

# 无线传感器网络中基于减法聚类的定位算法

陈洁洁

(湖北师范学院 计算机科学与技术学院, 湖北 黄石 435002)

**摘要:** 定位算法是无线传感器网络中的关键技术。文中在传统的 Dv-Hop 算法的基础上, 找出其产生误差的主要原因, 即对未知节点与锚节点之间的估计距离做出了修正, 提出一种无线传感器网络中基于减法聚类的定位算法。该算法用减法聚类的方法, 根据节点自身的密度, 选出锚簇头节点, 使锚簇头节点在锚节点密集处产生; 同时用所有锚簇头节点平均每跳距离的均值作为未知节点的网络平均每跳距离, 提高了定位精度, 减少了定位过程中的能量消耗。仿真实验表明, 该算法比 Dv-Hop 算法有更好的定位精度和鲁棒性。

**关键词:** 无线传感器网络; 减法聚类; 定位; 簇头; 锚

**中图分类号:** TP393

**文献标识码:** A

**文章编号:** 1673-629X(2011)10-0125-04

## Localization Algorithm Based on Subtractive Clustering in Wireless Sensor Networks

CHEN Jie-jie

(School of Computer and Technology, Hubei Normal University, Huangshi 435002, China)

**Abstract:** Node localization algorithm is one of the most critical technologies in wireless sensor networks. Based on Dv-Hop algorithm, point out the main reason for the error. Then a new algorithm named localization algorithm based on subtractive clustering in wireless sensor networks is proposed. In new algorithm, the subtractive clustering is used to choose the anchor cluster-head where has relatively high density; Besides, the average distance of all anchor cluster-heads is used as unknown nodes hop-distance in network. Compared with the traditional Dv-Hop localization algorithm, the improved algorithm is more reliable and more robust.

**Key words:** wireless sensor networks; subtractive clustering; localization; cluster-head; anchor

### 0 引言

在无线传感器网络<sup>[1]</sup> (Wireless Sensor Network, WSN) 领域中, 传感器节点随机部署, 导致了网络中绝大多数节点位置不能事先确定, 而与此同时在无线传感器网络中大量的应用都需要网络中节点的确切地理位置, 从而得到信息来源的准确位置。

基于无线传感器网络的定位机制的研究是目前国内外研究的热点, 所有 WSN 定位算法<sup>[2]</sup> 都是围绕是否测距而展开: 基于测距 (range-based) 和无需测距 (range-free) 的两类定位算法。基于测距定位算法是通过测量节点间点到点的距离或角度信息, 比如使用三边、三角测量法或最大似然估计法等方法来计算节点位置, 常用的算法有 TOA、TDOA、AOA 及 RSSI; 无需测距定位算法是利用节点间的邻近关系和连通性实现定位, 常用的算法有质心算法、Dv-Hop<sup>[3]</sup> 算法以及

APIT。

文献[4]提出一种无线传感器网络中模糊平均的定位算法, 通过寻找传播模式相似度足够大的节点组成聚类群, 并对聚类群内各节点 RSSI 值取平均得到, 减轻了随机性的对应传播模式路径损耗系数, 同时减轻小尺度随机衰落的影响, 最终提高定位精度。文献[5]提出一种无线传感器网络定位问题中的分簇算法, 这种算法实现多跳节点的分簇, 并且有较小的通信量。文献[6]提出一种基于减法聚类的无线传感器网络分簇算法, 考虑节点密集度进行分簇。从这些研究可以看出, 可以将模糊聚类的方法运用到定位算法中。

其中 Dv-Hop 算法是与距离无关的定位算法, 主要思想是将待定位的节点到锚节点之间的距离用网络平均每跳距离乘以待定位节点到锚节点间跳数来表示, 然后使用三边定位法或者最大似然估计法获得未知节点的位置信息。目前针对 Dv-Hop 改进算法的研究依然存在着如定位精度提高不明显, 增加了节点的通信开销、计算量、能量消耗等缺陷。

减法聚类<sup>[6-8]</sup> 是一种基于节点密度函数的方法,

收稿日期: 2011-03-15; 修回日期: 2011-06-22

基金项目: 湖北省高校重大项目基金 (Z200622002)

作者简介: 陈洁洁 (1982-), 女, 湖北黄石人, 硕士, 研究领域为无线传感器网络。

把每一个数据点作为一个潜在的中心,与数据的维数无关,所以不需要提前确定聚类的个数,直接根据数据的分布情况确定中心,同时与数据的个数成简单线性关系。因此,文中对定位算法进行了分析,提出了一种基于减法聚类的改进算法。

## 1 问题描述

由于文中的算法来源于 Dv-Hop,是美国路特葛斯大学的 Dragos Niculescu 等人提出的 Dv-Hop 算法,它分为以下 3 个阶段:

(1) 计算未知节点与锚节点的最小跳数。每个锚节点发出一个包括自己位置信息, ID 以及跳数数值为 0 的广播信号,在网络中被以泛洪的方式传播出去。每个锚节点的周围所有跳数为 1 的邻居节点接收该消息后,将对应的锚节点的位置和跳数记下来,同时将收到的信息包的跳数加 1,再向自己的邻居节点广播。这个过程一直持续到网络中所有的节点都获得与之相邻锚节点的位置信息和相应的跳数值,同时为了防止信息广播的无限循环,只有满足相应条件的信息才会被重新广播。

(2) 计算未知节点与锚节点的实际每跳的距离。只有在获得其它锚节点位置和相隔跳距之后,锚节点才开始计算网络平均每跳距离,然后将其作为一个校正广播到网络中。未知节点仅仅只记录接收到的第一个平均每跳距离,同时将其转发给邻居节点。未知节点接收到平均每跳距离后,根据记录相应的跳数,通过计算两者的乘积,就可以得到每个锚节点的距离。

(3) 未知节点利用第 2 阶段中计算到各个锚节点的距离,通过三边测量法或者极大似然估计法计算自身位置。

Dv-Hop 定位算法的主要误差来源是估算网络中平均每跳的距离,由于该未知节点到各个锚节点的节点分布情况大不相同,会导致较大的定位误差。考虑到锚节点密度,采用减法聚类的方法,基于锚节点自身节点密度选出锚节点的簇头,保证锚簇头节点在整个网络中合理均匀地分布,将所有锚簇头节点的平均跳数距离的均值作为未知节点网络平均每跳的距离。

## 2 无线传感器网络基于减法聚类定位算法

文中提出的算法主要是对上述 Dv-Hop 算法中第二阶段进行改进,根据锚节点周围的锚节点节点密度来计算该锚节点成为锚簇头的可能性,将具有最高的锚节点密度选为锚簇头节点,同时认定选出锚簇头节点附近的锚节点都不可能作为锚簇头。在第一个锚簇头节点被选出后,剩余可能作为锚簇头节点的锚节点继续采用类似的方法选择新的锚簇头节点。一直到所

有剩余的数据点作为锚簇头节点的可能性低于某个阈值则停止。改进的算法分为 4 个阶段:

- (1) 未知节点首先计算与锚节点的最小跳数;
- (2) 计算锚簇头节点;
- (3) 计算网络平均每跳的距离;
- (4) 通过极大似然估计法估计计算自身坐标。

其中第一、第四阶段和 Dv-Hop 算法过程一样,就不再描述了,下面分析新算法的第二和第三阶段。

### 2.1 锚簇头节点产生算法

采用减法聚类<sup>[6-8]</sup>方法计算锚簇头节点算法的具体步骤如下:

Step 1: 设无线传感器网络中  $n$  个坐标锚节点分别为  $y_i (i=1, 2, \dots, n)$ , 将输入所有的锚节点归一化到一个超立方体中,如公式(1)所示,每个节点都是锚簇头节点的候选者。

$$y_i = \frac{y_i - \min(y_i)}{\max(y_i) - \min(y_i)} \quad (1)$$

Step 2: 计算每个锚节点的密度指标:

$i = 1$

while ( $i \leq n$ )

{ 对应任意给定  $r_a > 0$ , 计算第  $i$  锚节点密度指标为:

$$D_i = \sum_{j=1}^n \exp\left(-\frac{\|y_i - y_j\|^2}{\left(\frac{r_a}{2}\right)^2}\right) \quad (2)$$

$i = i + 1;$

}

依次输出  $D_i (i = 1, 2, \dots, n)$

Step 3: 在 Step2 的结果里,选取最大的  $D_i = \max\{D_i | i = 1, 2, \dots, n\}$  为第一个锚簇头节点。

Step 4: 选取参数  $t_b$  大于  $t_a$ , 为了避免出现相近的锚簇头节点,一般取  $t_b = 1.5t_a$ 。

Step 5:  $y_{c1}$  为选中的锚簇头节点,  $D_{c1}$  为锚簇头节点的密度指标,那么其它每个锚节点的密度指标可以用下面的式子进行调整:

$$D_i = D_i - D_{c1} \exp\left(-\frac{\|y_i - y_{c1}\|^2}{\left(\frac{t_b}{2}\right)^2}\right) \quad (3)$$

Step 6: 找到第一个锚簇头节点同时修正  $n - 1$  个锚节点的密度指标,在其中找出密度指标最高的节点作为选定的新的锚簇头节点  $y_{c2}$ 。

Step 7: 对于任意给定的  $\partial$ , 且  $0 < \partial < 1$ 。以此来判断  $\frac{D_{c2}}{D_{c1}} < \partial$  是否成立,如果不成立,则转到 Step5 重新计算密度指标;如果成立则退出整个算法,那么就完成新算法中锚簇头节点选择。

### 2.2 计算平均每一跳的距离

Dv-Hop 定位算法<sup>[9,10]</sup>的第二阶段中,可以得知,该算法的误差主要是由于锚节点将自身的每跳平均距离广播到网络中,与此同时参与定位的未知节点使用最近点的每跳平均距离代替所有参与定位的锚节点的每跳平均距离,这个存在很大的误差。同时这个误差会带入循环进行累加,将会导致一些未知节点的定位精度误差过大。所以文中考虑将未知节点到每个锚节点的网络平均每跳距离等于所有锚簇头节点的每跳平均距离的均值。

当锚节点数量较少时,新算法的每个未知节点每一跳的距离,就是全网所有锚节点的平均每跳距离;当锚节点较多时,新算法的每个未知节点每一跳的距离,就是基于锚节点密度选出锚簇头节点的每跳平均距离的均值。

新算法通过第二阶段的锚簇头产生算法,可以得到锚节点  $n$  个簇头。 $n$  个锚簇头<sup>[11]</sup>节点(本身就是锚节点)距离已知,算出每个锚簇头节点的平均每跳距离。将  $n$  个锚簇头节点的平均每跳距离求均值,作为未知节点的每一跳的平均距离。

利用每个锚簇头节点  $(X_i, Y_i)$  从其它锚节点  $(X_j, Y_j)$  ( $j = 1, 2, \dots, n, j \neq i$ ) 收到的跳数信息,可以估计每个锚簇头节点平均每跳距离,并向整个网络广播,锚簇头节点  $i$  的平均每跳的距离为:

$$C_i = \frac{\sum_{j=1}^n \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2}}{\sum_{j=1}^n h_{ij}} \quad (4)$$

式中,  $h_{ij}$  为节点  $i$  和节点  $j$  之间的跳数。

求  $n$  个锚簇头节点平均每跳距离的均值:

$$C_{\infty} = \frac{\sum_{i=1}^n C_i}{n} \quad (\text{其中 } n \text{ 为锚簇头节点的个数})$$

将  $C_{\infty}$  值作为未知节点的校正值,并向整个网络广播,未知节点到每个锚节点的距离按照锚簇头节点均值  $C_{\infty}$  乘以相应的跳数来计算。

### 3 仿真实验和结果

为了考察该算法的性能,利用 Matlab 软件进行仿真实验,在相同的实验环境与 Dv-Hop 算法进行了比较,网络节点分布在区域尺寸为  $100\text{m} \times 100\text{m}$  的区域,锚节点和未知节点随机分布。仿真场景 1,锚节点总数为 50,所有的锚节点随机分布在  $100\text{m} \times 100\text{m}$  区域内。图 1 表示锚节点的原始的分布情况,图 2 表示通过减法聚类方法选出的锚簇头,用 \* 表示锚簇头,对比图 1 和图 2,在图 2 中锚节点根据自身密度不同生成了 5 个锚簇头节点。

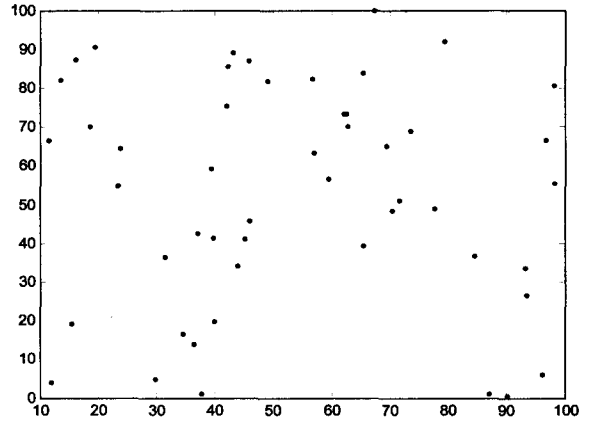


图 1 锚节点分布图

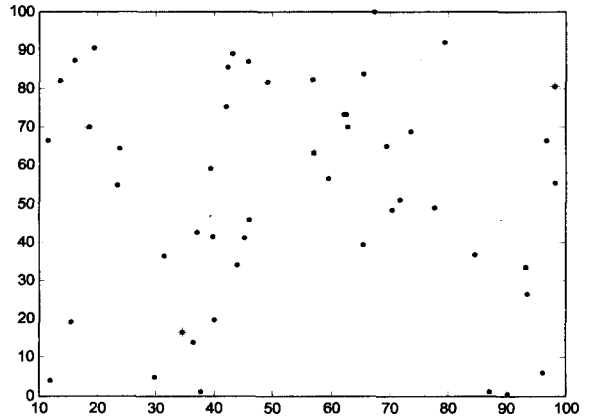


图 2 锚簇节点分布图

无线传感器网络的节点定位算法<sup>[12]</sup>评价指标包括精度、鲁棒性与可靠性等。新算法是基于网络平均每跳距离进行改进,主要是针对未知节点的定位误差进行的。同时用到的减法聚类是一种基于密度函数的方法,因此考虑以区域内包含锚节点的密度不同进行一组实验,来分析节点定位结果的精度的性能。

仿真场景 2 节点的分布区域为  $100\text{m} \times 100\text{m}$  的正方形,同时传感器中的锚节点和未知节点在仿真区域内随机分布,节点总数量 50 个,同时所有的节点具有相同的通信距离 10m。实验中将不同初始锚节点数量作为输入,分别选取 5, 10, 15, 20, 25 这五种初始锚节点个数进行对比。

图 3 ~ 图 5 分别是锚节点个数为 5、10、25 占区域所有节点的 5%、10%、50% 的每个未知节点误差图。从三个图中都可以看出,新算法中超过一半未知节点的误差比 Dv-Hop 算法误差小。新算法采用减法聚类的方法,使锚簇头节点在锚节点密集处产生,这种基于节点密度的锚簇头节点的产生方式保证锚簇头在整个网络中合理均匀分布,使未知节点每跳距离随网络中锚节点的密集度不同而不同,从而使未知节点每跳距离更加接近真实值,减少定位误差。同时当锚节点数量较少时,新算法的每个未知节点每一跳的距离,就是全网所有锚节点的平均每跳距离;当锚节点较多时,新

算法的每个未知节点每一跳的距离,就是基于锚节点密度的选出锚簇头节点的每跳平均距离的均值。

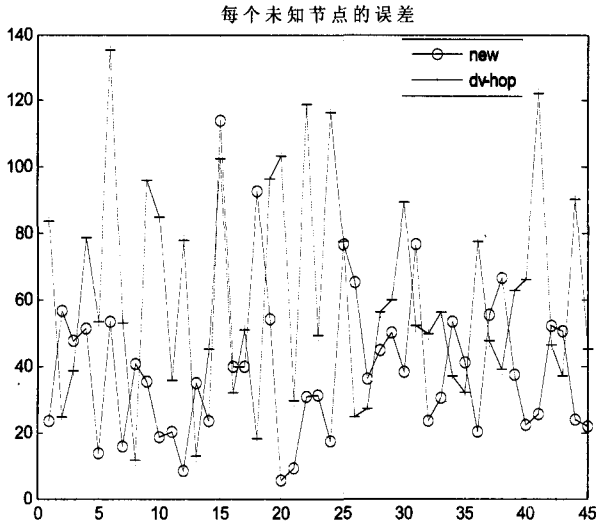


图 3 锚节点为 5,两种算法每个未知节点误差

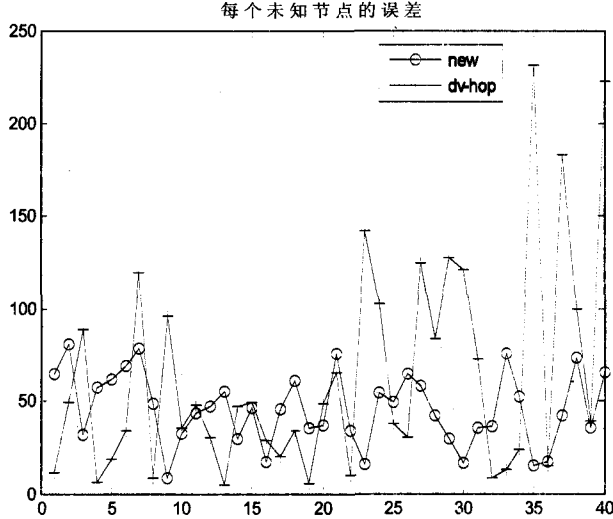


图 4 锚节点为 10,两种算法每个未知节点误差

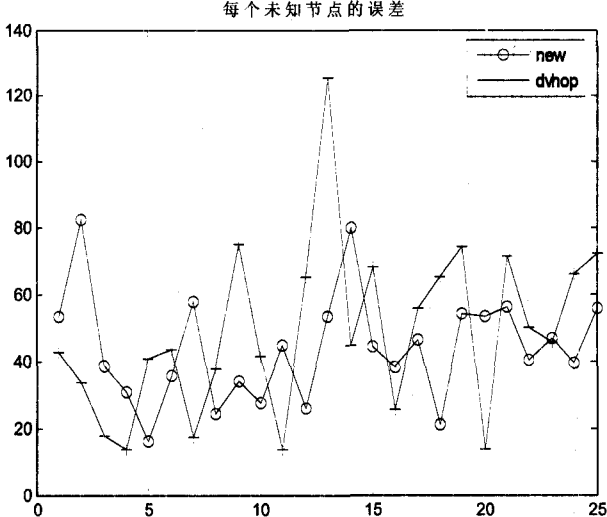


图 5 锚节点为 25,两种算法每个未知节点误差

图 6 是两种算法在不同锚节点密度下的整体误差,当锚节点的比率为 0.1:新算法的整体误差

51.0334, Dv - Hop 54.0180; 比率为 0.2:新算法 46.0104, Dv - Hop 49.4199; 比率为 0.3:新算法 36.8056, Dv - Hop 45.0267; 比率为 0.4:新算法 33.9300, Dv - Hop 43.9852; 比率为 0.5,新算法 32.3222, Dv-Hop 39.5021。画出图线如图 6 所示,当随着锚节点的数量增加,两种算法定位误差都在降低,新算法降低的趋势相对 Dv-Hop 显著些。由仿真结果可以看出,无论是锚节点密度低或者高的情况下,改进后的算法在改善定位精度上有着不错的表现,同时由于不需要计算所有锚节点的平均每跳的距离,只需要计算几个锚簇头节点的距离,减少了计算量和通信量,可以看出新算法能耗方面比原算法低。

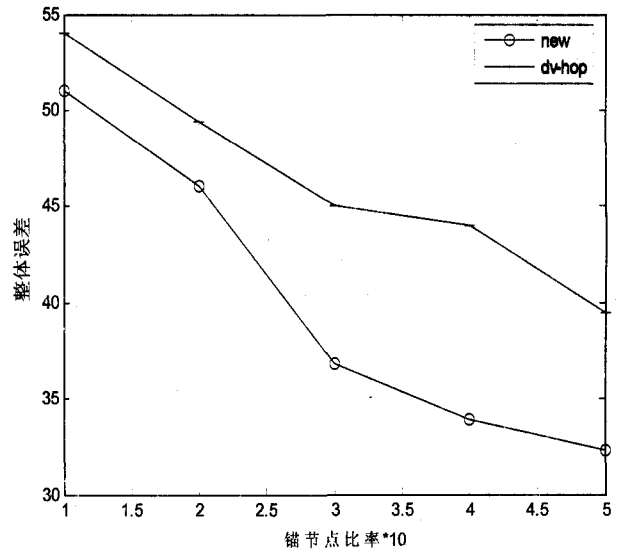


图 6 两种算法在不同锚节点的平均定位误差

### 4 结束语

文中考虑到锚节点密度,通过减法聚类的方法选出的锚簇头节点在整个网络中合理均匀地分布,同时将选出锚簇头节点的平均每跳距离的均值作为每个未知节点网络中平均每跳的距离,这种基于密度的锚簇头节点计算的平均每跳距离更接近与真实值,提高了定位精度。

### 参考文献:

- [1] 孙利民,李建中,陈渝,等.无线传感器网络[M].北京:清华大学出版社,2005.
- [2] He T, Huang C, Blum B, et al. Range - free localization schemes for large scale sensor networks [J]. MOBICOM, 2003,2(3):877-906.
- [3] Niculescu, Nath B. Ad - hoc positioning systems [C]//Proc. of the 2001 IEEE Global Telecommunications Conf. San Antonio:IEEE Communications Society,2001:2926-2931.
- [4] 候亚娜,胡维平.无线传感器网络中基于聚类平均的定位

```

<satellite: resolution>方位方向小于 30m</satellite: resolution>
<satellite: obliquity>98.2°</satellite: obliquity>
<satellite: semi-major-axis>7153.135km</satellite: semi-major-axis>
<satellite: flying-period>100.465 分钟</satellite: flying-period>
<satellite: breadth>98.2°</satellite: breadth >
</satellite>
</rdfs: RDF>

```

### 3 RDF 资源描述的基本实现方法

由上节卫星资源 RDF 描述的例子可以看出,RDF 描述资源的过程相当简单和直观,可概括为四点<sup>[12]</sup>:

- 第一,确定所要描述的资源及其各种属性;
- 第二,根据各种属性元素及其关系生成资源的 RDF 数据模型;
- 第三,根据资源的 RDF 数据模型再生成 RDF Schema;
- 第四,进一步生成资源描述的 RDF 文档。

运用 RDF 三元组原则,可以在任何需要的方面对资源进行任何描述,描述的方面越多,资源的描述信息就越丰富,在以后的实际应用中资源的利用率也就可能越高。如上卫星描述实例,当所有卫星资源经过上述的描述过程后,便拥有了一份完整的关于所有卫星资源的文档,这份文档中的每一项可以唯一标识一个特定的卫星资源,在浩瀚的军事网络中,可以根据这样的标识找到所需的卫星资源。这份卫星资源描述文档就好比一个索引一样,可以帮助我们管理卫星资源。

### 4 结束语

文中针对军事网格对资源描述的要求,提出了基于 RDF 的资源描述方法,并对 RDF 的定义、基本模型

以及 RDF Schema 进行了详尽的叙述,介绍了军事网格资源的定义和分类,并以军事网格中的卫星资源为例,进行了数据建模,生成了资源描述文档,阐述了 RDF 资源描述的基本实现方法,证明该方法对描述网格信息资源具有针对性和普遍性,适合描述网格资源,为后面的资源管理打下了良好的基础。

#### 参考文献:

- [1] Foster I. What is grid? A Three Point Checklist[EB/OL]. 2002. <http://www-fp.mcs.anl.gov/~foster/Articles/WhatIsTheGrid.pdf>. 2002.
- [2] 刘洪川,王廷申,汪艳伟. 未来的军事网格技术[J]. 国防科技,2008(2):7-10.
- [3] 朱莹,吴君华,汪婷婷,等. 网格资源描述技术的比较研究[J]. 微计算机信息,2006,22(4-3):176-178.
- [4] Manola F, Miller E. RDF Primer[EB/OL]. 2004-02. <http://www.w3.org/TR/rdf-primer>.
- [5] Florescu D, Kossman D. Storing and Querying XML Data Using a RDBMS[J]. IEEE Data Engineering Bulletin,1999,22(3):27-34.
- [6] 苏天醒. 基于 RDF 的信息网络服务平台资源管理的研究[D]. 北京:北京邮电大学,2009:14-24.
- [7] 任磊. 基于 RDF 元数据的网格资源发现模型研究[D]. 包头:内蒙古科技大学,2009:19-29.
- [8] RDF Vocabulary Description Language 1.0:RDFSchema[EB/OL]. 2004-02. <http://www.w3.org/TR/rdf-schema>.
- [9] 刘鹏,王立华. 走向军事网格时代[M]. 北京:解放军出版社,2004:26-32.
- [10] 任磊. 基于 RDF 元数据的网格资源发现模型研究[D]. 包头:内蒙古科技大学,2009:19-29.
- [11] 李菊芳,贺仁杰,姚峰. 卫星观测任务数据的 XML 描述[J]. 计算机工程与应用,2004(32):209-211.
- [12] 范晓静,姜月秋. 信息栅格信息资源描述分析[J]. 数字技术与应用,2009(10):83-84.

(上接第 128 页)

- 算法[J]. 计算机应用研究,2010,27(4):4446-4448.
- [5] 王珊珊,殷建平,张国敏,等. 一种无线传感器网络定位问题中分簇算法[J]. 计算机应用研究,2008,35(8):8346-8348.
- [6] 陈洁洁,樊晓平,瞿志华,等. 基于减法聚类的无线传感器网络分簇路由算法[J]. 信息与控制,2008,37(4):435-438.
- [7] Chiu S L. Fuzzy model identification based on cluster estimation[J]. Journal of Intelligent and Fuzzy Systems,1994,2(3):267-278.
- [8] 崔连延,徐林. 基于减法聚类与改进的模糊 C-均值聚类算法的说话人识别方法的研究[J]. 信息与控制,2010,37(3):318-321.
- [9] Liu Chong, Wu Kui, He Tian. Sensor Localization with Ring Overlapping Based on Comparison of Received Signal Strength Indicator[C]//Proceedings of Mobile Ad-hoc and Sensor Systems,2004 IEEE International Conference. Fort landerdale, Florida, USA:IEEE Press,2004:516-518.
- [10] 杨永雷,朱军. 无线传感器网络中异步成簇算法的研究[J]. 计算机技术与发展,2010,20(2):145-147.
- [11] 张强,卢潇,崔晓层. 基于分簇的无线传感器网络数据聚合方案研究[J]. 传感技术学报,2010,23(12):1778-1782.
- [12] 王书聪. 无线传感器网络分布式定位算法[J]. 计算机技术与发展,2008,18(11):62-65.