无线传感器网络中一种高效虚假数据过滤方案

赵巾帼[1] 朱凌志[1] 罗庆云[1] 梁俊斌[2]
(湖南工学院计算机与信息科学学院 衡阳 421002)[1] (广西大学计算机与电子信息学院 南宁 530004)[2]

摘 要 无线传感器网络中虚假数据过滤机制工作效率较低的根本原因在于在提高密钥共享度的同时无法保证密钥的安全性。提出了一种高效率的虚假数据过滤机制。构造簇头生成树,在源簇和中转簇之间建立关联;基于负荷指数提出了一种密钥分发策略,靠近源簇的中转簇存储源簇的多个密钥,提高了密钥共享度,且密钥在中转节点中分布较均匀;来自同一个源簇的多个密钥由中转簇中不同节点存储,保障了密钥的安全性。理论分析及仿真实验表明,该方案在提高虚假数据过滤效率的同时能均衡节点通信开销,并具有较低的能量消耗和存储开销。

关键词 无线传感器网络,虚假数据过滤,簇头树,负荷指数,密钥分发
中图法分类号 TP393 **文献标识码** A

Efficiency Filtering Scheme for False Data in Wireless Sensor Networks

ZHAO Jin-guo[1] ZHU Ling-zhi[1] LUO Qing-yun[1] LIANG Jun-bin[2]
(School of Computer and Information Science, Hunan Institute of Techology, Hengyang 421002, China)[1]
(School of Computer and Electronics Information, Guangxi University, Nanning 530004, China)[2]

Abstract The fundamental reason for low efficiency of false report filtering in wireless sensor networks is the lack of security assurance of the secrets while improving the degree of secret sharing. This paper presented an efficient false report filtering scheme. First we divided the network into clusters and constructed a cluster heads-tree, and then form relationships between the cluster heads were formed. By taking into account of the burden factor of nodes in the keys dissemination procedure, the scheme guarantees that the forwarding nodes closer to a cluster hold more authentication keys for the cluster than those nodes farther from it do, hence, the degree of secret sharing can be elevated. What's more, several keys from the same oral cluster are hold by different nodes in a forwarding cluster so as to guarantee the security of the keys. Analysis and simulation results show that the filtering procedure based on grouped cluster-keys outperforms existing schemes in terms of energy saves, storage overhead and filtering performances.

Keywords Wireless sensor networks, False data filtering, Cluster heads-tree, Burden factor, Keys dissemination

1 引言

无线传感器网络(Wireless Sensor Networks,简称WSNs)在军事和民用领域具有广泛的应用,如环境和交通监测、灾难救助甚至人体心脏监视,因此对WSN的研究是一个非常活跃的研究领域[1]。传感器节点通常部署在野外或者是敌方区域,簇头一般需要经过多跳才能将数据传送给基站,攻击者可以通过俘获节点并利用存储在节点内的秘密信息捏造事实上不存在的虚假事件,恶意篡改正在传送的数据包,发送重复数据包等[2]。若不加防范,这些虚假数据会引发错误警报,干扰用户决策,消耗有限的网络资源。因此,如何快速识别并过滤WSN中的虚假数据是一个具有挑战性的课题[3,4]。

WSN中虚假数据的识别和过滤研究已经取得了一些进展[3-11]。这些方法在技术上主要借鉴数字签名的思想,在待发送的数据包后额外附加 T 个MAC信息[3],并在数据转发的过程中对数据包中MAC的正确性进行验证,从而实现对虚假数据的识别和过滤。然而,由于缺少一种适合WSN的密钥分发方案,在提高密钥共享度的同时无法保证密钥的安全性,因此已有机制仍然存在过滤效率低下的问题;此外,它们没有考虑各节点的能量开销均衡性,导致靠近SINK的节点容易耗尽其能量而失效。密钥共享度是指一种安全机制在网络节点间部署的验证密钥的密度,是衡量过滤效率的关键指标。增大密钥共享度可以提高中转认证的概率,但同时必须给每个节点装载较多的秘密信息,从而造成密钥安全性下降。

基于以上考虑,本文提出一种保障密钥安全性的高效率虚假数据过滤机制。网络成簇后,把节点号与簇进行绑定,将妥协节点的破坏限制在本地。然后,利用生成树算法构造簇

到稿日期:2012-11-10 返修日期:2013-03-21 本文受国家自然科学基金项目(61103245),广西自然科学基金项目(2012GXNSEBA053163),湖南省教育厅科学研究一般项目(13C205)资助。

赵巾帼(1965-),女,硕士,副教授,CCF会员,主要研究方向为无线传感器网络,E-mail:zjg@hnpu.edu.cn;朱凌志(1979-),男,硕士,讲师,主要研究方向为无线传感器网络、物联网技术;罗庆云(1965-),男,教授,主要研究方向为网络安全;梁俊斌(1979-),男,博士,副教授,主要研究方向为无线传感器网络。

头树，基于负荷指数在关联簇之间进行验证密钥分发。源簇给中转簇分发多个验证密钥，且这些密钥由中转簇内不同节点存储，在提高密钥共享度的同时保障密钥安全性。此外，方案确保距离源簇近的中转簇所存储的验证密钥数量比距离较远的簇多，从而均衡验证密钥在网络节点中的分布。最后，基于簇组密钥对数据包进行认证和过滤。理论分析及仿真实验表明，该机制在提高虚假数据过滤效率的同时具备较低的能量消耗和存储开销。

2 相关工作

按照所采取的加密技术，可以将已有虚假数据过滤方案分为3大类：(1)基于对称密钥的方案；(2)基于公开密钥的方案；(3)基于数字水印技术的方案。

Fan Ye等率先利用对称密钥技术提出了数据包转发过滤机制SEF[3]。SEF将一个全局密钥池分成$n(n>T)$个密钥分区，每个分区包含m个密钥，且只有Sink节点知道全局密钥池的所有信息(包括密钥和密钥ID)。每个节点随机选取一个密钥分区存储。在转发过滤阶段，若拥有与检测节点相同的密钥，则中间节点利用该密钥重新产生一个MAC，并验证数据包中所携带的MAC的正确性。然而，由于没有将密钥与检测区域进行绑定，攻击者一旦俘获了T个不同的密钥分区，即可任意伪造假数据包而不被中转节点识别。

Zhu等提出了一种交叉、逐步的验证机制IHA[4]。IHA将节点组织成簇，并假设每个簇内包含至少$T+1$个节点。成簇后，网络将执行一个“协作发现”的过程在节点间建立密钥共享关系。在中转过滤阶段，各节点对数据包中由其上游协作节点所产生的MAC进行验证，一旦验证成功，再重新计算一个MAC并替换掉上游协作节点所产生的MAC，如此交叉、逐步地对数据包进行验证。然而，确定性的验证方式容易因路由变化而失效，且节点在部署后将裸密钥直接在节点间进行传输，容易造成密钥泄漏。

Li等提出了一种基于簇组织和投票机制的虚假数据过滤方案PVFS[5]。PVFS将节点组织成簇，并通过簇合并策略保证每个簇所包含的节点数量多于T个。接下来，各个簇头建立一条到Sink的最短路径。在转发过程中，转发簇头节点以一定概率对数据包进行验证。在转发过程中，与SEF、IHA等方案不同的是，中间节点检测到错误的投票后，并不马上将数据包丢弃，而是继续转发，仅当错误的投票达到某个累计阈值后，才将数据包丢弃。PVFS增加了假包在网络中传输的距离，不利于节省网络能量。此外，簇头之间需要以比普通节点更大的通信半径进行数据转发，簇头容易因能量耗尽而死亡。

Yang等提出了一种基于密码交换(commutative cipher)的路由过滤机制CCEF[10]。CCEF基于公钥模型，采用query response会话机制。在部署前，CCEF给每个节点配置一个互不相同的ID和密钥(称为主密钥)，并给Sink预置一个query-response会话序列。此外，每个节点还与Sink共享同一会话密钥。转发节点在不知道会话密钥的情况下，可以通过密码交换对数据包中的会话MAC进行验证，从而将假包过滤掉。当接收到数据包后，Sink对数据包中的检测节点利用主密钥产生的MAC的正确性进行验证，最终过滤所有假包；如果检测到假包，则更新会话密钥。

Ayday等提出基于位置的密钥(将节点位置、节点ID及节点私钥绑定)和基于位置密钥的邻居认证方案，在此基础上进一步提出了基于位置的门限数字签名的方法过滤虚假数据的机制LCNS[12]。邻居节点之间基于位置密钥建立对偶密钥，将妥协节点的破坏性限制在本地，防止串谋攻击，其引入带GPS功能的robots，为节点提供位置信息。收集到数据后，汇聚节点利用秘密共享算法将汇聚数据分成多个小份(share)，然后以单元块(cell by cell)的方式将每个share传递给Sink。由于数据包中包含一棵哈希树，因此转发节点能够对share进行验证。该机制的不足之处就是需要GPS和机器人的支持，以及对偶密钥的建立和维护等，开销太大。

Feng等提出了利用数字水印来保护传感数据的知识产权的机制[14]。国内学者Peng等论述了将数字水印应用到无线传感器网络中的数据安全保护，其水印信息的嵌入是通过SPPW技术来实现[15]的。Yi等则提出利用协作水印技术和多重半脆弱水印模型来过滤无线传感器网络中的虚假数据[16]。

对称密钥技术的计算复杂性小、实现简单，从能量有效及实用性等角度来说，为资源有限的传感器网络所青睐[19]。公钥机制对存储空间和计算能力要求较高，难以顺利应用在性能有限的无线传感器网络中。数字水印技术目前仍处于初级探索阶段，很多技术尚未成熟，因此在无线传感器网络中的应用还具有很大的局限性。本文基于对称密钥技术，主要研究过滤效率高、均衡节点开销的虚假数据过滤策略。

3 系统模型及相关假设

假设每个传感器节点的初始能量均为e_0，且每个节点具有唯一的ID标识。假设节点在布撒后的一段较短的时间内是安全的，本文利用成簇算法[5]将节点组织成簇，每个簇内包含的节点数为L，簇和簇头的标识都是C_i。假设节点的分布足够密集，簇内每次突发事件都有多个节点检测到。簇头使用比普通节点更大的传输半径进行通信，普通节点除了进行数据感知、签名以及验证外，仅与本簇簇头通信。

攻击者俘获节点后，可以利用存储在节点中的秘密信息伪造虚假数据报告并发送到网络中，发动虚假数据注入攻击[3]。此外，攻击者还可以利用妥协节点篡改传输中的合法数据包，或者发动重放攻击[3,4]等，但本文仅针对WSN中虚假数据注入攻击提出一种防范方案。若网络中特别是单个簇内有多个节点被俘，攻击者可以发动任何形式的攻击使得整个网络迅速瘫痪。因此我们假设每个簇内仅有少量节点被俘，且被俘节点数$N_c<T$，其中T为合法报告包含的签名数量。

假设SINK节点拥有全局密钥信息，能量充足，具备强大的计算和存储能力，作为虚假数据过滤的最后一道屏障，它可以识别并过滤所有漏过中转验证而到达基站的虚假数据。

4 高效的虚假数据过滤方案

4.1 预备知识

定义1(负荷指数B) 传感器网络中叶子节点的负荷指数为1，非叶子节点的负荷指数等于其各个子节点的负荷指

数之和，即

$$\begin{cases}\text{When } C_i \text{ is a leaf in the tree,}\\ \text{Then } B_i=1;\\ \text{Otherwise, note the sons of } C \text{ as } C_m(1\leqslant m\leqslant n),\\ B_i=\sum_{m=1}^{n}B_m\end{cases} \tag{1}$$

定义 2(关联关系 R)　关联关系 $R(M,N)$是指簇头树中源簇 M 和中转簇 N 之间的一种协作关系，通过将源簇内部分节点的主密钥分发给中转簇内节点进行存储以建立密钥共享关系，中转簇能够在数据转发过程中对源簇所产生的数据进行认证。其中源簇分发给中转簇的验证密钥通常有多个，我们将关联簇之间这种成组的验证密钥分发方式称为簇组密钥分发，簇组密钥的数量称为该关联关系的度，可通过簇头负荷指数计算得到，即

$$R_M^N=\left\lceil\frac{L}{B_N}\right\rceil\times\frac{B_M}{B_N} \tag{2}$$

4.2　初始化及验证密钥分发

节点部署前，密钥服务器给每个传感器节点装载一个不重复的来自全局密钥池 $G=\{K_i:0\leqslant i\leqslant N-1\}$的密钥，该密钥称为节点的主密钥。此外，本文采用对偶密钥管理机制[17,18]实现簇头和普通节点之间以及相邻簇头之间传输信息的保密性。

网络部署成簇后，假设簇头构成的集合为 $CH_v=\{C_1,C_2,\cdots,C_i\}$。由于每个簇内的节点数均为 L，且节点初始能量相同，因此每个簇内所有节点的初始能量之和也相同(以簇头指代簇，这里简称每个簇头的初始能量相同)。利用文献[20]中算法，将簇头集 CH_v 代替节点集 V 作为输入，即构造出一棵最大化网络生命周期的簇头树，如图 1 所示。该簇头树具备以下特性：(1)以 SINK 为根节点；(2)除 SINK 外，树中节点均为簇头；(3)因各簇头初始能量相同，故距离 SINK 跳数相同的节点所拥有的儿子节点数量的方差较小，结合定义 1 可知，网络中负荷因子相差不大的簇头所拥有的儿子节点数量也比较均衡。

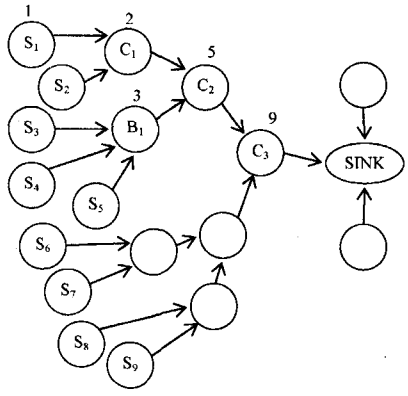

图 1　簇头生成树

然后，在簇头树中定义关联簇，并基于关联簇进行验证密钥分发，图 2 描述了这一过程。下面以图 1 中源簇 S_1 为例，介绍验证密钥分发的过程，记图 1 中源簇 S_1 到 SINK 的路径为 p:(S_1,C_1,C_2,C_3,SINK)。首先计算各中转簇头的负荷因子。根据定义 1，各叶子节点 S_1,…,S_9 的负荷因子为 1，节点 C_1 为 S_1 和 S_2 的父节点，其负荷因子等于 2，同理可得节点 C_2 和 C_3 的负荷因子分别为 5 和 9，图 1 中节点上方的值为节点负荷因子。其次，在路径 p 中，S_1 依次计算与各上游簇的关联度。令 $L=7$，根据定义 2 可知源簇 S_1 关于中转簇 C_1 的关联度为 4，同理可得源簇 S_1 关于中转簇 C_2,C_3 的关联度分别为 1，1。

关联度计算完毕后，源簇 S_1 依次向各中转簇分发验证密钥。首先，源簇簇头 S_1 随机挑选出 ω_1 个簇内节点，簇 C_1 中也相应地选出相同数量的存储空间较大的节点(均包括簇头)，形成两个候选节点集。接下来，两个候选节点集中的节点随机地一一映射，形成 ω_1 个节点对。最后，利用文献[17]中的算法在每个节点对之间建立一个会话密钥，源节点利用该会话密钥将主密钥安全地传输给目的节点。例如，源簇 S_1 随机指派 4 个节点采用会话方式将验证密钥传输给中转簇 C_1 中对应的节点，分发关联度为 4 的簇组密钥。

接下来，源簇 S_1 从剩余的 $L-\omega_1$ 个簇内节点中随机挑选 $\min\{L-\omega_1,\omega_2\}$个与中转簇 C_2 建立密钥共享关系。源簇 S_1 向中转簇 C_3 分发密钥的过程也类似。最后，源簇 S_1 发送一个包含簇内所有节点号的数据包(S_1,N_{k1},N_{k2},…,N_{kL})给所有关联簇的簇头进行存储，实现节点号与簇的绑定。其中 ω_1、ω_2 分别为 S_1 分发给 C_1 和 C_2 的密钥数量。

// * Let the path from source cluster S to SINK be p:(S_1,C_1,C_2,C_3, SINK), where m represents the hops between forwarding cluster C_m and SINK. Now we show the procedure of keys distribution using S as an example. * //

while(S. flag = false),

1. Each forwarding clusters C_1,C_2, …,C_m compute the burden factor according to the following formula,

$$\begin{cases}\text{When } C_i \text{ is a leaf in the tree, Then } B_i=1\\ \text{Otherwise, note the sons of C as } C_m(1\leqslant m\leqslant n), B_i=\sum_{m-1}^{n}B_m\end{cases}$$

2. The degree of association between source cluster S and forwarding cluster C_i is calculated as,

$$R_S^{C_i}=\left\lceil\frac{L}{B_i}\right\rceil\times\frac{B_S}{B_{C_i}}(1\leqslant i\leqslant m)$$

3. Source cluster S pick out λ nodes in-cluster randomly, forwarding cluster C_i pick out λ nodes with larger storage in-cluster too, both including the cluster head, where

$$\lambda=\min\{L-\sum_{j=1}^{i-1}R_S^{C_j},R_S^{C_i}\}$$

then the nodes from two cluster form λ node pair without crossings, $\{(N_{1\_S},N_{1\_C_1}),\cdots,(N_{1\_S},N_{1\_C_i})\}$

4. In each node pair, source node transmits its main key to target node using a session key($1\leqslant j\leqslant\lambda$)

5. Source cluster S bands the information of nodes in-cluster together, and transmits it to each forwarding cluster,

$S\rightarrow C_i:(S,N_{k_1},N_{k_2},\cdots,N_{k_L})$

6. S. flag=true.

7. Endwhile.

图 2　密钥分发算法伪码

簇头树中距离 SINK 在 k 跳之外的簇均给其关联簇分发验证密钥，而与 SINK 在 k 跳距离之内的簇则不往网络中分发验证密钥，其所产生的数据包传输较短的跳数后由 SINK 直接认证，这既可以节省这部分节点的密钥分发和维护开销，还可以减少靠近 SINK 的节点的存储和计算开销。参数 k 的

取值与实际应用中节点分布密度以及过滤概率等参数的设置有关，将在参数分析部分专门讨论。

4.3 数据报告生成

检测到突发事件后，簇头 C_i 收集簇内节点的感知数据并选取一个较完整的值 e 作为本次突发事件的描述，然后将 e 在簇内进行广播。簇内节点 N_i 将 e 与自身感知到的数据进行比较，若误差在一个规定的阈值之内，则利用与 SINK 共享的主密钥 K_i 对感知数据进行加密，生成 $MAC_i:K_i(e)$。接下来，N_i 利用对偶密钥机制[17,18]将签名加密并发送给簇头 C_i。簇头收集簇内 T 个不同节点的签名并形成数据包 R，同样采用对偶密钥机制将数据包加密后发送给下一跳簇头。

$$R:\{e,N_1,N_2,\cdots,N_T,MAC_1,MAC_2,\cdots,MAC_T\} \quad (3)$$

4.4 中转认证

当中转簇头 C_m 从上一跳簇头 C_n 接收到来自源簇 S_i 的数据包时，利用与 C_n 共享的对偶密钥解出数据包 R，然后按照以下步骤对 R 进行验证。

首先检查数据包中的节点号是否属于同一个源簇，以及数据包中 MAC 的数量是否为 T 个。若节点号或者 MAC 数量不满足要求，则直接丢弃 R；反之，则簇头检索簇内节点的验证密钥索引表，若簇内节点没有数据包 R 的验证密钥，则直接转发数据包；如果簇内节点拥有 R 的验证密钥，则簇头将数据包 R 的一部分，包括 e、待验证的签名以及相应的检测节点号，发送给簇内认证节点，并启动一个计时器 η。例如，假设数据包 R 中包含检测节点 N_b 产生的 MAC_b，且簇内节点 N_a 存储了 N_b 的验证密钥 K_b，则簇头 C_m 将短包(e,N_b,MAC_b)发送给 N_a。节点 N_a 利用验证密钥 K_b 重新对 e 产生一个签名并与 MAC_b 进行比较，若两个签名相等，则 N_a 发送验证成功的消息给簇头，反之则发送验证失败消息。簇头与簇内节点之间的通信采用对偶密钥管理机制[17,18]保障安全性。若簇内有数据包的多个验证密钥，则其余验证节点采用类似方式进行验证。

簇头 C_m 在等待时间段 η 后，若接收到的所有验证报告均为验证成功消息，则将数据包 R 转发给下一跳节点，反之，只要接收到一个验证失败消息，就丢弃数据包 R。

4.5 SINK 过滤

SINK 节点拥有全局密钥中的密钥信息，能量充足，且具备强大的计算和存储能力。当 SINK 接收到数据包时，对所有的签名进行重新校验，最终过滤掉所有漏过中转验证而最终到达的虚假数据。

5 性能分析

5.1 密钥分发效率及安全性分析

假设 C_{d0} 和 C_{d1} 为源 S 到 SINK 的转发路径中的两个中转簇，距离 SINK 的跳数分别为 d_0 和 $d_1(d_0\geqslant d_1)$。根据定义 2 知，与源 S 距离越小的中转簇所存储的验证密钥数量越多。另一方面，由负荷指数的定义可知 $B_{d0}\leqslant B_{d1}$，而负荷因子实际上刻画的是节点在网络中的平均通信能量负荷，即中转簇 C_{d0} 能量开销比簇 C_{d1} 大，存储了来自更多源簇的验证密钥。结合这些因素可知，验证密钥在网络节点中的总体分布是比较均匀的，从而均衡了各节点的通信能量开销。

对于节点随机均匀分布的网络，成簇后，其中簇的分布也比较均匀。结合簇头树的特性可知，距离 SINK 跳数相同的簇所拥有的子节点数也较均衡。为简便起见，接下来的分析中我们都采用二叉簇头树进行分析。表 1 列出了 8 跳二叉簇头树中运行密钥分发算法后各中转簇所存储的验证密钥情况，其中距离 SINK 3 跳之内的簇不往网络中分发验证密钥。从表中可以看出，距离 SINK 较近的簇存储的验证密钥数量较小，而其它中转簇所存储的验证密钥数量较平均。从第 8 跳源簇的分发结果看出，距离源簇越近的中转簇所存储的验证密钥数量比距离较远的中转簇多。

节点号与簇头绑定，攻击者俘获的来自不同簇的密钥无法用来发动同谋攻击，故攻击者必须俘获同一个簇内的至少 T 个节点才能伪造出不被中转节点识别的假数据，攻击难度较大。此外，每个中转节点存储的多个验证密钥分别来自不同的源簇，这些验证密钥也无法同时用来伪造假数据包，故我们的方案保证了密钥的安全性。

表 1 二叉簇头树密钥分发结果

Hops to SINK	Burden factor	Distribution of 8^{th} hops node	Distribution of single path	Total keys
1	2^9	0	1	8
2	2^8	0	2	12
3	2^7	0	6	16
4	2^6	0	6	16
5	2^5	0	6	16
6	2^4	0	6	16
7	2^3	1	6	16
8	2^2	2	5	12

5.2 虚假数据过滤效率

若簇 S 中有 N_c 个节点被妥协(包括簇头)，攻击者为了捏造一个假数据包，必须伪造$(T-N_c)$个假 MAC。因中转簇 C_i 中存储了来自源簇 S 的验证密钥，故簇 C_i 碰巧拥有这$(T-N_c)$个验证密钥中的一个的概率为

$$p_1=1-(1-\frac{T-N_c}{L})^{R_S^{C_i}} \quad (4)$$

从源簇 S 到 SINK 的 h 跳中转簇共存储了 M_h 个 S 的验证密钥，则源簇 S 伪造的假数据包传输 h 跳被过滤的概率为

$$M_h=\sum_{i=1}^{h}R_S^{C_i} \quad (5)$$

$$\begin{cases}\text{If } M_h=L, & p_h=1\\ \text{If } M_h<L, & p_h=\dfrac{T-N_c-1+M_h}{L}\end{cases} \quad (6)$$

图 3 所示为本方案及 SEF[3]的过滤概率 p 随传输跳数 h 的变化曲线，其中 $T=5$，$L=7$，同一个簇内的妥协节点数量考虑了 2 和 4 两种情况，网络节点形成二叉簇头树。

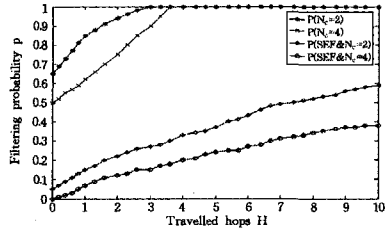

图 3 过滤概率 P 随传输跳数 H 的变化曲线

从图 3 中可以看出，本方案的过滤概率随传输距离增加而增大，第 1 跳的过滤概率达到 86%，前 3、4 跳即可保证将假数据过滤。本方案的第 1 跳过滤概率比 SEF 中传输 10 跳

的累计过滤概率还高，这是因为本方案中第一跳中转簇即存储了源簇的多个验证密钥，能对数据包中多个 MAC 同时进行认证，而 SEF 中每个中转节点仅能概率性地对数据包进行认证，传输多跳的累计过滤不高。另一方面，由于本方案中距离源簇较近的数个中转簇即存储了源簇的所有验证密钥，因此确保了在数跳之内将假包过滤掉。而 SEF 中假包在网络中传输的跳数较大，部分假包甚至能最终传输到 SINK。此外可以看出，当源簇内妥协节点数量增加时，SEF 的过滤概率下降明显，而本方案的过滤概率变化不大，因此，本方案抵抗节点妥协的弹性比已有方案好。

5.3 能量节省

过滤机制消耗的能量主要来自 3 个方面：(1)转发数据包的通信开销；(2)验证 MAC 时的计算开销；(3)数据包生成及验证时，簇头与簇内节点之间的通信开销。由于 SINK 节点能量充足，因此我们不考虑其能量消耗。文献[12]中指出，MAC 计算所消耗的能量跟传输数据包的能量相比可以忽略不计。此外，簇头与簇内节点传输的短消息仅包括事件 e、一个 MAC 及节点号，其传输一跳所消耗的能量也不考虑。因此在这里主要讨论转发数据包所消耗的能量。

方案给每个数据包增加了 T 个节点号和 T 个 MAC，这些额外负荷会增加传输中的能量消耗。但另一方面，方案通过尽早过滤大部分虚假数据来节省能量。

我们采用文献[3,5]中的模型进行能耗的量化分析。令 L_r，L_n 以及 L_M 分别表示不采用安全机制时的纯数据包长度、节点号长度以及 MAC 的长度。本方案中数据包长度为 $L_r' = L_r + (L_n + L_m) \times T$，传输(1 合法数据 $+\beta$ 虚假数据)所消耗的能量为 E，为便于比较，记采用 SEF 传输相同的数据量所消耗的能量为 E_{SEF}，传输距离均为 H 跳。

$$E_{SEF} = (1 + \frac{L_M + L_n}{L_r} \cdot T)(H + \beta \frac{1-(1-p_0)^H}{p_0}),$$

$$\text{here } p_0 = \frac{k_0(T - N_c)}{N} \tag{7}$$

$$\begin{cases} \text{If } M_{H_0-1} \leqslant L-(T-N_c-1) \leqslant M_{H_0}, \text{and } H_0 \leqslant H, \\ \quad \text{Then } E = [1 + \frac{(L_M+L_n)}{L_r} \cdot T] \times [H + \beta(H_0 - \sum_{i=1}^{H_0-1} p_i)] \\ \text{If } M_H < L-(T-N_c-1), \\ \quad \text{Then } E = [1 + \frac{(L_M+L_n)}{L_r} \cdot T] \times [H + \beta(H - \sum_{i=1}^{H-1} p_i)] \end{cases} \tag{8}$$

图 4 所示为本方案和 SEF 能耗的对比情况，其中传输距离为 20 跳，同一簇内妥协节点数为 $N_c = 4$，其它参数取值分别为 $L_r = 24$bytes，$L_n = 10$bits，$L_M = 64$bits，$N = 1000$，$k_0 = 50$。从图中可以看出，与 SEF 机制相比，本方案在能耗节省方面优势明显。

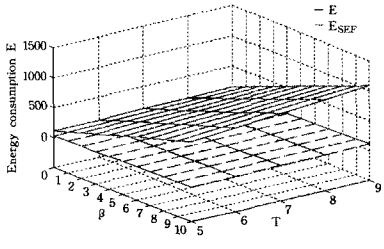

图 4 $H=20$，$N_c=4$ 时与 SEF 能耗对比

5.4 存储开销分析

本方案中普通节点的存储开销包括两方面：存储各源簇的验证密钥以及节点自身主密钥。对于簇头节点来说，还需存储各源簇的节点号。文献[3,6]指出，节点号的大小仅为数个字节，其存储开销与密钥的存储开销相比可以忽略不计，故簇头节点与普通节点的存储开销基本一致。簇 C_i 分配得到的验证密钥总量为

$$\sum_{j=1}^{B_i} R_{S_j}^{C_i} \tag{9}$$

故簇内每个节点存储的密钥数量约为

$$\frac{1}{L} \cdot \sum_{j=1}^{B_i} R_{S_j}^{C_i} \tag{10}$$

式中，B_i 为簇头 C_i 的负荷指数，L 为簇内节点数。以二叉簇头树为例，我们的方案仅需为每个节点配备约 0.3kB 的存储空间，根据当前主流节点的配置，如 UCB 研制的 MICA2 节点所使用的存储器为 4kB 的 SRAM 和 128kB 的 ROM[2,10]，显然能满足需求。

假设一个密钥的大小为 64bits，表 2 列出了各种主流过滤机制的存储开销和过滤性能。从表中可以看出，与已有机制相比，我们的方案具备较好的过滤性能，且拥有较低的存储开销。

表 2 虚假数据过滤机制的性能比较

Schemes	Storage burden	Filtering performance
SEF[3]	0.4kB	Weak
IHA[4]	4kB	Moderate
PVFS[7]	1.2kB	Weak
Fault localized filtering scheme[10]	1.5kB	Weak
Our scheme	0.3kB	Good

6 性能分析

参数值的合理选取能优化过滤机制的性能。我们对簇内节点数 L、数据包携带的签名数 T、验证密钥部署密度 δ 以及不分发密钥的簇与 SINK 的最大距离 k 等参数的取值进行分析。

假设节点部署密度为 ρ，节点感知半径为 r_s，则感知到同一事件的节点数为 $N_d = \rho\pi r_s^2$，因此在 $L \leqslant N_d$ 范围内只有选取 L 才能保证每个事件有充足的节点被检测到。T 的取值是安全性和能量消耗之间的一个折中，数据包携带的签名越多，攻击难度和传输过程中的能量消耗都越大。此外，T 的大小还受限于节点允许的最大数据包长度。例如，一些低端节点支持的最大数据包长度仅为 36bits，故 T 不能太大。验证密钥部署密度直接影响过滤效率。源簇在其上游中转节点中部署的验证密钥越多，即 δ 越大，数据包认证效率也越高。但是 δ 的取值受限于节点的存储能力，假设一个验证密钥大小为 64bits，存储 25 个验证密钥就需要 200bytes，占用了低端节点较大的存储空间。此外，一个中转簇需要存储多个源簇的验证密钥，也决定了 δ 不能太大。k 的取值主要与实际应用中的网络规模有关，对于大型网络，适当调大 k 可以节省很大部分节点的密钥分发开销以及密钥维护开销而不影响过滤能力，至于中小规模网络，k 对能量消耗以及过滤能力的影响均不大。

7 性能模拟

我们利用仿真实验进一步验证性能分析结果。在 400 * 40m² 区域中，约 700 个传感器节点随机均匀分布，区域左侧的 S 为静态源节点，每 2s 发出一个虚假数据包；右侧为静态的 SINK 节点。S 和 SINK 之间大约为 80 跳，成簇后，S 与 SINK 大约相距 10 个簇头。传输包和接收包的功耗分别为 60mW，12mW。使用一个大小为 1000 的全局密钥池，妥协节点数为 $N_c=4$。限于篇幅，这里仅给出了 $N_c=4$ 时本方案与 SEF[3] 以及 PVFS[5] 在过滤性能方面的对比，仿真结果取多次实验的平均值。

从图 5 中可以看出，基于簇组织的本方案和 PVFS 机制在过滤性能方面明显优于采用一般网络结构的 SEF 机制。此外，本方案所采用的簇组密钥认证方式确保虚假数据在传输 4 跳左右即可过滤，与同样基于簇组织、但采用概率性密钥认证方式的 PVFS 机制相比，过滤性能优势明显，这些都印证了前面的理论分析结果。

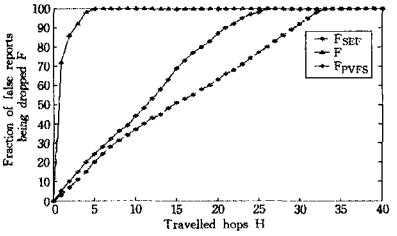

图 5 与 SEF，PVFS 的过滤概率对比

结束语 本文在深入分析 WSN 中虚假数据过滤机制工作效率低下的原因后，提出一种基于簇头生成树和节点负荷因子的验证密钥分发策略以及基于簇组密钥的认证机制。机制在提高密钥共享度的同时可保障密钥安全性，且均衡验证密钥在网络节点中的分布。理论分析及仿真实验表明，基于簇组密钥的认证机制在过滤性能、能量消耗以及存储开销等方面都取得了较好的效果。但是，在实现过滤机制时，仍必须对一些参数进行设定，这将是我们进一步研究的方向。

参 考 文 献

[1] 任丰原，黄海宁，林闯. 无线传感器网络[J]. 软件学报，2003，14(7)：1282-1291

[2] 苏忠，林闯，封富君，等. 无线传感器网络密钥管理的方案和协议[J]. 软件学报，2007，18(5)：1218-1231

[3] Ye F, Luo H, Zhang L. Statistical En-route Filtering of Injected False Data in Sensor Networks[C]// Proceedings of 23th Annual Joint Conference of the IEEE Computer and Communications Societies(INFOCOM 2004). 2004：2446-2457

[4] Zhu S, Setia S, Jajodia S. An interleaved Hop-by-Hop Authentication Scheme for Filtering of Injected False Data in Sensor Networks[C]// Proceedings of the IEEE symposium on Security and privacy(S&P'04). 2004：259-271

[5] Li F, Wu J. A probabilistic voting-based filtering scheme in wireless sensor networks[C]// Proceedings of the International Wireless Communications and Mobile Computing Conference (IWCMC 2006). 2006：255-265

[6] Yu Z, Guan Y. A Dynamic En-route Scheme for Filtering False Data Injection in Wireless Sensor Networks[C]// Proceedings of the 3rd international conference on Embedded networked sensor systems(SenSys'05). 2005：294-295

[7] Zhou L, Ravishankar C. A fault localized scheme for false report filtering in sensor networks[C]// Proceedings of the IEEE International Conference on Pervasive Services(ICPS'05). 2005：59-68

[8] Yang H, Ye F, Yuan Y, et al. Toward resilient security in wireless sensor networks[C]// Proceedings of the 6th ACM international symposium on mobile ad hoc networking and computing (MobiHoc'05). 2005：34-45

[9] Zhang Y, Yang J, Vu H. The Interleaved Authentication for Filtering False Reports in Multipath Routing based Sensor Networks[C]// Proceedings of 20th International Parallel and Distributed Processing Symposium (IPDPS'06). 2006：1-10

[10] Yang H, Lu S. Commutative Cipher Based En-route Filtering in Wireless Sensor Networks[C]// Vehicular Technology Conference(VTC'04). 2004：1223-1227

[11] Wang H, Li Q. PDF：A Public-key based False Data Filtering Scheme in Sensor Networks[C]// Proceedings of the International Conference on Wlreless Algorithms, Systems and Applications(WASA'07). 2007：129-138

[12] Ayday E, Delgosha F, Fekri F. Location-aware security services for wireless sensor networks using network coding[C]// Proceedings of the IEEE Conference on Computer Communications (INFOCOM 2007). 2007：1226-1234

[13] Ren K, Lou W, Zhang Y. Providing location-aware end-to-end data security in wireless sensor networks[C]// Proceedings of the IEEE Conference on Computing and Communicating(INFOCOM 2006). 2006：585-598

[14] Feng J, Potkonjak M. Real-time watermarking techniques for sensor networks[C]// Proceedings of the SPIE Security and Watermarking of Multimedia Contents. California, 2003：391-402

[15] 彭志娟，王汝传，王海燕. 基于数字水印技术的无线传感器网络安全机制研究[J]. 南京邮电大学学报：自然科学版，2006，26(4)：69-72

[16] 易叶青，林亚平，李小龙. WSN 中基于协作水印的虚假数据过滤算法[J]. 软件学报，2010，21(1)：107-118

[17] Liu D, Ning P. Location-Based Pairwise Key Establishments for Static Sensor Networks[C]// Proceedings of the 1st ACM workshop on Security in Ad Hoc and Sensor Networks, SASN'03. 2003：72-82

[18] Liu D, Ning P. Establishing Pairwise Keys in Distributed Sensor Networks[C]// Proceedings of the 10th ACM conference on computer and communications security, ACM'05. 2005：52-61

[19] Zhu S, Setia S, Jajodia S. LEAP：Efficient Security Mechanisms for Large-Scale Distributed Sensor Networks[C]// Proceedings of the 10th ACM Conference on Computer and Communications Security, CCS'03. 2003：：62-72

[20] Hussain S, Islam O. An energy efficient spanning tree based multi-hop routing in wireless sensor networks[C]// Proceedings of Wireless Communications and Networking Conference(WCNC'07). 2007：4383-4388