

基于 DS 证据理论和压缩感知的 WSN 数据融合策略设计

廖长荣

(重庆文理学院 电子电气工程学院, 重庆 402160)

摘要: 为了降低 WSN 数据量和延长网络生命周期, 设计了一种基于 DS 证据理论和压缩感知的 WSN 混合数据融合策略; 首先, 在分簇协议的基础上引入了基于 DS 证据理论和压缩感知的混合模型, 然后, 采用改进的 DS 对所有簇成员节点的基本信度分配函数进行加权处理, 在簇头处采用加权和归一化的信度分配函数计算证据对各命题的支持程度, 将支持程度较大的若干命题作为 DS 融合结果, 在此基础上采用压缩感知方法通过构造测量矩阵对融合结果进行稀疏化表示, 并在基站处对稀疏信号进行重构; 仿真实验表明, 文中方法能有效地实现数据融合, 且和其他方法相比, 具有重构误差较小和网络生命周期较长的优点, 具有较大的优越性。

关键词: 压缩感知; 证据理论; 无线传感器网络; 数据融合

Design of Data Fusion Strategy in Wireless Sensor Network Based on DS Evidence Theory and Compressing Sensing

Liao Changrong

(School of Electric & Electrical Engineering, Chongqing University of Arts and Sciences, Chongqing 402160 china)

Abstract: In order to reduce the amount of data and delay the network life cycle, a compound data fusion strategy based on DS evidence and compressing sensing was introduced based on the clustered protocol, then the improved DS is used to weight over the basic credit assigning function, the weight and normalized credit assign function is used to compute the support extent of evidence to proposition, using the larger value of support extent as the fusion result of DS, then using the CS and constructing the measure matrix to realize the sparse representing, and the construction for sparse signal is used in sink node. The simulation result shows the method in this paper can realize the data fusion effectively, and compared with the other methods, it has less construction error and long life cycle, so it has big priority.

Keywords: compressing sensing; DS theory; wireless sensor network; data fusion

0 引言

无线传感器网络 (wireless sensor networks, WSN)^[1]是由多门学科高度交叉组成的新兴学科, 目前已经在国际上受到广泛的关注。

近年来, WSN 的发展受到能量供应、数据存储、数据传输能力、数据处理能力和网络的可靠性等因素的制约和影响, 减少 WSN 数据传输的距离和数据量是提高 WSN 生命周期的关键^[2-3]。

数据融合是通过某种方法将 WSN 中大量传感器采集的原始信息进行有效处理, 去除其中的冗余信息, 从而有效地延长网络生命周期^[4]。

目前已有的数据融合方法主要可以分为: 基于生成树的数据融合方法^[5-6]、消除时空相关性的数据融合^[7-8]、基于预测的数据融合^[9]。基于生成树的数据融合方法存在的问题主要是当同时发送数据的传感器节点数量过多时, 会造成网络拥塞和能耗过高, 反而降低了数据包的传输速率。基于消除时空相关性的数据融合的方法计算量过大, 节点融合时带来的时延会影响

响系统的实时性和同步性。基于预测的数据融合方法具有融合效率不高的缺点。

上述工作具研究了 WSN 的数据融合方法, 但其往往在传感器节点上进行完全融合, 间接地增加了传感器节点的开销, 为此, 本文设计了一个基于压缩感知和 DS 证据理论的双层融合模型, 并最终将融合结果发送给基站。

1 基于 DS 知和压缩感混合融合模型

本文设计的基于 DS 和压缩感知的混合融合模型基于层次化的分簇路由协议, 可以采用 LEACH、LEACH-C 或者其改进版本^[9-10], 即将整个 WSN 的监控区域分为若干簇, 每个簇均包含一个簇头和若干个簇成员节点, 簇成员节点在采集了数据后, 将数据发送给簇头节点, 再由簇头节点发送给基站, 如图 1 所示^[11-12]。

基于压缩感知和 DS 理论的混合数据融合模型在簇头处进行两级数据融合:

第一级数据融合的目标是对簇成员节点发送的数据进行融合, 由于簇内成员节点位置相隔较近, 因此, 采集的数据在时间和空间上均具有一定的冗余性, 同时如果某些传感器节点由于发生故障或其它原因采集的数据和其它节点差别很大时, 采用 DS 证据理论对其融合不仅可以实现对数据量的大量压缩, 同时还能保证采集数据具有一定的精确性。

第二级数据融合是为了使得簇头传输到基站的数据量足够

收稿日期: 2014-04-25; 修回日期: 2014-05-26。

基金项目: 永川区自然科学基金(Yeste, 2013nc8007)。

作者简介: 廖长荣(1979-), 男, 重庆潼南人, 硕士, 讲师, 主要从事电子电路设计及单片机应用等方向的研究。

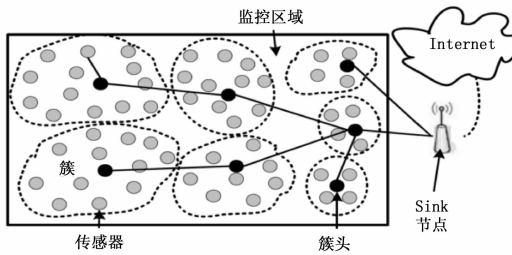


图 1 模型基于的网络拓扑结构

小，由于经过第一级数据融合后，簇头处的数据量已经大大减少，但是它仍然不是稀疏的，当簇头将其传送到基站时会耗费大量能量，为此，采用 CS 理论对数据进一步压缩，将数据表示为稀疏的向量，使得簇头发送给基站的数据量足够少，当基站接收到簇头发送的数据后，再采用某种重构算法如正交匹配分类算法对信号进行重构。

文中设计的基于 DS 和压缩感知的混合数据融合模型如图 2 所示。

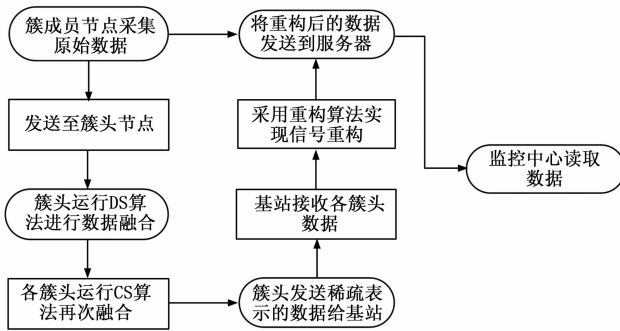


图 2 基于 DS 和压缩感知的混合融合模型

2 基于改进 DS 的第一级数据融合

2.1 DS 证据理论

定义 1: 概率分配函数，假设 Θ 为辨识框架，对于任意元素 $a, b \in \Theta$ ，均满足 $a \cap b = \Phi$ ，如果存在某函数 $m: 2^\Theta \rightarrow [0, 1]$ 满足：

$$\begin{cases} M(\Phi) = 0 \\ \sum_{A \in \Theta} M(A) = 1 \end{cases} \quad (1)$$

则称 m 为 Ω 上的基本概率赋值函数。 $\Omega = 2^\Theta$ 为 Θ 的幂集， Φ 为空集，代表不可能事件。

定义 2: Dempster 证据合成原则，假设 m_1, m_2, \dots, m_k 为 Θ 上的 k 个概率分配函数，则 DS 证据合成规则可以表示为：

$$\begin{cases} M(\Phi) = 0 \\ m(A) = \frac{1}{1-K} \times \sum_{\substack{\cap A_i = A \\ 1 \leq i \leq n}} \prod_{1 \leq i \leq n} m_i(A_i) \end{cases} \quad (2)$$

其中， K 表示证据之间的冲突部分，可以表示为：

$$K = \sum_{\substack{\cap A_i = \Phi \\ 1 \leq i \leq n}} \prod_{1 \leq i \leq n} m_i(A_i) \quad (3)$$

从式 (3) 中可以看出，当 $K = 1$ 时无法使用公式 (3)，当 K 趋向于 1 时，各证据之间的冲突很大，直接采用 DS 证据融合时往往会获得不正确的融合结果。

为解决上述问题，文中设计了一种基于权值的证据合成规则，即根据证据自身条件以及其对环境的影响，对不同的证据源提供的证据赋予不同的权值，并根据信任度分配函数的差值来构造相似度函数。

2.2 基于改进 DS 的数据融合设计

将 WSN 中的各簇成员节点采集的数据作为证据，对各证据采用改进的 DS 方法进行融合。

定义 3: 证据相似度，假设辨识框架中的 m 个完备而互不相容的命题为： $2^\Theta = \{A_1, A_2, \dots, A_m\}$ ， n 个证据 s_i 和 s_j ($i, j \in n, i \neq j$) 的信任度分配函数分别为 m_1, m_2, \dots, m_n ，则任意两证据 s_i 和 s_j 在命题为 A_k 时的相似度可以表示为：

$$sim_{ij}(A_k) = 1 - \frac{|m_i(A_k) - m_j(A_k)|}{\max(m_i(A_k), m_j(A_k))} \quad (4)$$

从式 (4) 可以看出，当证据对某命题的信度分配越一致，则他们的相似度越大，即可靠性更高。

定义 4: 信任度，证据 s_i 的信任度可以通过下式进行计算：

$$Cre_i(A_k) = \frac{sup_i(A_k)}{\sum_{j=1}^n sup_j(A_k)} \quad (5)$$

其中， $sup_i(A_k)$ 为证据 s_i 对命题 A_k 的综合支持度，可以表示为：

$$sup_i(A_k) = \prod_{j=1 \wedge i \neq j}^n sim_{ij}(A_k) \quad (6)$$

因此，文中簇头节点对多个传感器节点采集的数据进行融合的过程可以描述为：

- 1) 根据距离均匀地选取簇中的若干节点采集的数据作为命题，将所有传感器节点采集的数据作为证据；
- 2) 根据式 (5) 计算每个证据对某个命题 A_k 的信任度，并将其作为权值；
- 3) 采用权值对命题对应的所有焦元作加权平均：

$$m_{avg}(A_k) = \sum_{i=1}^n Cre_i(A_k) m_i(A_k) \quad (7)$$

- 4) 对加权平均结果进行归一化：

$$m_{avg}(A_k)' = \frac{m_{avg}(A_k)}{\sum_k m_{avg}(A_k)} \quad (8)$$

- 5) 将归一化的加权平均信度代替传统的信度分配函数，代入式 (2)，得到融合结果。

- 6) 重复 2) ~ 5)，直到所有命题均得到融合结果，此时将融合结果中 $m(A)$ 大于某阈值的结果作为最终的融合结果，即作为簇头发送给基站的数据。

3 基于 CS 压缩理论的数据稀疏表示

3.1 簇头发送基于 CS 的稀疏表示信号

香农-奈奎斯特采样定理指出：只有当信号采样频率大于或等于信号带宽的 2 倍时，才能保证任何一个连续带宽的有限信号的无损重构。

压缩感知理论能突破香农-奈奎斯特采样定理的限制，使其采样频率不仅受限于信号频率，而与信号结构相关。因此，能有效避免无用数据采集并节省采样时间和资源。

因此，当簇头接收的数据在经过 DS 融合后，可以再次通过 CS 压缩理论进行进一步压缩，将信号表示为稀疏形式，以

降低需要传输的数据量。

假设存在一个长度为 N 的离散信号 $X \in \mathbf{R}^N$ ，基向量为 $\varphi_i (i = 1, 2, \dots, N)$ ，则离散信号信号 X 可以表示为原子的线性组合：

$$X = \sum_{i=1}^N \alpha_i \varphi_i = \Psi \alpha \quad (9)$$

在式 (9) 中， α_i 是原始信号 X 在域 Ψ 上的等价表示，当 α_i 中大部分的元素为 0，仅有少量元素不为 0 时，则 α_i 是稀疏的，当 α_i 中仅有 K 个元素不为 0 且 $K \ll N, K$ 称为稀疏向量 α_i 的稀疏度，此时，原始信号 X 就称为可以稀疏表示的或者可以被压缩的。

因此，经过压缩的数据可以表示为：

$$Y = \Phi X = \Phi \Psi \alpha \quad (10)$$

其中， Φ 是一组与稀疏基无关的测量矩阵，其维数为 $M \times N (M \leq N)$ ， $\Phi = [\varphi_1, \varphi_2, \dots, \varphi_M]$ ，此时，将原始信号 X 投影到上得到一组测量值 $Y = [y_1, y_2, \dots, y_M]$ ，即：

$$y_i = \langle x, \varphi_i \rangle \quad (11)$$

N 维原始数据经过式 (9) 得到 K 个不为 0 的大系数，确定其位置，并删除剩余的 $N - K$ 个小系数，使得原始数据降到了 M 维，实现了数据的压缩，此时将这 K 个不为 0 的大系数和其位置进行编码，实现了数据的降维和压缩，并将其发送给基站。

3.2 基站重构稀疏信号

当基站接收到簇头发送的压缩数据后，需要对信号进行重构。由于式 (10) 中， $M \leq N$ ，因此方程是欠定的，即具有多个满足条件的解，无法重构数据。

但只要原始数据稀疏，且测量矩阵 Φ 满足条件：有限等距性质 (restricted isometric property, RIP 准则)，则 K 个不为 0 的大系数可以从 M 个测量值中准确重构，即：

$$\begin{cases} (1 - \delta_k) \|z\|_2 \leq \|Ab\|_2 \leq (1 + \delta_k) \|z\|_2 \\ \delta_k \in (0, 1) \end{cases} \quad (12)$$

通过求解最小 L_0 范数来根据测量值精确地对原始信号 X 进行重构：

$$\begin{cases} \alpha = \operatorname{argmin} \|\alpha\|_0 \\ s. t. Y = \Phi X = \Phi \Psi \alpha \end{cases} \quad (13)$$

在式 (13) 中， $\|\alpha\|_0$ 表示向量 α 中非零元素的个数，根据最小 L_0 范数在一定程度上与最小 L_1 范数具有等价性，式 (13) 可以更新为：

$$\begin{cases} \alpha = \operatorname{argmin} \|\alpha\|_1 \\ s. t. Y = \Phi X = \Phi \Psi \alpha \end{cases} \quad (14)$$

此时，可以根据一些全局优化算法或者正交匹配类的算法实现信号的重构。

4 仿真实验

在 Matlab 仿真环境下进行实验，实验参数设置如表 1 所示。

首先，将网络区域按照某种层次分簇路由由协议进行分簇，文中采用 LEACH-C 算法，在簇头处首先 DS 证据合成方法，将收集的簇成员节点发送的数据作为证据，将其中均匀采样获得的一些目标点作为命题，根据 DS 证据合成规则，得到每个命题的最终支持度，将具有最大支持度的命题作为数据融合的

结果。

表 1 数据融合仿真参数表

参数名称	参数值
节点数目/个	800
仿真时间/ms	300
基站位置(m, m)	(600, 550)
仿真区域/(m * m)	1 000 × 1 000
发送接收数据能量消耗/(nJ/bit)	$E_{elec} = 70$
数据融合能量耗费/(nJ/bit)	$E_{da} = 25$
节点初始能量/J	$E_0 = 2$
数据包发送率/(packet/s)	2
DS 融合阈值	0.8

然后，在各簇头处采用 CS 压缩感知理论进行稀疏化表示，并对稀疏表示的向量的 K 值和位置进行编码，将其发送给基站。

在同样采用 LEACH-C 协议和表 1 设定的参数的情况下，对不同的方法从重构误差和网络生命周期两个角度进行比较。

3 种方法得到的重构误差如图 3 所示。

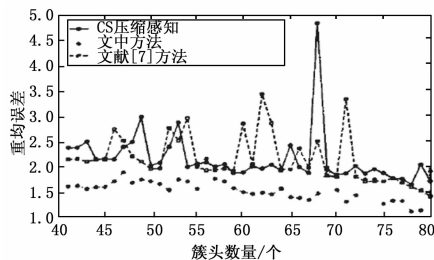


图 3 重构误差比较

从图 3 中可以看出，文中方法对应的重构误差最小，平均约为 1.342，而 CS 压缩感知方法的重构误差振荡较为剧烈，重构误差平均约为 3.4，文献 [7] 相对较为平缓，平均约为 2.8，显然，文中方法的效果最好，这是因为文中方法首先经过 DS 证据融合，去除具有较大冲突的数据，因此，重构误差较小。

3 种方法对应的网络生命周期如图 4 所示。

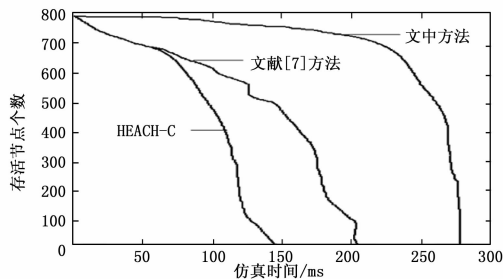


图 4 网络生命周期比较

从图 4 中可以看出，文中方法的最后一个节点死亡的时间为 280 轮，即网络生命周期为 280 轮，LEACH-C 方法为 142 轮，文献 [7] 方法为 207 轮，显然文中方法较其分别提高了 96.48% 和 35.27%，这是因为文中方法在 LEACH-C 协议的

