

文章编号:1674-2974(2016)10-0134-05

## 传感器网络中基于抽样的 带权近似 Top-k 查询算法\*

刘彩苹<sup>1</sup>, 蔡玉武<sup>1</sup>, 毛建旭<sup>2†</sup>, 龙亚辉<sup>1</sup>

(1. 湖南大学 信息科学与工程学院, 湖南 长沙 410082; 2. 湖南大学 电气与信息工程学院, 湖南 长沙 410082)

**摘要:**提出一种适用于传感器网络的抽样带权阈值过滤近似 Top-k 聚集查询算法。该近似算法会将无线传感器网络划成几个两两不相交的簇进行处理,在汇聚节点进行预处理以及在各个簇内进行抽样过滤处理,在抽样过程中给可靠而重要的节点赋上相应更大的权值,同时根据节点采集的信息具有时间相关特性,在簇内进行抽样阈值过滤处理,每个簇头节点都会接收到该簇内的 Top-k 候选子集,然后将每个簇的子集发送给 Sink 节点,该 Sink 节点将接收到能代表整网 Top-k 样本候选集。仿真实验结果显示该算法只需发送少量的数据,更小的抽样样本,并能满足任意精度要求。

**关键词:**无线传感器网络;抽样算法;Top-k 查询

**中图分类号:**TP212.9

**文献标识码:**A

## Research on the Approximate Algorithm of Top-k Query Based on Weighted Sampling in Wireless Sensor Network

LIU Cai-ping<sup>1</sup>, CAI Yu-wu<sup>1</sup>, MAO Jian-xu<sup>2†</sup>, LONG Ya-hui<sup>1</sup>

(1. College of Computer Science and Electronic Engineering, Hunan Univ, Changsha, Hunan 410082;

2. College of Electric and Information Engineering, Hunan Univ, Changsha, Hunan 410082)

**Abstract:** An approximate algorithm of Top-k query based on sampling and weight in wireless sensor network was presented. The algorithm divides the network into several disjoint clusters in the sink node and the nodes in cluster to take sampling process. In the process of sampling, greater weight for reliable and important sensor node is given. The sensor node sensing data has a time correlation, and sampling threshold filtering in the cluster. Each cluster head node receives a Top-k candidate subset of the cluster, and then sends the subset to the sink node. Finally, the sink node can receive a Top-k sample candidate that represents the whole network. Simulation experiments show that the algorithm only needs to send small data and smaller samples, and can satisfy arbitrary precision requirements.

**Key words:** wireless sensor networks; sampling algorithm; Top-k query

近年来,随着信息技术的快速发展,物联网时代 已经悄悄向我们走来,无线传感器网络是物联网技

\* 收稿日期:2015-06-27

基金项目:国家自然科学基金资助项目(61370096,61573134), National Natural Science Foundation of China(61370096,61573134); 国家科技支撑计划资助项目(2015BAF13B00); 湖南省科技计划资助项目(2012GK3158)

作者简介:刘彩苹(1978-),女,湖南邵阳人,湖南大学讲师,博士

† 通讯联系人, E-mail: maojianxu@hnu.edu.cn

术中关键技术之一。该技术广泛使用在现代化信息农业<sup>[1]</sup>、矿井智能化探测开采<sup>[2]</sup>和智能家居<sup>[3]</sup>等方面。传感器网络是由许多廉价的微型节点组织而成,可以在其监测范围内经由路由算法自组织成一个网络。用户在网络中会进行聚集查询处理,而 Top-k 查询是最常见的操作之一,具有非常大的实际意义,例如:用户在进行空气质量监测时,甲需要了解 PM2.5 值最大的  $k$  个值,乙需要了解空气质量指数最大的  $k$  个值,有时甲和乙对查询的精度标准和要求不一样,需要设计出能适应不同用户查询精度的 Top-k 聚集查询处理算法以便来满足多用户的实际应用需求。

由于传感器网络中的节点通信范围、计算处理、存储容量和能量大小都非常有限,聚集查询算法第一要务就要考虑节能,最大化网络的寿命。节点能量耗尽而失效,网络拓扑结构随时发生变化,而且节点在发送数据丢包和通信连接失败时,就会破坏生成的路由树,在很多情况下,传感器网络无法得到用户精确的查询分析结果。近年来,许多学者提出了许多能在传感器网络中进行近似查询的算法,近似算法能减少节点数据的发送量,节约节点的能量,最大化提高网络的寿命。

文献[4]提出一种垂直数据处理的 Top-k 算法。算法的主要思想是生成路由树,在路由树中进行 Top-k 查询,同时采用历史数据进行处理。文献[5]使用位图压缩机制减少节点间数据的发送量。节点能量和存储空间,但是恶意节点利用桶的信息可以估计出查询结果。文献[6]提出了一种近似 Top-k 聚集查询算法。该算法对传感器节点感知的数据进行抽样,并用线性模型得到满足用户精度要求的近似查询结果。文献[7]利用传感器节点感知数据的时空相关性。该算法采用综合样本抽样和数据压缩技术进行聚集查询。文献[8]提出了连续  $k$  近邻查询算法。算法的核心思想是采用环查询和变量维护技术减少节点数据的发送量。

文献[9]使用了一种概要查询处理技术。与其它查询算法不同的是在查询处理过程中只需要传输概要信息,而不需要传感器节点的感知值,这种技术可以减少节点的能量开销。文献[10]利用传感器网络的时间和空间相关性原理进行近似查询处理,减少节点数据的发送,降低能量的消耗。

实际上,节点在放置时往往出现分布不均匀的情况。目前的 Top-k 聚集查询算法有时并不能反映监测区域真实情况,例如下面情况:在环境污染监测的区域,越靠近人类居住或者水源的区域越重要,可

能这些区域的空气质量指数并不是很高,但是它已经对人们的健康产生了影响,需要给用户报警提示。还有传感器网络监测某一小区的噪声大小,查询者需要了解噪声大小会影响居民最高的  $k$  个点,在监测范围内,离居民比较近的点噪声分贝值虽然不是最大的,但是它对居民的影响超出了远处分贝值比它高的点。为了满足实际的需要,可以对重要的区域增加权值的参数,扩大它的作用效果,从而更加能够反映传感器网络监测区域的真实情况<sup>[11]</sup>。

传感器网络节点感知的数据具有时间和空间相关性,可以利用网络的历史查询信息估算出一个抽样阈值,在每个簇内进行抽样过滤处理时,只有大于阈值的感知值才会被发送给簇头节点,从而能够减少节点不相关信息的发送,节约能源,提高网络的寿命。

## 1 带权近似 Top-k 查询定义

如果在网络进行聚集查询时,网络是由  $N$  个节点组成,则把节点的集合记为  $I = \{1, 2, 3, \dots, N\}$ 。传感器网络存在一个能量和处理能力都强大的汇聚节点,根据路由算法以汇聚节点为根生成最小生成树,在进行聚集查询时汇聚节点会将用户的请求分发给所有节点。在  $t$  时刻,网络中各节点感知到的数据集合为  $D_i = \{d_1, d_2, \dots, d_N\}$ ,  $d_i = \{d_i.v, d_i.id, d_i.t\}$  其中  $d_i.v$  是节点的物理感知值,  $d_i.id$  是标识节点的唯一 ID,  $d_i.t$  是感知数据时的时间戳。

**定义 1**  $k$  是自然数,且  $|I| \geq k$ , 在  $t$  时刻时,数据集合  $D_i$  的带权 Top-k 集合记为  $W - \text{Top}(k, D_i)$ , 是  $I$  的一个真子集,同时  $W - \text{Top}(k, D_i)$  要符合以下要求:

- 1)  $W - \text{Top}(k, D_i) \subset I$ ;
- 2)  $|W - \text{Top}(k, D_i)| = k$ ;
- 3) 对于  $\forall i \in W - \text{Top}(k, D_i), \forall j \in I \setminus W - \text{Top}(k, D_i)$  均满足  $w_i d_i \geq w_j d_j$ 。

其中  $W = \{w_1, w_2, \dots, w_N\}$  是网络中节点赋与相应的权值,  $w_i \geq 1, 1 \leq i \leq N$ 。

**定义 2** 对  $\forall \epsilon > 0, D_i$  的近似  $W - \text{Top}(k, D_i)$  记为  $\epsilon - W - \text{Top}(k)$ , 其中,  $\epsilon - W - \text{Top}(k)$  要符合以下要求:

- 1)  $\epsilon - W - \text{Top}(k) \subset I$ ;
- 2)  $|\epsilon - W - \text{Top}(k)| = k$ ;
- 3) 设  $W - \text{Top}(k, D_i) = \{i_1, i_2, \dots, i_k\}$ , 满足条件  $w_{i_1} d_{i_1} \geq w_{i_2} d_{i_2} \geq \dots \geq w_{i_k} d_{i_k}$ ,  $\epsilon - W - \text{Top}(k) = \{j_1, j_2, \dots, j_k\}$  并且  $w_{j_1} d_{j_1} \geq w_{j_2} d_{j_2} \geq \dots \geq w_{j_k} d_{j_k}$ . 那么  $\max \left\{ \frac{w_{i_m} d_{i_m}}{w_{j_m} d_{j_m}} \mid 1 \leq m \leq k \right\} \leq 1 + \epsilon$ .

如果存在  $m$  使得  $1 + \epsilon \geq \frac{\omega_{im}d_{im}}{\omega_{jm}d_{jm}}$ , 则  $\omega_{jm}d_{jm} \geq \frac{\omega_{ij}d_{ij}}{1 + \epsilon}$ , 故  $\frac{\omega_{im}d_{im} - \omega_{jm}d_{jm}}{\omega_{im}d_{im}} \leq \frac{\epsilon}{1 + \epsilon}$ . 由  $\epsilon - W - \text{Top}(k)$  的形式化定义知, 近似查询的结果相对误差不大于  $\frac{\epsilon}{1 + \epsilon}$ .

**定理 1**<sup>[12]</sup> 独立同分布的中心极限定理: 设随机变量  $X_1, X_2, \dots, X_n, \dots$  相互独立, 并且服从同一分布, 并且随机变量具有数学期望为  $E(X_k) = \mu$ , 方差为  $D(X_k) = \sigma^2 > 0 (k = 1, 2, \dots)$ , 则进行 Top-k 抽样分布近似地服从参数为  $\mu_x = \mu, \sigma_x^2 = \frac{\sigma^2}{n}$ ,  $\sigma_x = \frac{\sigma}{\sqrt{n}}$  的正态分布, 3 个参数分别表示随机样本的方差、均值和标准差.

**定理 2** 将无线传感器网络划分成  $r$  个两两不相交的簇, 设  $i = 1, 2, \dots, r, q_1 = \frac{\text{Max}(D_i)W_{\max 1}\varphi_{\delta/2}^2}{\text{NWMin}(D_i)\epsilon^2 + \text{Max}(D_i)W_{\max 1}\varphi_{\delta/2}^2}$ , 则  $\frac{q_i}{q_1} = \frac{W_{\max i}}{W_{\max 1}}$ , 其中:  $W_{\max i}, W_{\min i}$  分别表示网络第  $i$  个簇中所有节点权值的最大值和最小值,  $\text{Max}(D_i), \text{Min}(D_i)$  分别是传感器网络中感知数据集的最大值和最小值;  $\varphi_{\delta/2}^2$  是标准正态分布的  $\delta/2$  临界值;  $\bar{W} = \text{Sum}(W)/N$ .

**证明** 汇聚节点首先得到传感器网络中的最大和最小权值  $W_{\max}, W_{\min}$  和每个节点权值的和. 在进行 Top-k 查询时, 汇聚节点会根据  $q_1 = \frac{\text{Max}(D_i)W_{\max 1}\varphi_{\delta/2}^2}{\text{NWMin}(D_i)\epsilon^2 + \text{Max}(D_i)W_{\max 1}\varphi_{\delta/2}^2}$  计算出  $q_1$  的值, 并且汇聚节点会将  $q_1$  和最大权值  $W_{\max}$ , 最小权值  $W_{\min}$  广播到网络中各簇的簇头节点, 每个收到广播的传感器簇头节点  $i$  根据簇内权值的最大值  $W_{\max i}$  计算出该簇相应的抽样概率. 即  $q_i = \frac{q_1 W_{\max i}}{W_{\max 1}}, i = 1, 2, \dots, r$ .

**定理 3** 存在关系  $S(p, \epsilon, \delta) = \Theta(\frac{1}{p\epsilon} \cdot \ln \frac{1}{p\delta})$ , 对于任意的输入  $X \in \chi(W, n, p)$ , 对  $X$  进行误差为  $\epsilon$  的 Top-k 近似查询时抽样的样本大小至少为  $S(p, \epsilon, \delta)$ , 并且概率至少为  $1 - \delta$ .

**证明** 如果  $i \in \text{top}(k, X), j \notin \text{top}(k, X)$ , 并且  $p_X(i) > (1 + \epsilon) \cdot p_X(j)$ , 则称  $(i, j)$  对是相关联的.

在样本  $S$  中, 如果  $(i, j)$  对是相关联的并且满足  $p_S(i) \leq p_S(j)$ , 则  $(i, j)$  对是弱交换对,  $\epsilon - W - \text{Top}(k, S)$  是  $\text{Top}(k, S)$  误差为  $\epsilon$  的近似值当且仅

当在  $S$  中没有弱交换对.

对任意  $i \in \text{top}(k, X)$  定义阈值  $t_i$  满足  $\frac{p_X(i)}{1 + \epsilon} < t_i < p_X(i)$ , 对任意  $j \notin \text{top}(k, X)$  定义阈值  $t'_j = \min \{t_i \mid (i, j) \text{ 是相关联的} \}$ , 记  $E_i$  是  $S$  中满足  $p_S(i) \leq t_i$  的随机事件, 也是  $t_i$  阈值补集的随机事件,  $E'_j$  是样本  $S$  中满足  $p_S(j) \geq t'_j$  的随机独立事件, 也是  $t'_j$  阈值补集的随机独立事件.

从上面阈值的属性可知, 如果  $i \in \text{top}(k, X)$  并且不是  $t_i$  阈值补集,  $j \notin \text{top}(k, X)$  并且不是  $t'_j$  阈值补集, 则在  $S$  中不存在弱交换对.

## 2 基于抽样带权近似过滤 Top-k 查询处理算法

基于抽样的带权近似过滤 Top-k 查询处理算法所涉及参数定义如下所示:

$D_t = \{d_1, d_2, \dots, d_N\}$ , 表示在  $t$  时刻, 传感器网络中各节点感知到的数据集合,  $d_i = \{d_i.v, d_i.id, d_i.t\}$  中  $d_i.v$  是节点的物理感知值,  $d_i.id$  是标识节点的唯一 ID,  $d_i.t$  是感知数据时的时间戳.

$W - \text{Top}(k, D_t)$ , 表示在  $t$  时刻时, 数据集合  $D_t$  的带权 Top-k 集合.

$W = \{\omega_1, \omega_2, \dots, \omega_N\}$  表示由可信度赋与节点的权值,  $\omega_i \geq 1, 1 \leq i \leq N$ .

$\epsilon$ , 且  $\epsilon \geq 0$ , 相对误差.

$\delta$ , 且  $0 \leq \delta < 1$ , 失败概率, 即近似聚集查询的结果误差不大于  $\frac{\epsilon}{1 + \epsilon}$ .

$q_i$ , 表示第  $i$  个簇的抽样概率, 其中  $i = 1, 2, \dots, r$ .

$H(k)$  节点抽样时的过滤阈值.

基于抽样的带权阈值过滤近似 Top-k 聚集查询算法如下:

**输入**

1) Sink 节点存有网络中所有节点的权值, 其中  $W_{\max}$  和  $W_{\min}$  为最大最小权值.

2)  $t$  时刻网络中节点采集到的数据集合  $D_t = \{d_1, d_2, \dots, d_n\}$ .

3) 样本容量  $S$ .

**输出**  $\epsilon - W - \text{Top}(k)$  候选集

算法详细步骤如下所示:

i) 将节点的所有权值均保存在 Sink 节点上, Sink 节点根据定理 1 计算出  $q_i$  的概率值. 并由  $(\epsilon, \delta)$  和以前查询统计数据估算本次抽样的样本大小, 记为  $S$ .

ii) 根据传感器网络中时间相关性原理, 汇聚节点由  $\epsilon - W - \text{Top}(k)$  的历史查询信息计算节点在抽样时的过滤阈值  $H(k)$  来筛选数据以便减少网络中数据的发送量.

$$H(k) = \frac{W - \text{Top}(k)^{t-1} + W - \text{Top}(k)^{t-2} + \dots + W - \text{Top}(k)^{t-d}}{d(1+\epsilon)},$$

式中  $W - \text{Top}(k)^{t-1}$  表示抽样算法返回  $t-1$  时刻  $\epsilon - W - \text{Top}(k)$  集合排列时在集合中第  $k$  位的值,  $d$  的大小可以根据实际情况设置.

iii) 汇聚节点给第  $R_i$  个簇头节点发送信息  $(q_i, W_{\max}, W_{\min}, S, H(k))$ ,  $1 \leq i \leq r$ . 第  $R_i$  个簇头节点接收到汇聚节点发送过来的信息后就进行簇内节点抽样算法.

4) 汇聚节点最终将会收到所有簇头节点发送回来的信息  $\epsilon - W - \text{Top}(k)$  候选集的子集.

### 3 簇内抽样带权阈值过滤近似 Top-k 聚集查询算法

1) 假如设第  $R_i$  个簇的编号集合为  $\{R_{i1}, R_{i2}, \dots, R_{ir_i}\}$ . 当  $R_i$  簇的簇头节点接收到汇聚节点发送的  $(q_i, W_{\max}, W_{\min}, S, H(k))$ ,  $1 \leq i \leq r$  时, 簇头节点  $i$  会根据簇内权值的最大值  $W_{\max i}$  计算出该簇相应的抽样概率,  $q_i = \frac{q_1 W_{\max i}}{W_{\max 1}}, i = 1, 2, \dots, r$ .

2) 簇头节点会根据抽样概率  $q_i$  计算第  $R_i$  个簇的样本大小是  $S * q_i$ , 并记为  $S_i$ . 簇头节点随机独立均匀地产生  $S_i$  个属于  $\{R_{i1}, R_{i2}, \dots, R_{ir_i}\}$  的节点 ID 集合, 并且把 ID 集合为  $M_i$ . 然后簇头节点会将集合  $M_i$  及  $H(k)$  值传送给该簇内的所有节点.

3) 若  $j \in M_i$  且节点  $j$  的感知信息满足  $d_j.v \geq H(k)$ , 则把节点  $j$  的感知信息传送给簇头节点, 否则节点  $j$  不向簇头节点传送数据, 最终簇头节点会收集到该簇内所有抽样的数据  $\epsilon - W - \text{Top}(k) - \text{Sub}_i$ .

由上面的分析可知, 基于抽样带权阈值过滤近似 Top-k 聚集查询算法如下:

The algorithm of  $\epsilon - W - \text{Top}(k)$

Input:  $W, S, D_i, \epsilon, \sigma$

Output: The set of  $\epsilon - W - \text{Top}(k)$

The sink get  $q_i$  and  $H(k)$  by theorem 1 and history data

For any cluster head  $i, i=1, 2, \dots, r$

{

$$q_i \leftarrow \frac{q_1 W_{\max i}}{W_{\max 1}} \quad // \text{计算第 } i \text{ 个簇的抽样概率}$$

$$S_i \leftarrow S * q_i$$

```
//簇头节点随机独立均匀地产生  $S_i$  个属于  $\{R_{i1}, R_{i2}, \dots, R_{ir_i}\}$  的节点 ID 集合
 $M_i \leftarrow \{R_{i1}, R_{i2}, \dots, R_{ir_i}\}$ 
 $j \leftarrow 0$ 
while  $j < |M_i|$  do
  if  $(d_j.v \geq H(k))$ 
     $\epsilon - W - \text{Top}(k) - \text{Sub}_i \leftarrow d_j.v$ 
   $j \leftarrow j + 1$ 
end while
}
```

return  $\epsilon - W - \text{Top}(k)$

节点发送所有权值的时间复杂度为  $O(n)$ , 簇内节点抽样算法的时间复杂度为  $O(rR_i)$ , 故最差时间复杂度为  $O(n^2)$ .

### 4 实验仿真及结果分析

在伯克利分校研究者研发的 TAG<sup>[13]</sup> 平台上进行实验. 仿真实验中的感知数据来自于 Berkeley Intel 实验室传感器网络监测真实环境时得到的温度数据.

选取简单的 Top-k 抽样近似算法和本文的基于抽样带权阈值过滤近似 Top-k 聚集查询算法进行对比实验.

简单 Top-k 抽样近似查询算法由 Sink 节点随机独立地产生样本大小为  $S$  的样本, Sink 节点将样本  $S$  广播到网络中进行抽样, 如果节点编号在抽样样本中, 就将该节点的信息发送给汇聚节点, 汇聚节点收到数据后输出前  $k$  个最大值作为 Top-k 查询的近似结果. 而带权近似 Top-k 查询抽样算法只需抽样本簇中少量数据发送给簇头节点, 大大减少了数据的发送量, 节省了网络中节点的能量. 图 1 可以说明这种变化趋势.

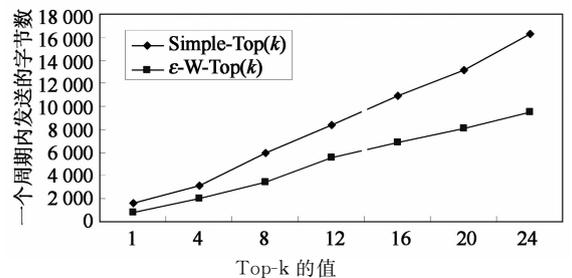


图 1 一个周期内不同 Top-k 值下节点发送的字节  
Fig. 1 Bytes sent by nodes in a cycle with different Top-k values

选取简单 Top-k 簇内抽样近似查询算法和本文带权过滤簇内抽样算法在不同的网络规模下应当选取样本容量的大小关系如图 2 所示.

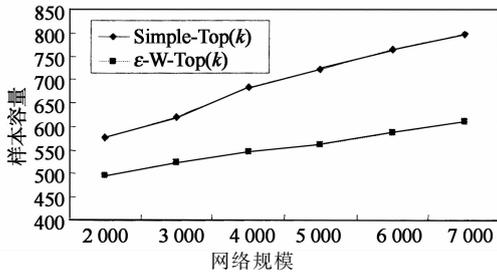


图2 不同网络规模下样本  $S$  大小的比较  
Fig. 2 Comparison of the sample size of  $S$  under different network sizes

在  $(\epsilon, \delta)$  和  $k$  一定的条件下, 简单 Top-k 抽样近似查询算法进行查询时随机独立地产生样本大小为  $S$  的样本. 而抽样带权近似过滤 Top-k 查询处理算法通过给定的  $(\epsilon, \delta)$  和历史信息估计确定样本的容量  $S$ , 在查询过程中根据权值确定查询概率计算每个簇内的抽样样本以及依据历史信息过滤抽样的样本值, 得到能符合  $(\epsilon, \delta)$  精度的近似 Top-k 查询数据. 从图 2 可以看出, 在  $(\epsilon, \delta)$  和  $k$  确定的条件下, 带权近似 Top-k 查询抽样算法在同样网络规模大小下样本容量比简单 Top-k 抽样近似查询算法都要小.

给定  $k$  值进行带权近似 Top-k 查询抽样算法时, 抽样的样本容量与用户查询的精度大小和网络规模有着密切的关系, 网络中节点的数量越多, 查询精度要求越低, 抽样的样本大小就越大, 从图 3 的实验结果可以知道这种变化趋势.

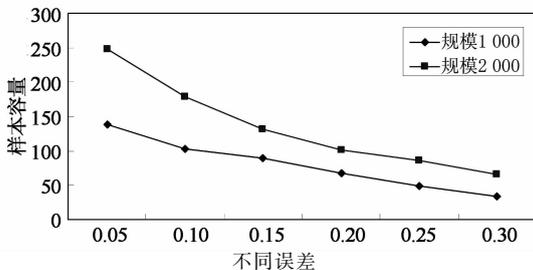


图3 不同误差情况下样本容量大小

Fig. 3 Sample size under different error conditions

## 5 结 语

本文提出一种适用于传感器网络的抽样带权阈值过滤近似 Top-k 聚集查询算法. 与其它 Top-k 查询处理算法不同, 本算法进行分簇抽样处理, 并根据实际情况给节点加入了权值参数, 同时每个簇内进行抽样, 根据历史信息过滤掉不必要的的数据, 每个簇头节点会将抽样子集传递给汇聚节点, 最终汇聚节点收集到全网的满足用户查询精度的 Top-k 集合. 带权近似 Top-k 查询抽样算法可以有效地减少

节点不相关信息的发送, 节约能量, 提高网络的生命周期. 与普通抽样算法相比, 只需要更少的样本容量即可, 同时可以满足不同用户对查询精度不同的请求. 所以, 带权近似 Top-k 查询抽样算法适用于注重节约节点能量的无线传感器网络, 同时能满足多用户的精度要求, 以及 Top-k 查询的近似结果可以反应传感器网络监测区域的真实情况.

## 参考文献

- [1] BARBAGLI B, BENCINI L, MAGRINI I, *et al.* A real-time traffic monitoring based on wireless sensor network technologies[C]//Proceedings of the 7th International Wireless Communications and Mobile Computing Conference, Istanbul, Turkey: IEEE Computer Society, 2011:820-825.
- [2] ZHANG F, DISANTO W, REN J, *et al.* A novel cps system for evaluating a neuralmachine interface for artificial legs[C]//Proceedings of IEEE/ACM International Conference on Cyber-Physical Systems. Chicago, USA: IEEE Computer Society, 2011:67-76.
- [3] BOCCA M, TOIVOLA J, ERIKSSON M, *et al.* Structural health monitoring in wireless sensor networks by the embedded goertzel algorithm[C]//Proceedings of IEEE/ACM International Conference on Cyber-Physical Systems. Chicago: IEEE Computer So. Ciety, 2011:206-214.
- [4] CHU D, DESHPANDE A, M. HELLERSTEIN J M, *et al.* data collection in sensor networks using probabilistic models [C]//In Proceedings of the 22th International Conference on Data Engineering. Georgia, USA; April 3-7, 2006; 234-251.
- [5] CAO Q, ABDELZAHER T, HE T, *et al.* Towards optimal sleep scheduling in sensor networks for rare event detection [C]//Proceedings of IPSN. CA, USA; IPSN, 2005; 20-27.
- [6] KOUSHANFAR F, TAFT N, POTKONJAK M. Sleeping coordination for comprehensive sensing using isotonic regression and domatic partitions [C]//Proceedings of IEEE Infocom. Barcelona, Spain; 2006;45-58.
- [7] LI J, LI Z. Data sampling control, compression and query in sensor[J]. International Journal of Sensor Networks, 2014, 2 (1/2): 53-61.
- [8] YAO Yu-xia, TANG Xue-yan, LIM Ee-peng. Localized monitoring of knn queries in wireless sensor networks[J]. The VLDB Journal, 2014, 18(1):99-117.
- [9] NASRIDINOV A, PARKY H. Optimal aggregator node selection in wireless sensor networks[C]//Proceedings of ICCA. Seoul, Korea; ICCA, 2013. ASTL, 2013; 37-39.
- [10] DELIGIANNAKIS A, PROCESSING Y K. Approximate aggregation queries in wireless sensor networks[J]. Information Systems, 2013, 31(8):770-792.
- [11] BI Ran, LI Jian-zhong, CHENG Si-yao. Approximate Top-k query processing algorithm in wireless sensor networks[J]. Journal on Communications, 2011, 32(8):45-54.
- [12] BEMSEIN S, BERNSTEIN R. Elements of statistics II: descriptive statistics[M]. Newyork, USA McGraw-Hill, 2004.
- [13] MADDEN S, FRANKLIN M J, HELLERSTEIN J M. TAG: a tiny aggregation service for Ad-Hoc sensor networks[C]//Symposium on Operating Systems Design and Implementation. Boston, MA, 2002;131-146.