

基于 LEACH 协议的多因子可靠数据融合优化策略

王振飞 余 利 郑志蕴

(郑州大学信息工程学院 郑州 450001)

摘要 针对无线传感网中节点部署密度大、数据冗余度高、能量有限和易遭攻击等问题,提出基于 LEACH 算法的多因子可靠数据融合优化策略。该策略对 LEACH 算法做出 3 方面优化改进:在数据融合的相似度计算中增加两个可靠性优化因子 MN-LEACH 和 LF-LEACH;在数据传输中采用多路径传输优化因子 MT-LEACH。在敌对环境中,为避免大量恶意节点干扰真实数据,提高数据融合精确度,首先对突发的非线性噪声数据采用拉普拉斯函数而非高斯函数进行过滤,然后采用 MN-LEACH 优化因子计算相似度,通过加权平均进行数据融合,最后采用 LF-LEACH 优化因子对链路传输进行检测,并在传输时采用 MT-LEACH 优化因子,从而使链路负载更加均衡。实验结果表明,该策略与传统 LEACH 算法相比,在数据融合精确度、信噪比及能耗等方面有明显的优势。

关键词 无线传感网, LEACH 算法, 数据融合, 链路故障, 精确度

中图分类号 TP393 **文献标识码** A

Reliable Data Aggregation Strategy of Multi-optimized Factors Based on LEACH Algorithm

WANG Zhen-fei YU Li ZHENG Zhi-yun

(School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract Owing to the deployment of nodes in high density and limited energy, the wireless sensor nodes have the problems of high data redundancy and are vulnerable to be attacked, and so on. We proposed a reliable data aggregation strategy of multi-optimized factors based on LEACH algorithm in wireless sensor networks. The strategy makes some improvements to optimize the LEACH algorithm in three aspects: two reliable optimizing factors MN-LEACH and LF-LEACH are adopted during the calculation of data fusion in the similarity; the optimization of multi-path transmission factor MT-LEACH is used for data transmission. Firstly, in a hostile environment, Laplace is used instead of Gauss function to filter the emergent and nonlinear noise data in order to avoid a large number of malicious nodes interfering with real data and improve the accuracy of data fusion, and then we used MN-LEACH optimizing factor to calculate similarity and average the weight for data aggregation. Finally, we used optimization of LF-LEACH factor to detect the link transmission, and adopted the MT-LEACH optimizing factor in the process of transmission as well, so that the load of links will be balanced. The experimental results show that the strategy outperforms the traditional LEACH strategy, in terms of accuracy of data aggregation, signal-to-noise ratio and energy consumption, and so on.

Keywords Wireless sensor networks, LEACH algorithm, Data aggregation, Link fault, Accuracy

1 引言

无线传感网络(wireless sensor networks, WSN)是由数量众多、体积微小、仅电池供电的传感器节点通过无线通信的形式形成的自组织网络。特点是规模大、密度高、存在大量冗余。传感器节点大多部署在无人看管或几乎不能更换电池的环境下长时间工作,如何有效降低能耗、延长网络生命周期和提高数据融合精确度便成为重要问题^[1]。

由于相邻传感器节点感知的数据通常高度相关,甚至相同,网络中感知的冗余数据大量传给基站(BS)处理,降低了网络性能,因此数据融合技术的使用使无线传感网络的生命

周期、数据精确度和可靠性等网络性能得到重大改善。数据融合是对多份数据压缩组合出有效、符合用户需求的数据的过程。利用算法对多传感器观测的信息进行检测、相关估计和综合,产生比单一传感器更精确可靠的信息,从而提高数据精确度,减少能耗和延长网络生命周期^[2]。

LEACH 路由协议是一种低功耗的自适应分簇算法^[3],将 WSN 网络中的节点分为若干簇,在簇建立阶段,随机产生簇首并形成簇域;在簇通信阶段,簇内节点将感知的数据直接传送给簇首,簇首对数据融合后传送给基站。本文在 LEACH 算法基础上,采用多路径多跳传输的网络拓扑结构,对感知或收集的数据采用多因子优化,进行可靠数据融合。

本文受国家自然科学基金项目(61379079)资助。

王振飞(1973—),男,博士,副教授,主要研究方向为分布式计算、信息安全, E-mail: iezfwang@zzu.edu.cn; 余 利(1989—),女,硕士生,主要研究方向为无线传感网信号处理和信息安全; 郑志蕴(1962—),女,博士,教授,主要研究方向为分布式计算、智能信息处理。

2 相关工作

在无线传感网络中,为提高数据融合的精确度,文献[4]采用卡尔曼滤波(kalman filter, KF)融合方法,建立信号和噪声的状态空间模型,利用前一时刻系统状态估值和现时刻系统观测值得到现实状态的最优估计。该方法对线性高斯噪声数据进行高效处理,一般应用于线性高斯的情况,但对非线性的环境预测,预测误差偏大。文献[5]对 KF 进行改进,采用无迹卡尔曼滤波算法(unscented kalman filter, UKF)。首先建立非线性的状态方程,选择一组样本点确定状态向量的概率密度函数,然后进行无损变换得到最优估计,最后进行数据融合处理。然而样本点的随机选择带来大量恶意节点入侵产生的安全隐患。文献[6]采用的粒子滤波算法,使用蒙特卡罗方法来实现最优的递归贝叶斯估计,并选用合适的重要性函数代替状态后验概率分布,是非高斯非线性状态下的“最优”滤波器。但负责存储、计算的传感器节点的通信开销能耗也会增加。文献[7]采用灰度模型和卡尔曼滤波相结合的方法,首先设置双队列机制以降低误差率,对同步传感节点和簇首进行数据融合估计,然后采用卡尔曼滤波对噪声数据移除,最后进行加权融合。该策略一定程度上提高了数据融合精确度,但没有考虑链路故障对融合精确度的影响。

文献[8]提出 OPAG 算法以实现可靠数据融合。利用无计算错误的内网进行数据融合,能容忍适度的数据丢失。但融合节点通常离源数据节点太远,导致数据精确度降低。为增强数据融合的可靠性,文献[9]提出 EEHA 算法,采用 TAG 算法为节点建立一棵数据融合树,利用分片重组技术增强对隐私数据的保护。分片重组及单路径传输的策略增加链路负载,从而增加能耗。文献[10]提出 P-SMART-CLPNT 算法,在 SMART^[11]隐私保护方案基础上加入多类优化因子以减少数据传输中的碰撞,从而提高数据融合精确度,但没有考虑使链路负载均衡的多路径传输策略。

针对上述文献的不足,本文提出基于 LEACH 算法多优化因子的可靠数据融合策略(reliable data aggregation strategy of multi-optimized factors based on LEACH algorithm, MFO-LEACH)。该策略对 LEACH 协议做出 3 方面的改进。在敌对环境下,为提高数据融合精确度,避免大量恶意节点干扰真实数据,对突发的非线性噪声数据采用拉普拉斯函数而非高斯函数进行过滤,利用 MN-LEACH 优化因子计算相似度,以加权平均进行数据融合。为增强数据传输可靠性,采用 LF-LEACH 优化因子对链路传输进行检测,并在传输中采用 MT-LEACH 优化因子,使链路负载均衡。

3 研究背景与网络模型

3.1 研究背景

如图 1 和图 2 所示,随着链路故障率 ℓ 的增加,数据平均融合精确度 \mathcal{R} 随之降低,网络生命周期呈下降趋势。当 $\ell = 0.210$ 时, $\mathcal{R} = 4.000$, 网络生命周期降为 0.345(s)。特别地,当 $\ell = 0.312$ 时,其 $\mathcal{R} = 2.521$, 对应的网络生命周期降为 0.292(s)。说明 LEACH 算法中链路故障的处理并不理想,其

是导致数据融合精确度和网络生命周期降低的主要因素。

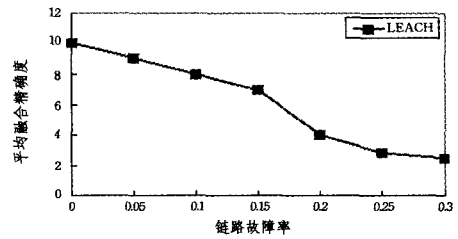


图1 链路故障率与平均融合精确度之间的关系

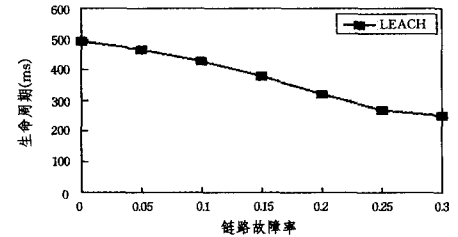


图2 链路故障率与生命周期之间的关系

基于仿真结果分析,本文从 3 个角度进行改进,在 LEACH 算法中加入 3 个优化因子以提高数据融合的精确度与可靠性。研究框架如图 3 所示。

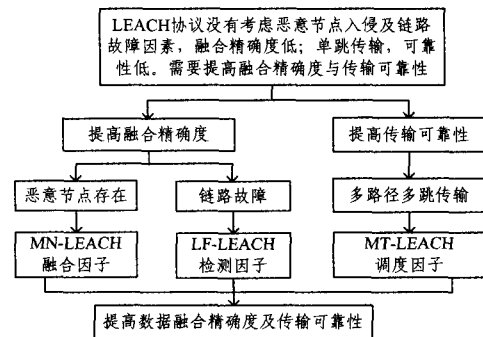


图3 研究框架

3.2 网络模型

假定网络采用数据融合层和路由层两层结构。数据融合层采用基于优化因子的加权平均计算方法;路由层采用基于 LEACH 协议的多路径多跳传输策略。网络模型如图 4 所示。

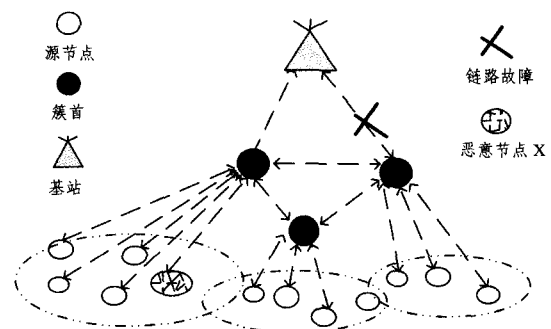


图4 基于 LEACH 协议的多路径多跳传输模型

4 多因子优化策略

为提高数据融合的可靠性,从恶意节点入侵、链路故障和多路径传输 3 个角度提出相应的优化因子,具体分析 3 种优

化因子并给出相应的算法。

4.1 节点入侵优化因子(MN-LEACH)

由于传感器的分布位置、任务需求不同,因此其感知的数据也不同。

定义 1 不同应用需求的传感器节点数为 N 个,节点感知环境所得的不同数据的特征属性为 D 个,即存在 $N \times D$ 节点-数据矩阵 S 。 $s_{n,d} = x(x \in \{1, 2, \dots, S_{max}\})$,表示节点 n 对数据属性 d 的评分, $s_{n,d}$ 的默认值为 0。节点-数据矩阵表示如下:

$$S = [n_1, \dots, n_N]^T, n_i = [S_{i,1}, \dots, S_{i,m}, S_{i,n}, \dots, S_{i,D}]^T$$

$$(i = 1, \dots, p, q, \dots, N, (p \neq q))$$

$$S = [d_1, \dots, d_D], d_j = [S_{1,j}, \dots, S_{p,j}, S_{q,j}, \dots, S_{N,j}]$$

$$(j = 1, \dots, m, n, \dots, D, (m \neq n))$$

其中, n_i 表示节点 i 对所有数据属性的评分, d_j 表示所有节点对数据 j 属性的评分。

为规避特殊情况,如恶意节点伪造虚假数据使得节点对数据的评分近似于平均分值,但实际评分态度和平均分值态度相反,针对这种突发的干扰噪声,本文采用拉普拉斯随机变量对节点进行估计,而不是采用只对线性噪声过滤的高斯随机变量。利用信息向量判断节点对数据评分的真实态度,并依据信息向量矩阵计算节点间相似度。针对大量恶意节点过高的评分,采用降低权重的方法进行约束以提高数据融合精确度。

假定 $S_{1,j}, S_{2,j}, \dots, S_{N,j}$ 是独立且服从均匀分布的样本。由概率密度函数公式,对拉普拉斯随机变量 μ_j 和 b_j 进行最大似然估计。概率密度函数表达式由式(1)计算。

$$f(S|\mu, b) = \frac{1}{2b} \exp\left(-\frac{|S-\mu|}{b}\right) \quad (1)$$

其中, μ 是位置参数, b 是大于 0 的范围参数。 μ_j 和 b_j 最大似然估计表达式分别由式(2)和式(3)计算。

$$\hat{\mu}_j = \frac{1}{D} \sum_{i=1}^D S_{p,j} \quad (2)$$

$$\hat{b}_j = \frac{1}{D} \sum_{i=1}^D |S_{p,j} - \hat{\mu}_j| \quad (3)$$

定义 2 节点 p 的信息向量 R_p

$$R_p = [S_{p,1}, \dots, S_{p,D}]^T$$

$$S_{p,j} = \text{sgn}(S_{p,j} - \hat{\mu}_j) \times I(S_{p,j}) \quad (p = 1, \dots, N, j = 1, \dots, D)$$

其中, $\text{sgn}(S_{p,j} - \hat{\mu}_j)$ 表示节点 p 对数据属性 j 的态度。根据 j 的平均态度,对 $\text{sgn}(S_{p,j} - \hat{\mu}_j)$ 做出正或负的态度表示。 $I(S_{p,j})$ 是对 $S_{p,j}$ 评分的信息量,由式(4)计算。

$$I(S_{p,j}) = -\ln(f(S = S_{p,j} | \hat{\mu}_j, \hat{b}_j))$$

$$= \ln(2\hat{b}_j) + \frac{|S_{p,j} - \hat{\mu}_j|}{\hat{b}_j} \quad (4)$$

为防止敌方设置多个恶意节点,伪造虚假数据,形成评分过高的干扰,导致真实数据丢失的情况,采用降低权重因子过高的约束方法。该因子添加相关重量级权重因子以降低相似的权重。该权重由小部分合作评分的数据度量,由式(5)计算。

$$\text{sim}_I(n_p, n_q) = \frac{\text{Min}(|I_{n_p} \cap I_{n_q}|, \eta)}{\eta} \times \frac{\sum_{\{i|S_{p,j}, S_{q,j} \neq 0\}} R_{p,j} \times R_{q,j}}{\sqrt{\sum_{\{i|S_{p,j}, S_{q,j} \neq 0\}} R_{p,j}^2} \times \sqrt{\sum_{\{i|S_{p,j}, S_{q,j} \neq 0\}} R_{q,j}^2}} \quad (5)$$

其中, $|I_{n_p} \cap I_{n_q}|$ 是节点 n_p 和 n_q 合作评分的数据值。如果该数据值小于 η ,则这些节点的相似度将会降低。节点相似性权重由式(6)计算。

$$w_i = \begin{cases} \text{sim}_I(n_p, n_q), & \text{sim}_I(n_p, n_q) > 0 \\ 0, & \text{sim}_I(n_p, n_q) \leq 0 \end{cases} \quad (6)$$

对权重加权平均,由式(7)计算出融合数据。

$$w' = \frac{1}{N} \sum_{i=1}^N w_i \quad (7)$$

若基站没有检测到恶意节点,则簇首用修正的余弦计算节点间相似度,由式(8)计算^[12]。

$$\text{sim}(n_p, n_q) = \frac{\sum_{j \in S_{pq}} (S_{p,j} - \bar{S}_p) \times (S_{q,j} - \bar{S}_q)}{\sqrt{\sum_{j \in S_p} (S_{p,j} - \bar{S}_p)^2} \times \sqrt{\sum_{j \in S_q} (S_{q,j} - \bar{S}_q)^2}} \quad (8)$$

其中, $S_{p,j}$ 和 $S_{n,j}$ 分别表示节点 p 和 q 对数据属性 j 的评分, S_{pq} 表示 S_p 和 S_q 的评分集合, \bar{S}_p 和 \bar{S}_q 分别表示节点 p 和 q 对数据的平均评分。

算法 1 MN-LEACH 算法

输入: w_i 和 net 或者 NeigID

// w_i 表示权重,net 表示网络,NeigID 表示邻居节点的 ID

输出: AggData // AggData 融合数据

1. 簇首选择阶段

由 LEACH 算法选择 CH // CH 表示簇首节点

采用 TDMA 为每个 CH 内成员分配时间片 // 避免簇内节点发送数据时产生冲突

2. 数据融合阶段

FOR($n_i = 1; n_i < N; n_i++$)

{ FOR($d_j = 1; d_j < D; d_j++$)

{ 由式(2)计算出未知数据的评分

计算差值 $F = \text{EXP}_i - T_i$ 。

// EXP_i 表示节点 n_i 对数据 d_j 评分预期, T_i 表示实际数据评分

IF($F > S_{thre}$) // S_{thre} 表示评分阈值

NodeNor = 0 // 存在恶意节点

ELSE

NodeNor = 1

}

}

IF(NodeNor = 0)

由式(1)~式(5)计算相似度

由式(6)、(7)进行数据融合

ELSE

由式(8)计算相似度

平均加权融合

3. 数据传输阶段

由 MT-LEACH 算法将融合后的数据传输给基站。

4.2 链路故障优化因子(LF-LEACH)

针对网络中链路节点能耗过低或遭恶意节点入侵而导致

整个链路发生异常的情况,采用链路故障优化算法。核心理想是利用链路检测模型判断链路传输是否正常。若链路传输正常,则返回正常标志 1,否则返回 0,并由基站广播 PROBE 故障消息给所有簇首。簇首判断故障链路上的节点是否隶属于本簇,若是,则簇首转发故障消息给对应链路的邻居节点。然后调用 MN-LEACH 算法检测邻居节点是否正常,最后簇首将数据重新融合,使用 MT-LEACH 优化算法转发给基站。若不是本簇,则不作响应。

链路检测模型由式(9)计算^[13]。

$$\theta \leq \frac{P_{xy}/d_{xy}^\alpha}{\epsilon_0 + \sum_{L_{gh} \in \psi_{xy} - \{L_{xy}\}} P_{gh}/d_{gh}^\alpha} \quad (9)$$

其中, ψ_{xy} 表示并发传输的链路集合, L_{xy} 表示链路集合, P_{xy} 和 P_{gh} 分别表示链路 L_{xy} 和 L_{gh} 的发射器传输功率, d_{xy} (d_{gh})表示链路 L_{xy} (L_{gh})发射器与接收器间的距离, α 表示路径衰减率(取值在 2 和 6 之间)。 ϵ_0 表示噪声, θ 是信号干扰噪声率(SINR)成功传输的阈值($\theta \geq 1$)。当 $SINR \geq \theta$ 时,链路 L_{xy} 传输正常。

算法 2 LF-LEACH 算法

输入: AggData

输出: PostAggData //最终融合的数据

1. 簇首选择阶段:同 MN-LEACH 算法的步骤 1

2. 数据融合阶段

IF(SINR $\geq\theta$)//由式(9)计算

LinkNor=1 // 链路传输正常

ELSE

LinkNor=0

IF(LinkNor=0)

IF(CH.CID==SourID.CID)//是否属于本簇

转发故障信息 PROBE 消息给 NeigID

MN-LEACH(NeigID) //采用 MN-LEACH 算法步骤 2 进行处理

ELSE

不作处理

ELSE

输出 PostAggData

3. 数据传输阶段:同 MN-LEACH 算法步骤 3。

4.3 多路径传输优化因子(MT-LEACH)

为减少数据传输中的碰撞,提高数据传输可靠性,簇首采用基站的优化调度算法。按照路径列表传输数据,基站利用能耗评估模型计算各路径的剩余总能量,判断链路传输是否正常,并链路排序,由链路有效和能量最多的路径传输重要数据。其他数据由路径向量表在 TDMA 时间内并行传输。

路径向量表的数据包格式如表 1 所列。

表 1 路径向量表的包格式

SourID	DestID	LID	E_{lr}
PreHop	HopCount	NextHop	TVaild

SourID 表示源数据节点 ID, DestID 表示目的节点 ID, LID 表示路径编号, E_{lr} 表示路径剩余能量,由式(10)一式(13)计算。TVaild 取值布尔型,表示是否在 TDMA 机制分配的时间片内,当 TVaild=1,表示在转发数据的时间范围内,否则此时间片不转发该数据。PreHop 表示前一跳, HopCount 表示跳数, NextHop 表示下一跳。

节点发送和接收 k (bit)数据所消耗的能量分别由式(10)和式(11)计算^[14]。

$$E_t(k, d) = E_{t-e}(k) + E_{t-a}(k, d) \\ = E_e \times k + \xi_a \times k \times d^\lambda \quad (10)$$

$$E_r(k) = E_{r-e}(k) = k \times E_e \quad (11)$$

节点 i 在某一时刻的剩余能量 E_{ir} 由式(12)计算。

$$E_{ir}^i = E_p - (E_t(k, d) + E_r(k)) \quad (12)$$

其中, $E_t(k, d)$ 表示 k 个数据包传输 d 距离的能耗, E_e 表示发射电路和接收电路模块处理 1bit 数据的能耗, d 表示数据传输距离, ξ_a 表示发射功率放大器处理 1bit 数据的能耗, λ 是无线信道决定的系数常量(一般 λ 取 2 或 4), $E_r(k)$ 表示接收 k 个数据包所消耗的能量, E_p 表示节点当前电池能量, E_{thre} 表示传输数据最低能量。若 $E_{ir} \geq E_{thre}$, 则计算路径总剩余能量, 否则发送链路故障报告。

路径剩余能量由式(13)计算。

$$E_{lr} = \sum_{i=1}^N E_{ni} \quad \forall i=1, \dots, N \in l \quad (13)$$

路径选择由式(14)计算。

$$E_k = \max_{l=1, \dots, L} \{E_{lr}\} \\ = \max_{l=1, \dots, L} \left\{ \sum_{i=1}^N E_{ni} \right\} \quad \forall i=1, \dots, N \in l \quad (14)$$

其中, E_{lr} 表示路径所有节点剩余总能量, E_k 表示各路径剩余能量最大值。

算法 3 基站调度 Link-Schedule

输入: 路径向量表

输出: LID 列表

由式(9)计算

IF(LinkNor=0)

发送 PROBE 故障消息

ELSE

由式(10)一式(13)计算出各条路径的总的剩余能量

快速排序

由式(14)选择对应的链路

输出 LID 列表。

算法 4 MT-LEACH 算法

输入: LID 列表

输出: PostAggData

1. 簇首选择阶段:同 MN-LEACH 算法步骤 1

2. 数据融合阶段:同 MN-LEACH 算法步骤 1 和 LF-LEACH 算法步骤 2

3. 数据传输阶段

IF(LID.LVaild=1)//在 TDMA 分配的时间片内

转发 PostAggData

ELSE

等待下一个 TDMA 时间片

PROBE 消息的包格式在可变量段中增加了前 4 个字段。

固定字段:

SourID	DestID	SeqNum
--------	--------	--------

可变量段:

LID	NeigID	TVaild	Payload
NodeNor	LinkNor	CID	

NeigID 表示一维数组,存储邻居节点 ID, Payload 表示数据载荷, NodeNor 和 LinkNor 均为布尔型, 1 表示正常, 0 表示不正常, CID 为整型, 表示簇首 ID。

5 性能分析

采用 OPNET 仿真实验平台, 在链路故障时, 分别对传统 LEACH 算法和 MFO-LEACH 策略在网络生命周期、数据融合精确度、信噪比及能耗 4 个方面进行比较与分析。仿真环境参数配置如表 2 所列。

参数名称	配置值
传感节点数目	300
场景大小	100m×100m
拓扑配置模式	随机
感知数据的大小	默认: 2048bits
报告间隔	0.2ms
路径数	6 条
仿真时间	300s
路由协议	LEACH
最大传输范围	60m
数据传输速率	1 Mbps
基站电池能量	∞ (无穷大)
SN 节点的电池能量	25Joules

5.1 网络生命周期

网络生命周期由式(15)计算^[15]。

$$T[L] = \frac{(E_0 - E[Ew])}{(Pc + \lambda E[Er])} \quad (15)$$

其中, E_0 表示网络初始能量, $T[L]$ 表示网络平均生命周期, Pc 表示网络持续的恒定能耗, $E[Ew]$ 表示网络崩溃时的总剩余能量, λ 表示传感器平均报告速率, $E[Er]$ 是消耗的总能量。

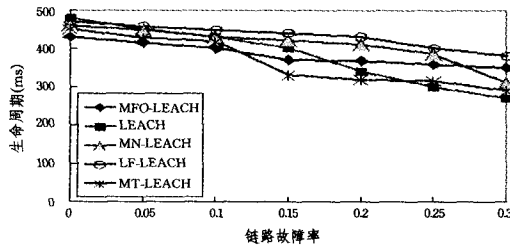


图 5 生命周期与链路故障率之间的关系

网络生命周期与链路故障之间的关系如图 5 所示。当 $\ell \in (0, 0.05)$ 时, 传统 LEACH 算法的生命周期要高于其他优化策略, 这是因为 MN-LEACH 算法需要检测恶意节点, 若存在恶意节点, 则采用不同的计算方法进行数据融合, 这将导致计算量和能耗增加, 但能耗增加是微小的(由图 8 可知), 因此生命周期降低微小。LF-LEACH 算法需要检测链路, 若存在链路故障, 则消息的故障节点的邻居节点需要重新感知融合, 这必然会增加能耗, 因此生命周期比传统 LEACH 略微降低。MT-LEACH 算法增加数据多路径传输机制, 需要计算路径剩余能量并判断路径是否有效, 因此其生命周期比传统 LEACH 低 [22, 30]ms, 同理, MFO-LEACH 策略增加 MN-LEACH、LF-LEACH 和 MT-LEACH 3 个优化因子, 因此生命周期在开始时低于传统 LEACH 的生命周期。

当 $\ell \in (0.15, 0.30)$ 时, 随着链路故障率的增加, 传统 LEACH 的生命周期近似线性降低, 这是由于 LEACH 没有恶意节点入侵、链路故障及多路径传输的处理机制, 链路不可

靠, 导致网络生命周期大幅下降。此时, 其生命周期要远远低于其他的优化策略。另外, 在链路故障增加的情况下, MFO-LEACH 策略的网络生命周期总体比较稳定, 说明该策略具有较高的可靠性。

5.2 数据融合精确度

利用基站均方差 \mathcal{R} 来度量融合精确度, 如式(16)所示。

$$\mathcal{R} = \frac{1}{T} \sqrt{\sum_{i=1}^N (T_i - C)^2 / N} \quad (16)$$

其中, C 表示源节点的真实值, T_i 表示 t 时刻基站融合后的值, N 表示选取时间点, T 表示 $t \in [1, N]$ 时间段内基站融合后的值。

平均数据融合精确度与链路故障之间的关系如图 6 所示。3 个单因子的优化策略和 MFO-LEACH 优化策略整体上均比传统 LEACH 算法的平均融合精确度高。这是因为 MN-LEACH、LF-LEACH 和 MT-LEACH 3 个单因子优化方法分别从恶意节点数据融合的计算、链路故障处理和多路径传输调度机制 3 个角度均增强了数据融合和传输的可靠性, 因此提高了数据平均融合精确度。MFO-LEACH 策略的平均融合精确度比传统 LEACH 高出 5—8 个点。特别地, 当 $\ell \in (0.15, 0.21)$ 时, 传统 LEACH 策略平均融合精确度大幅降低, 说明当链路故障率大于一定阈值时(此阈值随网络拓扑的变化、网络规模的大小等因素动态变化), 传统 LEACH 策略数据融合的质量较低。由此说明 MFO-LEACH 策略的可靠性比传统 LEACH 高。另外, MFO-LEACH 策略的数据平均融合精确度均比 3 个单因子优化方法精确度高, 并且随着链路故障率的增加, 其数据平均融合精确度降幅比 3 个单因子方法低, 说明 3 种因子分别对节点数据融合、链路和传输的优化对数据融合的精确度及可靠性有明显的优化效果。

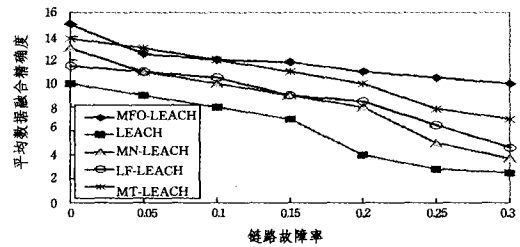


图 6 数据融合精确度与链路故障率之间的关系

5.3 信噪比

平均信噪比与链路故障之间的关系如图 7 所示, 当 $\ell \in (0, 0.30)$ 时, MFO-LEACH 策略的平均信噪比仍然保持在 [38.53, 46.79] 区间, 说明 MFO-LEACH 策略的信噪比比传统 LEACH 算法的高, 抗噪声干扰能力优势明显, 几乎保持平稳信噪比, 同时其平均信噪比比分别采用 MN-LEACH、LF-LEACH 和 MT-LEACH 单一的优化因子的高。这是因为单一的优化因子, 只能单纯提高某一方面的抗噪声干扰性, 不能兼具防御故障的处理能力。特别地, MFO-LEACH 策略的平均信噪比比传统 LEACH 算法的高出 [19.21, 29, 83] (dB)。另外, 3 个单一优化因子方法的平均信噪比均比传统 LEACH 算法的高, 这说明随着链路故障率的增加, MFO-LEACH 策略可靠性比传统 LEACH 在可靠性和信噪比方面均有提高。

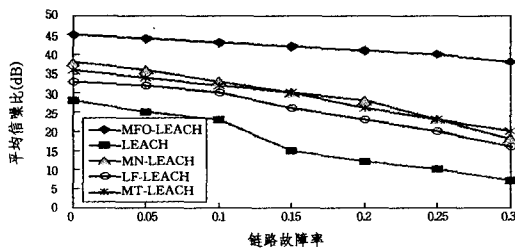


图7 平均信噪比与链路故障率之间的关系

5.4 能耗

关于能耗的计算,本文已在 MT-LEACH 部分中作了介绍,由此定义出链路故障率 ℓ 。

$$\ell = \frac{\sum_{l=1}^L \sum_{i \in l} E_{nr}^i}{\sum_{l=1}^L E_l} \quad (E_{nr} < E_{thre})$$

其中, l 表示某条路径, L 表示路径总条数,由于应用和网络拓扑的不同, E_{thre} 有不同的要求。本文 E_{thre} 的取值范围为 $E_{thre} \in (8, 90, 10, 18)$ 焦耳。

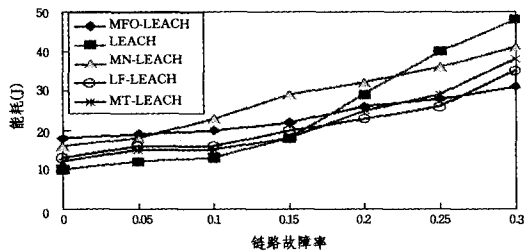


图8 网络能耗与链路故障率之间的关系

网络能耗与链路故障率的关系如图8所示。当 $\ell \in [0, 0.16]$ 时,传统 LEACH 算法的能耗均低于其他优化因子的能耗。这是因为传统 LEACH 算法没有增加处理可靠性的能耗。此时,MT-LEACH 优化因子的能耗低于其他优化因子算法的能耗。说明多路径策略一定程度上使链路负载均衡,降低了网络能耗。但当 $\ell \geq 0.16$ 时,传统 LEACH 算法能耗近似线性增长,但加入可靠性优化因子的 LEACH 算法能耗的增长幅度较低,因为随着链路故障率的增长,传统 LEACH 对恶意节点入侵、链路故障及数据传输没有对应的处理机制,所以能量消耗增多。特别地,当 $\ell \geq 0.18$ 时,MFO-LEACH 策略的能耗要低于传统 LEACH 算法的能耗,说明多因子综合优化的算法对链路故障增加的抵抗力较强,并且其能量消耗较低。当 $\ell \in [0, 0.08]$ 时,MFO-LEACH 算法的能耗比 MN-LEACH、LF-LEACH 和 MT-LEACH 算法的都高,这是因为采用三重可靠性保障的算法比单个因子的可靠性算法略微复杂。但当 $\ell \geq 0.08$ 时,MFO-LEACH 算法的能耗增长幅度开始低于其它单因子的优化算法,说明 MFO-LEACH 算法在增强可靠性的同时其链路负载的均衡也对网络能耗的降低起平衡作用。

结束语 本文提出无线传感网中基于 LEACH 算法的多因子可靠数据融合优化策略,在 LEACH 算法基础上,融入了 MN-LEACH、LF-LEACH 和 MT-LEACH 3 个优化因子,分别针对大量恶意节点干扰真实数据,作出数据融合方法优化、链路故障预测和数据多路径多跳传输策略的优化。仿真实验结果表明该策略可有效地提高数据融合的精确度,降低能耗,

并且多路径传输可有效避免链路故障或者节点故障造成的数据丢失,从而提高数据传输的可靠性。

参考文献

- [1] Renjith P N, Baburaj E. An analysis on data aggregation in wireless sensor networks[C]// Proceedings of International Conference on Radar, Communication and Computing(ICRCC). Tiruvannamalai, India, 2012; 62-71
- [2] Yuan Fei, Zhan Yi-Ju, Wang Yong-Hua. Data density correlation degree clustering method for data aggregation in WSN[J]. IEEE Sensors Journal, 2014, 14(4): 1089-1098
- [3] Fan Xiang-ning, Song Yu-lin. Improvement on LEACH protocol of wireless sensor network[C]// Proceedings of International Conference on Sensor Technologies and Applications(ICSTA). Valencia, Spain, 2007; 260-264
- [4] Jain A, Chang E Y, Wang Y-F. Adaptive stream resource management using Kalman filters[C]// Proceedings of ACM SIGMOD International Conference on Management of Data. New York, USA, 2004; 11-12
- [5] Kim J, Jung K, Kim J, et al. Positioning accuracy improvement of laser navigation using unscented Kalman filter[J]. Intelligent Autonomous Systems, Springer, 2013, 193(1): 807-816
- [6] Chan Fu-kai, Wen Chih-yu. Adaptive AOA/TOA localization using fuzzy particle filter for mobile WSNs[C]// Proceedings of the 73rd IEEE Vehicular-Technology Conference(VTC). Budapest, Hungary, 2011; 1-5
- [7] Wei G, Ling Y, Guo B, et al. Prediction-based data aggregation in wireless sensor networks; Combining grey model and Kalman filter[J]. Computer Communications, 2011, 56(3): 359-370
- [8] Chen Z, Shin K G. OPAG: Opportunistic data aggregation in wireless sensor networks[C]// Proceedings of Real-Time System Symposium. Barcelona, Spain, 2008; 345-354
- [9] Li H, Lin K, Li K. Energy-efficient and high-accuracy secure data aggregation in wireless sensor networks[J]. Computer Communication, 2011, 34(4): 591-597
- [10] 杨庚, 李森, 陈正宇, 等. 传感器网络中面向隐私保护的高精确度数据融合算法[J]. 计算机学报, 2013, 36(1): 189-200
- [11] He W B, Liu X, Nguyen H, et al. PDA: Privacy-preserving data aggregation in wireless sensor networks[C]// Proceedings of the 26th IEEE International Conference on Computer Communications. Anchorage AK, USA, 2007; 2045-2053
- [12] 潘谦红, 王炬, 史忠植. 基于属性论的文本相似度计算[J]. 计算机学报, 1999, 22(6): 651-655
- [13] Li Hong-xing, Wu Chuan, Hua Qiang-Sheng, et al. Latency-minimizing data aggregation in wireless sensor networks under physical interference model[J]. Elsevier Ad Hoc Networks, 2014, 12(1): 52-68
- [14] Sahoo R R, Singh M, Sardar A R, et al. TREE-CR: trust based secure and energy efficient clustering in WSN[C]// IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology (ICECCN). Tirunelveli, India, 2013; 532-538
- [15] Chen Yun-xia, Zhao Qing. On the lifetime of wireless sensor networks[J]. IEEE Communications Letters, 2005, 9(11): 976-978