传感器网络中一种基于地理位置的虚假数据过滤方案

刘志雄，王建新
(中南大学 信息科学与工程学院，湖南 长沙 410083)

摘 要：提出了一种基于地理位置的虚假数据过滤方案 GFFS。在 GFFS 中，节点在部署后将地理位置预分发给部分其他节点存储，每个数据报告必须包含 t 个具有不同密钥分区检测节点的 MAC 以及地理位置，转发节点既对数据分组中包含的 MAC 和地理位置的正确性进行验证，还对地理位置的合法性进行验证。理论分析及仿真实验表明，GFFS 能有效地过滤不同地理区域的多个妥协节点协同伪造的虚假数据，且具备远强于已有方案的妥协容忍能力。

关键词：无线传感器网络；虚假数据过滤；妥协节点；协同攻击
中图分类号：TP393 **文献标识码**：A **文章编号**：1000-436X(2012)02-0156-08

Geographical information based false report filtering scheme in wireless sensor networks

LIU Zhi-xiong, WANG Jian-xin
(School of Information Science and Engineering, Central South University, Changsha 410083, China)

Abstract: A geographical information based false reports filtering scheme (GFFS) in sensor networks was presented. In GFFS, each node distributes its location information to some other nodes after deployment. When a report was generated for an observed event, it must carry not only MACs from t detecting nodes with distinct key partitions, but also locations of these nodes. Each forwarding node checks not only the correctness of the MAC and the locations carried in the report, but also the legitimacy of the locations. Analysis and simulation results demonstrate that GFFS can resist collaborative false data injection attacks and thus can tolerate much more compromised nodes than existing schemes.

Key words: wireless sensor network; false report filtering; compromised node; collaborative attack

1 引言

无线传感器网络(WSN, wireless sensor network)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[1]。传感器节点通常部署在野外或者敌方区域，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，发动虚假数据注入攻击[2]。这类攻击不仅引发错误警报，同时也可以耗尽宝贵的网络资源[3,4]。

鉴于虚假数据的安全威胁，不少学者提出了一些解决办法[3~11]，它们的共同特点是在待发送的数据报告中附加 t 个 MAC(message authentication code)，并在数据转发的过程中对 MAC 进行验证，从而实现对虚假数据的识别和过滤，这里 t 是系统参数。这些方案在过滤性能和安全性等方面都获得了较好的效果，但没有考虑节点密钥与地理位置的相关性，故无法过滤不同地理区域多个妥协节点协同伪造的虚假数据。

本文提出一种基于地理位置的虚假数据过滤方案 GFFS，将节点地理位置预分发给中间节点存储，并在数据报告中同时附带检测节点产生的 MAC

收稿日期：2011-01-04；修回日期：2011-05-20

和地理位置，然后由中间节点在转发过程中同时对数据分组中包含的 MAC 和地理位置进行验证，从而将不同地理区域的多个妥协节点协同伪造的虚假数据过滤掉。

2　相关工作

文献[3]率先提出了传感器网络中虚假数据转发过滤的基本框架 SEF。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。每个节点在部署前随机地从全局密钥池中选取一个密钥分区，并从中任选 k 个密钥进行存储。事件发生后，多个检测节点联合产生一个包含 t 个互不相同的 MAC 数据报告。在数据分组转发过程中，与检测节点拥有相同密钥分区的中间节点以概率 k/m 对数据分组中的一个 MAC 进行验证。在 SEF 中，攻击者只要俘获了 t 个不同密钥分区，即可通过协作伪造出不被识别的假分组。

例如，当 t=5 时，假设攻击者俘获了图 1 中的拥有不同密钥分区的节点 $S_1,\cdots,S_5$，尽管这 5 个节点位于网络中不同的地理区域，但攻击者仍然可以利用它们协同地伪造发生在 A 点(或者其他任何位置)的假分组 R，并通过妥协节点 S_1 发送给 S_1 的邻居节点，而转发节点和 sink 都无法对 R 进行检测和过滤。

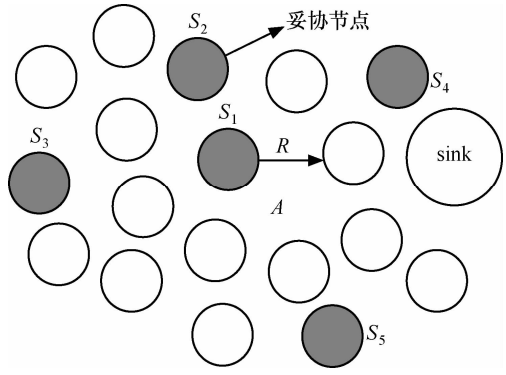

图 1　多个妥协节点协同伪造虚假数据

文献[4]提出了一种交叉、逐步的认证机制 IHA。IHA 将节点组织成簇，簇头建立到 sink 的路径，并基于路径进行交叉式对称密钥分发，然后在转发过程中由中间节点对簇头产生的数据分组进行逐步、交叉的验证。IHA 采用一种确定性的方式进行密钥分发和数据验证，限制了网络中不同区域的妥协节点协同地伪造假数据分组。然而，一旦路由发生变化，这种确定性的数据验证方式也随之失效，重新建立路由并分发密钥需要较大的维护开销。

文献[5]提出了一种基于爬山 (hill climbing)策略和多径路由的虚假数据过滤方案 DEFS。DEFS 将网络节点组织成簇，簇头采用类似“爬山”的策略进行密钥分发：以较高的概率将簇内节点的密钥预分发给离簇头较近的中间节点存储，而以较低的概率将簇内节点的密钥预分发给离簇头较远的节点存储。检测到突发事件后，簇头将数据分组沿多条路径并发地向 sink 转发，中间节点分别以一定概率对数据分组进行验证。爬山策略有利于减轻靠近 sink 节点的工作负担，然而多路径并发传输数据分组的开销太大，不利于节省传感器节点的能量。

文献[6]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。PVFS 将节点组织成簇，源簇头建立一条到 sink 的路径。和 IHA 不同的是，转发节点均为簇头，以概率 d_i/d_0 存储源簇内一个节点的密钥，其中，d_0 和 d_i 分别表示源簇头和转发簇头距离 sink 的跳数。事件发生时，簇头收集簇内 t 个节点产生的投票生成数据分组，转发簇头节点以一定概率对数据分组进行验证。PVFS 由簇头转发数据分组，簇头之间需要以比普通节点较大的通信半径进行数据转发，簇头容易因能量耗尽而死亡。

文献[7]提出了一种不受门限值限制的过滤机制 RSFS。节点与 sink 形成星形拓扑，事件发生后，由一种比普通节点能量强得多的节点作为簇头完成数据聚合，聚合结果必须附加簇内每个普通节点产生的 MAC。目的节点在接收到汇聚结果时通过对聚合结果及附加 MAC 信息的校验，实现对虚假数据的过滤。该方案不受门限值的限制，但假数据分组必须传输到 sink 才能被检测到，而无法由转发节点进行检测和过滤，不利于节省传感器网络的能量。

文献[8]的作者认为以上基于对称密钥机制的安全性不够，提出了一种基于密码交换(commutative cipher)的路由过滤机制 CCEF。该机制假设各节点与基站共享一会话密钥(session key)，路由中报文转发节点不知道报文源节点的会话密钥，但可通过交换密码对报文进行检测，从而提高了安全性。文献[9]利用椭圆曲线密钥技术来改进文献[8]中的方案，进一步提高安全性。文献[10]将整个网络划分为不

重叠的单元(cell)，以单元块(cell by cell)的方式传递数据，采用门限共享技术和公开密钥技术来保证数据的安全。文献[11]在文献[10]的基础上，将网络划分为不重叠单元，在特定的路由上对数据加密进行研究。

最近研究表明[2,4]，公开密钥技术虽然比对称密钥技术具备更好的安全性，但对存储空间和计算能力要求较高，无法顺利应用在性能有限的传感器网络中；而对称密钥技术的计算复杂性较小，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。已有基于对称密钥技术的方案，如 SEF、DEFS、PVFS、RSFS 等，没有将节点密钥与节点地理位置绑定，故无法检测不同地理区域多个妥协节点协同伪造的虚假数据。本文基于对称密钥技术，研究如何过滤不同地理区域多个妥协节点协同伪造的虚假数据。

3 基于地理位置的虚假数据过滤方案 GFFS

3.1 系统模型及相关假设

假设传感器节点分布密度较大，事件 e 发生后，有多于 t 个节点同时检测到。各个检测节点利用密钥对事件进行加密后生成 MAC，然后将 MAC 和位置信息发送给一个中心节点(CoS, central of stimulas)[3]。此外，CoS 联合其他检测节点，采用感知范围叠加的办法对事件发生地点进行定位，得到 L_e，这里 L_e 是取多个节点感知范围重叠区域中某个点的位置。CoS 将事件位置 L_e、各个检测节点的位置以及 MAC 附加在事件 e 后面，生成数据报告。

假设攻击者可以俘获网络中的多个节点，然后利用存储在这些节点中的秘密信息伪造虚假数据报告并发送给邻居节点，发动虚假数据注入攻击[3,4]。假设 sink 节点无法被俘获，且其拥有全局密钥信息，能量充足，并具备强大的计算和存储能力，能够过滤所有最终到达的假数据分组。

3.2 节点部署与初始化

部署前，给每个传感器节点分配唯一的 ID 标识。存在一个全局密钥池 $G=\{K_i:0\leqslant I\leqslant N-1\}$，密钥池大小为 N，分为 n 个不重叠的密钥分区$\{U_i,0\leqslant I\leqslant n-1\}$，每个分区包含 m 个密钥($N=nm$)。每个节点随机选择一个密钥分区，并任取其中 k 个密钥存储。

各个传感器节点 S_i 通过 GPS 或者其他方法[13]可以获得其地理位置，记为 L_i：(X_i, Y_i)，其中，X_i、Y_i 分别表示节点 S_i 在整个监控区域的横坐标和纵坐标。接下来，各节点 S_i 利用 Bubble-geocast 算法[14,15]将 c 个数据分组(S_i, L_i, U_i)分发给中间节点存储，使得中间节点各以概率 c/N_a 存储该数据分组，其中 N_a 为网络中节点总数量，U_i 为节点 S_i 存储的密钥分区索引。

3.3 数据报告生成

事件发生后，检测节点共同选举一个中心节点 CoS。CoS 将感知数值 e 发送给各检测节点，检测节点收到数据 e 后，将自己的感知数值与 e 进行比较，若误差在预定的阈值范围内，则随机选取一个存储的密钥 K_i 对 e 进行加密，生成消息认证码 $M_i:K_i(e)$。接下来，各检测节点将节点号、地理位置以及 MAC 发送给 CoS。中心节点 CoS 将事件位置以及 t 个拥有不同密钥分区的检测节点的节点号、密钥索引、MAC 和地理位置附加在感知数据 e 后面，产生数据报告 R:$\{e, L_e; i_1, i_2, \cdots, i_t; M_{i1}, M_{i2}, \cdots, M_{it}; j_1, j_2, \cdots, j_t; L_{j1}, L_{j2}, \cdots, L_{jt}\}$。其中，$i_1, i_2, \cdots, i_t$ 为密钥索引，$j_1, j_2, \cdots, j_t$ 为节点号。

3.4 转发过滤

由于随机预载了一个密钥分区中的部分密钥以及部分节点的地理位置和密钥分区索引，故中间节点可以以一定概率对数据分组中的 MAC、节点位置以及密钥索引进行验证。

当接收到转发数据分组 R 时，中间节点 S_i 首先对数据分组中附带的节点 ID、密钥索引、MAC 以及地理位置的数量进行检查，然后核对这些密钥索引是否属于不同密钥分区，接下来检查各个地理位置与事件发生位置 L_e 之间距离的合法性，最后验证 MAC、地理位置和密钥索引的正确性。中间节点对数据分组 R 的具体验证步骤如下。

1) 检查数据分组 R 中是否各包含 t 个节点 ID、密钥索引、MAC 以及地理位置。若任何一项的数量不符合要求，则丢弃 R。

2) 检查 t 个密钥索引是否属于不同的密钥分区，否则丢弃 R。

3) 检查各个地理位置与 L_e 之间的距离是否都小于节点感知半径 r_s，否则丢弃 R。

4) 若存储了 R 中某个检测节点 $S_{jv}(1\leqslant v\leqslant t)$的地理位置 L_{jv} 和密钥分区索引 U_{jv}，则先判断 R 中附带的 S_{jv} 的密钥索引 i_v 是否属于密钥分区 U_{jv}，接下来判断 R 中附带的 S_{jv} 的地理位置是否与存储的 L_{jv} 相等。若任何一项不满足，则丢弃 R。

5) 若存储了 R 中某个密钥索引 i_v，则利用存储的密钥 K_{iv} 对 e 重新计算一个 M 并与 R 中附带的 M_{iv} 比较，若二者差值小于某个阈值 ε，则说明 M_{iv} 正确，否则丢弃 R。

最后，若以上验证都通过，则将 R 转发给下一跳节点。

图 2 为转发过滤算法伪代码。

/* on receiving report R */

1. Check that t $\{i_v, M_{iv}\}$ tuples exist in R; drop R otherwise.

2. Check the t key indices $\{i_v, 1 \leqslant v \leqslant t\}$ belong to t distinct partitions; drop R otherwise.

3. Check that t $\{j_v, L_{jv}\}$ tuples exist in R; drop R otherwise.

4. Check that $(|L_e, L_{jv}| \leqslant r_s, 1 \leqslant v \leqslant t)$; drop R otherwise.

5. If it has one location $L \in \{L_{jv}, 1 \leqslant v \leqslant t\}$ and key partition U, it first check the key index i_v in R belongs to U; drop R otherwise. It then check the location in R is the same as L; drop R otherwise.

6. If it has one key $K \in \{K_{iv}, 1 \leqslant v \leqslant t\}$,it computes $M = K(e)$ and see if the corresponding M_{iv} is the same as M; drop R otherwise.

7. Send R to the next hop.

图 2　转发过滤算法伪代码

3.5　Sink 验证

Sink 节点拥有全局密钥信息以及所有节点的位置信息，且具备强大的计算、存储能力以及充足的能量，能够过滤所有漏过中转验证而最终到达的虚假数据。当 sink 接收到数据分组时，对所有的 MAC 以及检测节点位置信息进行重新校验，如果所有 MAC 以及位置信息都正确，则接收数据分组并执行相应的决策，否则将数据分组丢弃。

4　性能分析与仿真结果

4.1　防范协同攻击的能力

为了简化分析，假设网络监控区域为直径 $2r_a$ 的圆形区域，N_a 个感知半径为 r_s 的传感器节点随机部署在里面。

SEF 是在数据分组中附带 t 个 MAC，由中间节点通过 MAC 验证来过滤虚假数据。攻击者在俘获了网络中任意地理区域的，且拥有不同密钥分区的 t 个节点后，即可通过协作伪造出 SEF 无法过滤的假分组。例如，当 t=5 时，攻击者俘获了图 3 中的拥有不同密钥分区的节点 S_1，…，S_5 之后，即可伪造出 SEF 无法过滤的假分组。

而 GFFS 在数据分组中同时附带 t 个 MAC 和检测节点的地理位置，中间节点既要对 MAC 进行验证，又要验证地理位置的正确性和合法性(即各个检测节点的地理位置与事件位置 L_e 之间的距离不能大于节点感知半径 r_s)。若攻击者利用节点 S_1，…，S_5 在 GFFS 中伪造假分组，由于节点 S_1 和 S_4 之间的距离大于 $2r_s$，故无论怎样伪造 L_e，都无法使得节点 S_1，…，S_5 与 L_e 之间的距离都小于 r_s，因此假分组将在第一跳便被过滤掉。因此，和 SEF 不同，GFFS 可以防范不同地理区域的多个妥协节点协同伪造虚假数据。

4.2　妥协容忍能力

根据 GFFS 的过滤规则，攻击者若想伪造出中间节点无法识别的假分组，必须俘获 t 个拥有不同密钥分区、且都位于某个直径为 $2r_s$ 区域内的节点。如图 3 所示，当 t=5 时，假设攻击者俘获了直径为 $2r_s$ 区域 A 中的节点 S_6，…，S_{10}，且这些节点分别存储不同的密钥分区，则可以伪造发生在区域 A 中圆心位置的事件 L_e，并构造 t 个正确的 MAC 以及 t 个合法的地理位置，使得 GFFS 无法识别和过滤。

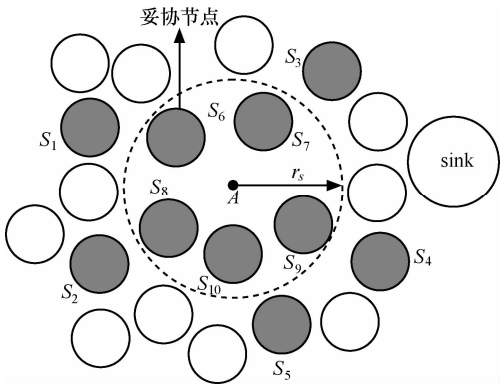

图 3　直径为 $2r_s$ 区域内的妥协节点

定理 1　攻击者在随机俘获网络中的 N_c 个节点后($N_c \geqslant t$)，可以获得至少 t 个不同密钥分区的概率为

$$p_S = \frac{\sum_{i=t}^{n}\left[C_n^i \sum_{j=0}^{i}((-1)^j C_i^{i-j}(i-j)^{N_c})\right]}{n^{N_c}} \tag{1}$$

证明　先考虑攻击者获得刚好 t 个不同密钥分区的情形。首先，从 n 个密钥分区中任取 t 个的方法有 C_n^t 种。其次，记 N_c 个妥协节点组成的集合为 B_1，选取出来的 t 个密钥分区的集合为 B_2。那么，集合 B_1 中的每个节点各从集合 B_2 中任取 1 个密钥分区，使得 B_2 中每个密钥分区都至少被取到 1 次，

其方法数量有

$$\begin{aligned}Q_a &= C_t^t \cdot t^{N_c} - C_t^{t-1} \cdot (t-1)^{N_c} + C_t^{t-2} \cdot (t-2)^{N_c} - \cdots + \\ &\quad (-1)^{t-2} \cdot C_t^2 \cdot 2^{N_c} + (-1)^{t-1} \cdot C_t^1 \cdot 1^{N_c} \\ &= \sum_{j=0}^{t-1} ((-1)^j C_t^{t-j} (t-j)^{N_c}) \\ &= \sum_{j=0}^{t} ((-1)^j C_t^{t-j} (t-j)^{N_c}) \end{aligned} \tag{2}$$

因此，刚好获得 t 个密钥分区的方法数量为 $C_n^t \times Q_a$。同理可求出刚好获得 t +1，t +2，…，n 个密钥分区的方法数量。因此，至少获得 t 个密钥分区的方法总数为

$$\sum_{i=t}^{n} \left[C_n^i \sum_{j=0}^{i} ((-1)^j C_i^{i-j} (i-j)^{N_c}) \right] \tag{3}$$

最后，N_c 个节点中的每个节点各从 n 个密钥分区中任选 1 个的方法总数为 n^{Nc}。故攻击者至少获得 t 个拥有不同密钥分区节点的概率为 P_S。得证。

定理 2 攻击者在随机俘获网络中的 N_c 个节点后($N_c \geqslant t$)，获得至少 t 个拥有不同密钥分区，且都位于某个直径为 $2r_s$ 的圆形区域中的妥协节点的概率为

$$p_G = \sum_{i=t}^{N_c} \frac{C_{N_c}^i \left(\frac{r_s^2}{r_a^2} \right)^i \left(1 - \frac{r_s^2}{r_a^2} \right)^{N_c - i} P_n^i}{n^i} \tag{4}$$

证明 先考虑刚好获得 t 个拥有不同密钥分区，且都位于某个直径为 $2r_s$ 的区域中(假设为区域 A)妥协节点的情形。由于每个节点以概率 r_s^2/r_a^2 分布在 A 中，则 A 中刚好有 t 个节点的概率为 $p_b=C_{N_c}^t \times (r_s^2/r_a^2)^t \times (1-r_s^2/r_a^2)^{Nc-t}$。接下来，这 t 个节点被分配了不同密钥分区的概率为 $p_c=P_n^t/n^t$。因此，攻击者刚好获得 t 个拥有不同密钥分区，且都位于区域 A 中的妥协节点的概率为 $p_d = p_b \times p_c$。

同理可求出刚好获得 t +1，t+2，…，N_c 个拥有不同密钥分区，且都位于区域 A 中的妥协节点的概率。因此，至少获得 t 个拥有不同密钥分区，且都位于某个直径为 $2r_s$ 区域中的妥协节点的概率为 p_G。得证。

图 4 给出了当 t=5，r_s/r_a =1/2，n=15 时，p_S 和 p_G 的理论分析曲线和仿真实验结果，其中仿真结果是在相同参数条件下随机测试 10 000 次的平均值。如图 4 所示，攻击者在俘获少量节点后即可以以较高概率攻破 SEF，而攻击者需要俘获较多的节点才能攻破 GFFS，例如当 N_c =10 时，攻击者有 99.2% 的概率可以攻破 SEF，而攻破 GFFS 的概率仅为 3%。这是因为 SEF 没有考虑节点密钥与地理区域之间的关联关系，攻击者只要随机从网络中俘获 t 个密钥分区便可攻破 SEF；而 GFFS 通过引入地理位置信息可以将妥协节点的破坏性限制在本地，攻击者必须俘获 t 个拥有不同密钥分区，且同处于某个直径为 $2r_s$ 的圆形区域内的节点才能攻破 GFFS。理论分析和仿真实验结果都说明，GFFS 的妥协容忍能力远强于 SEF。

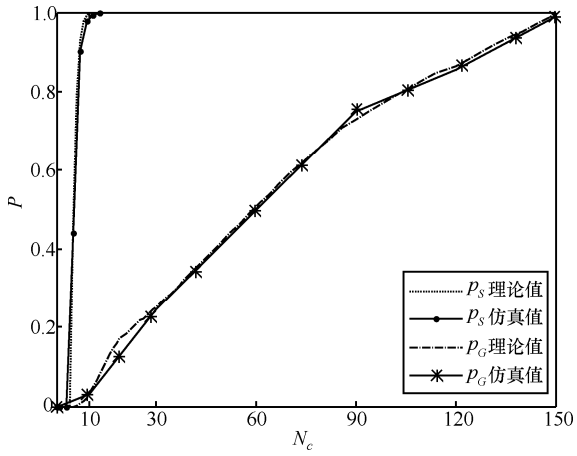

图 4 p_S、p_G 的理论值与仿真结果

4.3 能量开销

GFFS 消耗的能量主要来自 4 个方面：①各个节点在初始化阶段分发 c 个地理位置数据分组的开销；②CoS 与其他检测节点在产生数据分组时的通信开销；③中间节点进行 MAC 验证时的计算开销；④中间节点转发数据分组的通信开销。

在初始化过程中，各个传感器节点分发的数据分组很短（仅包括节点 ID、节点位置和密钥分区索引），且分发时间较短；在生成数据分组的过程中，CoS 和其他检测节点之间传输的数据分组长度也很短（仅包括 MAC、节点号和地理位置各 1 个），且传输距离仅为 1 跳，这 2 种能耗与完整数据分组在网络中传输的能耗相比可以忽略不计。另外，文献[3]指出，MAC 计算所消耗的能量比传输数据分组的能量低了几个数量级，故在此也忽略不计。因此在这里只考虑转发数据分组所消耗的能量。

和 SEF 相比，GFFS 给每个数据分组增加了 t 个节点号及 t 个地理位置。令 I_r，I_n，I_k 以及 I_u 分别表示不采用安全机制时的纯数据分组长度、节点号长度、地理位置长度以及 MAC 的长度，则 GFFS

和 SEF 中数据分组长度分别可表示为 $I_{r0}=I_r+I_k+(I_n+I_k+I_u)t$，$I_{r1}=I_r+(I_n+I_u)t$。例如，当 I_r＝24byte，I_n＝10bit，I_k＝10bit，I_u＝64bit 时，GFFS 数据分组和 SEF 数据分组的长度分别为 77.8byte 和 70byte。这些额外的负荷会增大数据分组传输的能耗，但考虑到 GFFS 具备防范协同攻击的能力，这样的额外开销是可以承受的。此外，若攻击者利用来自不同地理区域的妥协节点协同伪造假分组并注入网络中，则 GFFS 可以通过尽快将假分组过滤而比 SEF 节省能量，本文将在仿真实验部分对此进行验证。

4.4　存储开销

GFFS 中每个传感器节点需要存储一个密钥分区中的 k 个密钥以及 c 个节点位置。假设每个密钥和节点位置的长度分别为 64bit 和 10bit，则存储 5 个密钥和 60 个节点位置约需耗费 115byte 的存储空间。当前主流节点(如 UCB 研制的 MICA2 节点)配置 3KB 以上的 SRAM 和 128KB 以上的 ROM[2]，显然能满足需求。此外，可以采用布隆过滤器[12](Bloom filters)将节点地理位置映射成字符串，以降低节点存储开销。

4.5　参数分析

全局密钥池参数(包括全局密钥池大小 N、每个节点存储的密钥数量 k 和密钥分区数量 n)对过滤概率的影响在 SEF[3]中已详细分析，主要对参数 t(每个合法数据分组携带的 MAC 和地理位置数量)，参数 c(每个节点存储的地理位置数量)以及参数 N_c(攻击者俘获的妥协节点数量)的取值进行分析。

t 的取值是安全性和能量消耗之间的折中。数据分组携带的 MAC 和地理位置越多，攻击者需要俘获更多的节点才能成功伪造出安全机制无法过滤的假分组，从而加大了攻击难度，增强了安全性。但是，这样也增大了数据分组的长度，从而导致在传输过程中消耗更多的能量。此外，t 的大小还受限于节点允许的最大数据分组长度。例如，一些低端节点支持的最大数据分组长度仅为 36bit，因此 t 不能太大[3]。

c 的取值也是安全性和存储开销之间的折中。节点预存储的地理位置数量越多，则中间节点能够以更大的概率对假分组中的地理位置进行检测，从而增大了过滤假分组的概率。但 c / N_a 不能太大，否则攻击者在俘获少量节点后即可获取网络中所有节点的地理位置。此外，在实际应用中，c 的取值还受限于节点存储能力。例如，假设一个地理位置的长度为 10bit，存储 100 个验证密钥就需要 1KB，占用了传感器节点较大的存储空间。

攻击者俘获的妥协节点数量 N_c 越大，则攻击者可以以较大的概率获取一些相邻节点的秘密信息，从而在伪造假分组时所需捏造的假 MAC 和假地理位置数量也越少，相应地降低了假分组被中间节点识别的概率。当 N_c 足够大时，由于获取的秘密信息足够多，攻击者可以伪造出 GFFS 无法过滤的假分组。

4.6　仿真实验

为了进一步验证 GFFS 的性能，本文和 SEF[3]一样，利用 C++语言建立了模拟仿真平台。仿真实验中采用的 GFFS 数据分组大小为 77.8byte，SEF 数据分组大小为 70byte，节点发送和接收 1 个 77.8byte 数据分组的功耗分别为 6.6×10^{-3}J，1.3×10^{-3}J，发送和接收一个 70byte 数据分组的功耗分别为 6×10^{-3}J，1.2×10^{-3}J[3]。仿真环境如下：在 1 个圆形网络区域中，1 个静态源节点和 1 个静态 sink 节点分别位于圆心和圆周上，其他节点随机分布在圆形区域中。其他仿真参数的设置见表 1。限于篇幅，仅给出了在 c=0，20，40 的情况下，GFFS 和 SEF 在妥协容忍、过滤概率以及能耗方面的实验数据。取 10 次仿真实验的平均值作为实验结果。

表 1　　仿真参数的设置

参数	参数值
网络大小	$(\pi\times50\times50)m^2$
节点数量	400
源节点发包间隔	2s
源节点发包总数	100
节点通信半径	2.5m
节点感知半径	5m
全局密钥池大小 N	150
密钥分区数量	15
节点存储密钥数量	5

为有效评价相关机制的假分组过滤能力，本文提出利用单跳可过滤数据分组的累积值进行性能评价，如式(5)所示，其中，h 为传输跳数，N_h 为第 h 跳过滤的假分组个数。若机制的过滤性能越好，则假分组在网络中传输的跳数越少，相应地 f ($0\leqslant f\leqslant100$)越大。反之，f =0 则说明由于太多的节点被妥协，攻击者伪造的假分组无法被安全机制过滤。

$$f=\sum_{h=1}^{\infty}(N_h/h) \tag{5}$$

图 5 给出了 f 随妥协节点数量 N_c 和每个节点存储的地理位置数量 c 的变化情况。从图 5 中可以得出如下几点：①c 越大时，GFFS 的过滤能力越强。以 N_c=80 为例，当 c=20 时，f 仅为 37，而当 c =40 时，f 为 47。因为 c 越大时，节点存储的地理位置数量越多，中间节点对假分组中包含的假地理位置进行验证的概率也越高，故 GFFS 的过滤能力也越强。②当 c =0 或 $N_c \geqslant$130 时，GFFS 的过滤能力和 SEF 相同，而在其他情况下，GFFS 的过滤能力强于 SEF。因为当 c =0 时，中间节点不存储其他节点的地理位置，此时 GFFS 和 SEF 都无法对地理位置进行验证，因此过滤能力相同；当 $N_c \geqslant$130 时，此时攻击者俘获了足够多的秘密信息，从而可伪造出 GFFS 和 SEF 都无法过滤的假分组；当 c>0 且 N_c<130 时，SEF 仅能对假分组中的 MAC 进行验证，而 GFFS 则能同时对 MAC 和地理位置进行验证，故 GFFS 的过滤能力强于 SEF。③GFFS 能容忍的妥协节点数量为 130 个，远多于 SEF 能容忍的 12 个。这是因为 SEF 无法防范来自不同地理区域的妥协节点的协同攻击，故攻击者在俘获少量节点后即可攻破 SEF；而 GFFS 能通过地理位置验证防范来自不同地理区域的妥协节点的协同攻击，故攻击者攻破 GFFS 需要俘获大量节点。

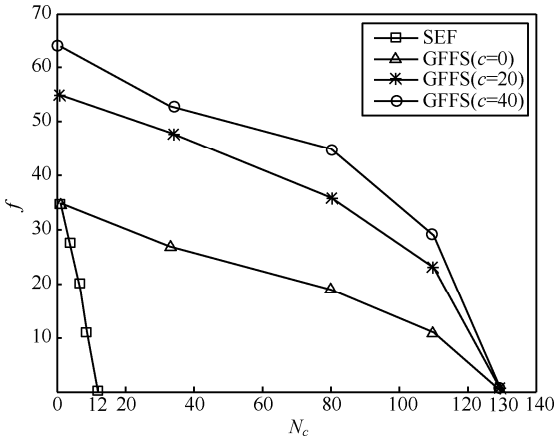

图 5 过滤概率

图 6 所示为源节点产生的 100 个假数据分组在 GFFS 和 SEF 中的传输能耗对比情况。从图 6 中可以看出，只有在 c =0，且 N_c<12 的情况下，GFFS 的能耗比 SEF 稍大，而其他时候 GFFS 的能耗都小于 SEF。这是因为当 c=0，且 N_c<12 时，GFFS 的过滤能力和 SEF 一致，而 GFFS 的数据分组比 SEF 数据分组长，因此 GFFS 能耗比 SEF 大；在其他情况下，尽管 GFFS 数据分组长度大于 SEF，但 GFFS 能通过尽早将假数据分组过滤掉而达到比 SEF 更节省能量的目的。

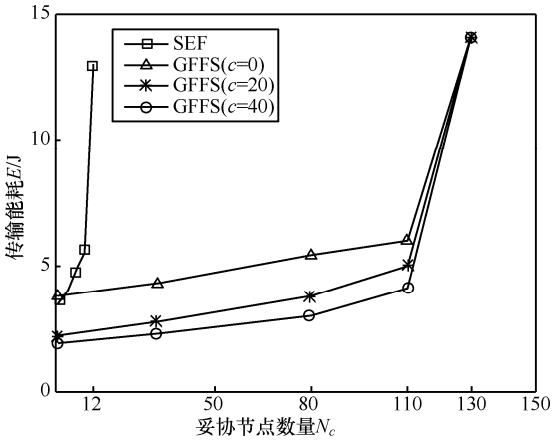

图 6 能量消耗

5 结束语

本文在深入分析当前方案无法对传感器网络中不同区域的妥协节点协同伪造的虚假数据进行检测和识别的原因后，提出一种基于地理位置的过滤方案 GFFS。将节点位置信息预分发给中间节点存储，由中间节点在转发过程中同时对数据分组中所包含的 MAC 和检测节点位置信息进行验证，从而将不同区域的妥协节点协同伪造的虚假数据过滤掉。理论分析及仿真实验表明，GFFS 可以有效地防止协同攻击，并具备较好的妥协容忍能力和较低的能量开销。然而，获取节点绝对位置信息增大了传感器节点的开销，因此，研究利用邻居节点相对位置信息防止来自不同地理区域的妥协节点的协同攻击将成为进一步的工作。

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.
REN F Y, HUANG H N, LIN C. Wireless sensor networks[J]. Journal of Software, 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.
SU Z, LIN C, FENG F J, *et al*. Key management schemes and protocols for wireless sensor networks[J]. Journal of Software, 2007,18(5): 1218-1231.

[3] YE F, LUO H, ZHANG L. Statistical en-route filtering of injected false data in sensor networks[A]. Proceedings of 23th Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM'04[C]. Hong Kong, China, 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[A]. Proceeding IEEE Symposium on Security and Privacy, S&P'04[C]. Berkley, CA, USA, 2004.259-271.

[5] YU Z, GUAN Y. A Dynamic en-route scheme for filtering false data injection in wireless sensor networks[A]. Proceedings of the 3rd International Conference on Embedded Networked Sensor Systems, SenSys'05[C]. San Diego, 2005.294-295.

[6] LI F, W J. A probabilistic voting-based filtering scheme in wireless sensor networks[A]. Proceedings of the International Wireless Communications and Mobile Computing Conference, IWCMC'06[C]. Vancouver, Canada, 2006.255-265.

[7] MA M. Resilience of sink filtering scheme in wireless sensor networks[J]. Computer Communications, 2006, 30(1):55-65.

[8] YANG H, LU S. Commutative cipher based en-route filtering in wireless sensor networks[A]. Vehicular Technology Conference, VTC'04[C]. Los Angeles, USA, 2004.1223-1227.

[9] WANG H, LI Q. PDF: a public-key based false data filtering scheme in sensor networks[A]. Proceedings of the International Conference on Wireless Algorithms, Systems and Applications, WASA'07[C]. Vaasa, Finland, 2007.129-138.

[10] REN K, LOU W, ZHANG Y. Providing location-aware end-to-end data security in wireless sensor networks[A]. Proceedings of the IEEE Conference on Computing and Communicating, INFOCOM'06[C]. Barcelona, Spain, 2006.585-598.

[11] AYDAY E, DELGOSHA F and FEKRI F. Location-aware security services for wireless sensor networks using network coding[A]. IEEE Conference on Computer Communications, INFOCOM'07[C]. Anchorage, Alaska, USA, 2007.1226-1234.

[12] BLOOM B. Space/Time trade-offs in hash coding with allowable errors[J]. Communications of the ACM, 1970,13(7):422-426.

[13] MAO G, FIDAN B, ANDERSON B. Wireless sensor network localization techniques[J]. Computer Networks: The International Journal of Computer and Telecommunications Networking, 2007.2529-2553.

[14] BOSE P, MORIN B, STOJMENOVIC I, URRUTIA J. Routing with guaranteed delivery in ad hoc wireless networks[J]. Wireless Networks, 2001, 7(6): 609-616.

[15] PENG S L, LI S S, LIAO X K, PENG Y X, XIAO N. Estimation of a population size in large-scale wireless sensor networks[J]. Journal of Computer Science and Technology, 2009,24(5):987-996.

作者简介：

刘志雄（1982-），男，湖南娄底人，中南大学博士生，主要研究方向为无线传感器网络。

王建新（1969-），男，湖南邵阳人，博士，中南大学教授、博士生导师，主要研究方向为计算机网络优化理论、计算机优化算法等。