.

6 结论

本文基于Agent的智能特性,提出了一个面向IoT的异构实体间关系服务模型,该模型主要解决了目前在处理IoT异构实体数据更新和场景服务的准确率稳定性问题.本文方法区别于现有的传统本体语义技术的优势如下:

本文给出了全新的实体动态关系计算公式和动态影响因子计算方法,在面对大量且不断更新实体数据时,有较高且稳定的计算结果和高效的计算性能;将实体关系计算结果结合Agent,构建动态逻辑服务序列,降低了服务的冲突,同时可以满足单场景和多场景的服务需求,且在服务准确率和冲突处理方面有较优的效果.

本文提出的方法更加适应IoT复杂多样的数据环境,弥补了目前对于IoT异构实体的数据处理和服务构建方法存在的不足,并在单场景和多场景的实验结果中证实了本文方法的优势.

参考文献

- [1] BERMUDEZ-EDO M, ELSALEH T, BARNAGHI P, *et al.* IoT-Lite: A lightweight semantic model for the Internet of things [C]//Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress. IEEE, 2017: 1-13.
- [2] WU Z, XU Y, ZHANG C, *et al.* Towards semantic web of things: From manual to semi-automatic semantic annotation on web of things [C]//International Conference on Big Data Computing and Communications. Springer International Publishing, 2016: 295-308.
- [3] KHODADADI F, SINNOTT R O. A semantic-aware framework for service definition and discovery in the Internet of things using CoAP[J]. *Procedia Computer Science*, 2017, 113: 146-153.
- [4] YAO L, SHENG Q Z, NGU A H H, *et al.* Things of interest recommendation by leveraging heterogeneous relations in the Internet of

- things[J]. *ACM Transactions on Internet Technology*, 2016, 16(2): 9.
- [5] XIAO B, KANTER T, RAHMANI R. Logical interactions for heterogeneous IoT entities via virtual world mirrors in support of ambient assisted living [J]. *Journal of Ambient Intelligence & Smart Environments*, 2016, 8(5):565—580.
- [6] RESTREPO S E, PEZOA J E, NAEINI M R. A model for optimal service allocation in a smart environment [J]. *Pervasive & Mobile Computing*, 2017, 42(7):45—47.
- [7] KAMARA-ESTEBAN O, AZKUNE G, PIJOAN A, *et al.* MASSHA: An agent-based approach for human activity simulation in intelligent environments [J]. *Pervasive & Mobile Computing*, 2017, 40(5):279—300.
- [8] SHIN D. A socio-technical framework for Internet-of-Things design: A human-centered design for the Internet of Things [J]. *Telematics & Informatics*, 2014, 31(4):519—531.
- [9] YANG L, LI W, GHANDEHARI M, *et al.* People-centric cognitive Internet of things for the quantitative analysis of environmental exposure [J]. *IEEE Internet of Things Journal*, 2017, 5(4):2353—2366.
- [10] ZHANG M, QIU Y, ZHENG R, *et al.* A novel architecture for cognitive Internet of things [J]. *International Journal of Security & Its Applications*, 2015, 9(9):235—252.
- [11] RAZZAQUE M A, MILOJEVIC-JEVRIC M, PALADE A, *et al.* Middleware for Internet of things: A survey [J]. *IEEE Internet of Things Journal*, 2017, 3(1):70—95.
- [12] BELLIFEMINE F, RIMASSA G, POGGI A. Developing multi-agent systems with a FIPA-compliant agent framework [J]. *Software Practice & Experience*, 2015, 31(2):103—128.
- [13] OTTER H S, VEEN A V D, VRIEND H J D. ABL0oM: Location behaviour, spatial patterns, and agent-based modeling [J]. *Journal of Artificial Societies & Social Simulation*, 2017, 4(4):2.
- [14] HONG Y, CHEN G, BUSHNELL L. Distributed observers design for leader-following control of multi-agent networks [J]. *Automatica*, 2017, 44(3):846—850.
- [15] TAPIA D I, RODRIGUEZ S, CORCHADO J M. A distributed ambient intelligence based multi-agent system for alzheimer health care [J]. *International Journal of Ambient Computing & Intelligence*, 2017, 1(1):15—26.
- [16] 刘平峰, 章佩璐, 陈冬林. 电子目录本体合并中的实例消重机制 [J]. *计算机工程*, 2012, 38(2):32—35.
- LIU P F, ZHANG P L, CHEN D L. Instance duplicate elimination mechanism in electronic catalog ontology merging [J]. *Computer Engineering*, 2012, 38(2):32—35 (In Chinese)
- [17] PIAO G, BRESLIN J G. User modeling on twitter with WordNet synsets and DBpedia concepts for personalized recommendations [C] // *ACM International on Conference on Information and Knowledge Management*. ACM, 2016:2057—2060.
- [18] HICKS A, RUTHERFOLD M, FELLBAUM C, *et al.* An analysis of WordNet's coverage of gender identity using twitter and the national transgender discrimination survey [C] // *Global Wordnet Conference*. 2016.
- [19] 曾雪琳, 吴斌. 基于位置的社会化网络的并行化推荐算法 [J]. *计算机应用*, 2016, 36(2):316—323.
- ZENG X L, WU B. Parallelized recommendation algorithm in location-based social network [J]. *Journal of Computer Applications*, 2016, 36(2):316—323. (In Chinese)
- [20] GRAVINA R, ALINIA P, GHASWZADEH H, *et al.* Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges [J]. *Information Fusion*, 2016, 35:68—80.
- [21] YANG X. An adaptive mechanism for inconsistent context resolution in ubiquitous computing [C] // *International Conference on Control Engineering and Communication Technology*. IEEE Computer Society, 2012:703—706.
- [22] 李昉松, 陶先平, 吕建, 等. 面向动作的上下文感知应用的规约与运行时验证 [J]. *软件学报*, 2017, 28(5):1167—1182.
- LI X S, TAO X P, LÜ J, *et al.* Specification and runtime verification for activity-oriented context-aware applications [J]. *Journal of Software*, 2017, 28(5):1167—1182. (In Chinese)
- [23] 刘瑜, 龚俐, 童庆禧. 空间交互作用中的距离影响及定量分析 [J]. *北京大学学报(自然科学版)*, 2014, 50(3):526—534.
- LIU Y, GONG L, TONG Q X. Quantifying the distance effect in spatial interactions [J]. *Acta Scientiarum Naturalium Universitatis Pekinensis*, 2014, 50(3):526—534. (In Chinese)
- [24] 石立兴, 胡访宇. 二维空间上基于呼叫详细记录数据的用户移动性分析 [J]. *计算机应用*, 2015, 35(9):2453—2456.
- SHI L X, HU F Y. Users' mobility analysis based on call detail record in two-dimensional space [J]. *Journal of Computer Applications*, 2015, 35(9):2453—2456. (In Chinese)
- [25] 马立平. 统计数据标准化——无量纲化方法——现代统计分析方法的学与用(三) [J]. *数据*, 2000(3):34—35.
- MA L P. Standardization of statistical data—dimensionless methods—learning and use of modern statistical analysis methods (3) [J]. *Data*, 2000(3):34—35. (In Chinese)
- [26] 刘一松, 朱丹. 基于聚类与二分图匹配的语义 Web 服务发现 [J]. *计算机工程*, 2016, 42(2):157—163.
- LIU Y S, ZHU D. Semantic web service discovery based on clustering and bipartite graph matching [J]. *Computer Engineering*, 2016, 42(2):157—163. (In Chinese)