doi:10.3969/j.issn.1000-436X.2012.10.007

安全透明的无线传感器网络数据汇聚方案

郭江鸿，马建峰
（西安电子科技大学 计算机学院，陕西 西安 710071）

摘 要：提出了一种安全透明的传感器网络数据汇聚方案，汇聚节点在不对加密数据进行解密的情况下通过散列函数与异或操作完成数据完整性检查、数据源身份认证、数据汇聚等功能，保证了数据在汇聚及传输过程中的隐私性。与相关数据汇聚方案相比，除了提供密钥安全性，所提方案可有效抵抗主动攻击、节点妥协攻击及 DoS 攻击等恶意行为，具有高的安全性；同时，方案的汇聚结果提供了数据的全局分布信息。

关键词：传感器网络；数据汇聚；数据隐私；网络安全
中图分类号：TP301 **文献标识码**：A **本文编号**：1000-436X(2012)10-0051-09

Secure and transparent data aggregation for wireless sensor networks

GUO Jiang-hong, MA Jian-feng
(School of Computer, Xidian University, Xi′an 710071, China)

Abstract:A secure and transparent data aggregation scheme for wireless sensor networks was proposed. Using hash function and XOR operation, the aggregation node completed the data integrity checking, source identity authentication and data aggregation without decrypting the encrypted data, ensured the data privacy in the process of transmitting and aggregation. Compared with related data aggregation schemes, except providing high security of encrypt key, the proposed scheme had better performances in resilient against active attack, node compromise attack and DoS attack. Also, the proposed scheme provides the information of the global data distribution.

Key words:sensor network; data aggregation; data privacy; network security

1 引言

无线传感器网络 (WSN, wireless sensor network)由大量资源传感器节点组成，彼此通过无线链路进行通信，在战场监控、灾难拯救、目标跟踪、野生动物保护等方面得到了广泛应用。无线传感器网络的一个主要功能是由节点对所处环境的某些物理参数进行测量，并将结果送往远方的服务器(或基站)进行进一步处理。由于邻近传感器可能测得相同数据，即传感器测得的原始数据中有冗余信息，数据汇聚技术成为减少数据传输量的重要手段之一。

在一些恶意环境下，敌手可能通过节点妥协、伪造数据、伪造身份等手段发动攻击以降低数据的可用性，甚至通过恶意操作向网络中注入大量虚假数据，意图发动以消耗节点计算能耗及通信能耗，进而减短传感器网络生存期的 DoS 攻击。在此情况下，数据汇聚方案不仅要保证数据本身的机密性与

收稿日期：2011-04-20；**修回日期**：2011-10-15
基金项目：国家高技术研究发展计划(“863”计划)基金资助项目（2007AA01Z429,2007AA01Z405）；国家自然科学基金重点资助项目（60633020）；国家自然科学基金资助项目（60573036, 60702059, 60503012）
Foundation Items: The National High Technology Research and Development Program of China (863 Program)(2007AA01Z429 2007AA01Z405); The Key Program of National Natural Science Foundation of China (60633020); The National Natural Science Foundation of China(60573036, 60702059, 60503012)

完整性，而且要提供抵抗 DoS 攻击的能力及数据源身份鉴别。在大部分现有的传感器网络数据汇聚方案中，数据多以明文传送或由汇聚节点对加密数据进行解密来完成数据汇聚，不能很好地保护数据的隐私性及安全性；同时，大部分现有的数据汇聚方案主要考虑汇聚效率，过滤的信息过多，不利于基站对全网数据分布情况的统计分析。

Govindan 等[1]提出的数据汇聚方法中，通过选择优化的经验路径进行数据汇聚及网内处理，有效地降低了传感器网络的传输能耗，但该方案并未考虑数据的安全性。Przydatek 等[2]提出的一种安全的消息汇聚协议，但该方案主要考虑的是数据汇聚的安全性，数据依旧用明文传输，难以保证数据的隐私性；Wagner 等[3]研究了传感器网络中对数据汇聚的攻击行为，并提出了一个评价数据汇聚方案安全强度的理论框架，但数据的隐私性并未包含在内；Cam 等[4]提出了能量有效的基于模式码安全数据汇聚协议(ESPDA)。节点建立与原始数据对应的模式码并将其发往簇头，簇头不需要对加密数据进行解密，根据模式码实现了数据汇聚。但是，该方案中每个节点都发送模式码导致能耗不够理想，同时，没有考虑到多跳转发的数据认证，可能导致主动攻击及 DoS 攻击；Acharya 等[5]提出的端到端的加密算法允许汇聚节点通过对密文的操作完成数据汇聚，保护了数据的隐私性。但该方法中指数级的计算复杂度导致过大的计算开销，且该方案在唯密文攻击下并不安全；Huang 与 Tygar 等[6]提出了安全的加密数据汇聚(SEDA, secure encrypted data aggregation)方案，该方案使用散列函数与异或操作在不对密文解密的情况下完成数据汇聚，保护了数据的隐私性且能耗较优。但该方案存在以下问题：1)要求汇聚节点预装入$(N-1)$对密钥异或值，N 为网络中节点总数，网络的可扩展性差，2)没有提供数据认证，不能抵抗主动攻击及 DoS 攻击等恶意行为，安全性差。同时，以上方案的汇聚结果并不能使基站了解各种数据在全网的分布情况，不利于整体的统计分析。

针对以上问题，本文提出了一种安全透明的传感器网络数据汇聚(STDA, secure and transparent data aggregation)方案，汇聚节点在不对密文解密的情况下通过低能耗的散列与异或运算完成数据汇聚，保护了数据的隐私性，且解决了 SEDA 方案中簇头存储开销过大的问题，提高了网络的可扩展性；通过附加 MAC 有效抵抗主动攻击、妥协攻击及 *DoS* 攻击等多种恶意行为，提供了高的安全性；同时，本文方案仅对数据进行汇聚，而具有相同数据节点的标识不被过滤，基站可通过汇聚结果了解全网的数据分布状况。

2 预备知识

2.1 ESPDA 及 SEDA 方案简介

Cam 等[4]提出的 ESPDA 方案简介如下。

1) 基站为每个节点 N_i 预装入 ID、与基站的配对密钥 k_i 以及公共密钥 k。

2) 对应每次数据收集，基站选取数据收集密钥 k_b，广播 $E_k[k_b]$，每个节点 N_i 用 k 解密消息并计算 $K=k_b \oplus k_i$ 作为本次加密密钥。

3) 对应每次数据收集，簇头选取随机种子 S，簇内广播 $E_k[S]$,节点 N_i 解密得到 S，并根据 S 计算模式码序列，序列中每个模式码对应一个取值范围；N_i 发送与自己测量数据对应的模式码到簇头。

4) 簇头对模式码进行比较及汇聚，根据汇聚结果要求部分簇内节点发送数据。

5) 节点发送<ID，t, $E_K[Data]$, $MAC(K, Data)$>到基站，t 为时戳。基站根据 ID 计算 $K=k_b \oplus k_i$ 并完成数据解密及 MAC 验证。

Huang 等[6]提出 SEDA 方案简介如下。

1) 基站为每个节点 N_i 建立与基站的配对密钥 k_i，选取单向函数 $f()$且 $f(x \oplus y)=f(x) \oplus f(y)$，$N_i$ 预装入 ID、k_i 及 $f()$。

2) 设网络中节点总数为 N，ID 分别为 N_1，N_2，…，N_N，与基站的配对密钥分别为 k_1，k_2，…，k_N，基站计算密钥序列 $L=<k_1 \oplus k_2$，$k_2 \oplus k_3$，…，$k_{N-1} \oplus k_N>$并将之预装入所有簇头节点。

3)对应每次数据收集，节点 N_i 生成不同的随机数 r_i，发送<N_i，$m_i \oplus r_i \oplus f(r_i)||k_i \oplus r_i$ >到簇头。m_i 为原始数据。

4)对于不同节点 N_i 与 N_j 的消息，簇头先通过密钥序列 L 得到 $k_i \oplus k_j$，进而得到 $r_i \oplus r_j$，再计算如下：$V=m_i \oplus r_i \oplus f(r_i) \oplus m_j \oplus r_j \oplus f(r_j) \oplus (r_i \oplus r_j) \oplus f(r_i \oplus r_j) = m_i \oplus m_j$ $V=0$，则 $m_i=m_j$，簇头保留其中一个，另一个被视为冗余数据丢弃。数据汇聚完成后，簇头将汇聚结果发往上游汇聚节点，直到基站接收到最终汇聚结果。

5) 基站使用与 N_i 的配对密钥 k_i 解密 $k_i \oplus r_i$，再利用 r_i 解密消息得到 m_i。

2.2　网络模型

STDA 方案采用与 CAM 及 Huang 方案相同的网络结构——分簇传感器网络，并作如下假设：

1) 节点间可通过合适的密钥协议建立配对密钥;

2) 簇头具有较高的安全性;

3) 已建立相应数据汇聚树，基站为根节点。

因本文主要关注数据汇聚，且现已有诸多文献对建立配对密钥、构建汇聚树等问题进行研究并取得了系列成果，所以做出上述假设而不对其进行具体讨论。

为便于分析，本文采用一个理想的分簇传感器网络模型，如图 1 所示。

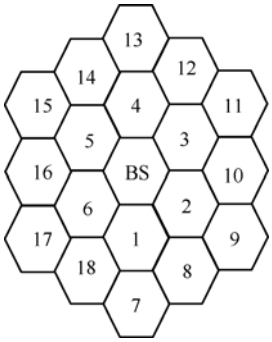


图 1　理想的分簇传感器网络模型

图 1 中，BS 为基站，网络共分为 19 个簇，每个簇用一个正六边形表示，设正六边形边长为 40m，簇头位于正六边形中央，BS 为中央簇头。传感器节点通信半径为 40m，每个簇内平均有 n 个节点。进行数据汇聚所使用的汇聚树如图 2 所示。

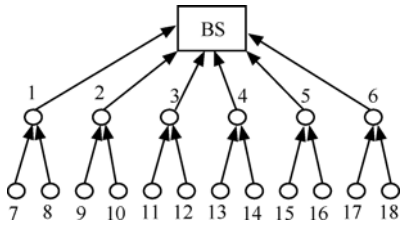


图 2　数据汇聚树

2.3　攻击者模型

本文假设攻击者模型为 Dolev-Yao 模型，攻击者可以控制整个通信网络，除了可以窃听、截获所有经过网络的消息外，还具备以下知识和能力。

1) 熟悉加解、解密、散列(hash)等密码运算，拥有自己的加密密钥和解密密钥。

2) 熟悉网络中各节点的 ID。

3) 具有密码分析的知识和能力。

4) 可以发动以下攻击：

① 由于临近的传感器节点可能得到相同的数据且敌手可以得到节点发送的密文信息，因此假设敌手可发动已知明文/密文攻击；

② 一般说来，传感器节点妥协难以避免，敌手可发起妥协攻击在物理上俘虏节点，获取其秘密信息；

③ 敌手可以重放以前的合法消息或假冒身份向汇聚节点发送虚假消息发动主动攻击；

④ 敌手可以向网络中注入大量虚假数据发起 DoS 攻击。

3　STDA 方案简介

STDA 方案主要由系统初始化、消息加密、数据汇聚、基站解密等部分组成。

3.1　系统初始化

设网络共有 N 个节点，ID 分别为：N_1，N_2，…，N_N，$|ID|$=2byte，部署前，基站(等同可信第三方)选取 N 个 l bit 的随机密钥 S_1，S_2，…，S_N，l 为安全参数。同时选取 2 个单向函数 $f(\,)$，$g(\,)$，具有下列性质

$$\begin{cases} f(x \oplus y) = f(x) \oplus f(y) \\ g(x \oplus y) \neq g(x) \oplus g(y) \end{cases} \tag{1}$$

对任意的 $i \in [1, N]$，服务器将 S_i，N_i，$f()$，$g()$，c 以及密钥材料预装入节点 N_i。其中，S_i 为节点 N_i 与基站的配对密钥，c 为序号，初始为 1。

3.2　消息加密

传感器部署后，节点通过分簇算法(如 ACE 算法[7]等)建立分簇网络，并完成节点间的配对密钥建立。当收到基站数据请求或到达周期性数据采集时间后，N_i 采集数据并构造消息(I)：

$$< N_i \| c \| (m_i \oplus f^l(r_i)) \| (s_i \oplus r_i) \| f^l(f^l(r_i)) \| MAC > \quad \text{(I)}$$

其中，N_i 为节点 ID；m_i 为 N_i 采集的数据；r_i 为 N_i 生成的随机数(每次使用不同的随机数掩盖数据)；c 为序号；$f^l()$表示取 $f()$的前 l bit。MAC 为消息认证码，计算如下

$$MAC = g^l(k_i \| c \| (m_i \oplus f^l(r_i)) \| (s_i \oplus r_i) \| f^l(f^l(r_i))) \tag{2}$$

其中，k_i 为节点 N_i 与簇头 A_i 的配对密钥，$g^l()$表示取 $g()$的前 l bit。N_i 发送消息(I)到簇头 A_i，同时更新 c 为 c+1。

3.3 数据汇聚

A_i 维持一张用于数据汇聚的邻居信息表，如表 1 所示。

表 1 邻居信息

邻居 ID	配对密钥	序号
N_1	k_1	c’
N_2	k_2	c’
N_3	k_3	c’
…	…	…

表 1 中的 c’初始为 0，N_i 为簇内节点。

A_i 接收到 N_i 的消息后，进行如下验证。

1) 新鲜性验证：通过序号 c 检查消息新鲜性，若消息中的序号与 A_i 自身的序号一致则转 2)，否则丢弃。

2) 检查邻居信息表中对应的 c’，若 c−c’=1 则通过验证并转 3)，否则丢弃。

3) 数据源身份认证及消息完整性验证：A_i 按照式(2)计算 MAC，与消息中 MAC 一致则接受该消息并更新邻居信息表中与 N_i 对应的 c’为 c’+1；因敌手在未捕获节点的情况下，无法获取节点与簇头的配对密钥，因此 MAC 验证可同时进行数据源身份认证和消息完整性验证。

4) 数据汇聚：对于来自 N_i 及 N_j 的消息 M_i、M_j，簇头计算如下

$$\begin{aligned} &f^l(m_i\oplus f^l(r_i)\oplus m_j\oplus f^l(r_j))\oplus f^l(f^l(r_i))\oplus f^l(f^l(r_j))\\ &=f^l(m_i\oplus m_j)\\ &=V \end{aligned} \tag{3}$$

显然，若 V≠f^l(0)，则 $m_i\neq m_j$，M_i 及 M_j 都被保留；若 V=f^l(0)，则 $m_i=m_j$ 的概率为 $1-1/2^l$，当 l 足够大时，可认为 $m_i=m_j$，A_i 保留 M_i 或 M_j 中的一个，并暂存消息 M_i 如格式(II)。

$$M_i =< N_i \| (m_i \oplus f^l(r_i)) \| (s_i \oplus r_i) \| f^l(f^l(r_i)), \quad IDList > \tag{II}$$

其中，*IDList* 为与 N_i 具有相同数据的节点 ID 列表。A_i 在不对数据解密的情况下完成了数据汇聚，保证了数据的隐私性。

设 M_{Ai} 为 A_i 对 A_i 所在簇的簇内数据及下游汇聚节点所发送汇聚数据的汇聚结果，K_{Ai} 为 A_i 与上游汇聚节点 A_j 的配对密钥，则 A_i 构造消息(III)发往上游汇聚节点 A_j 后更新自身的序号 c 为 c+1。

$$< A_i \| M_{Ai} \| c \| g^l(K_{Ai} \| M_{Ai} \| c) > \tag{III}$$

A_j 收到消息(III)后，先通过 c 及 MAC 验证检查发送方身份、消息完整性及消息新鲜性，然后与其他下游汇聚节点的汇聚结果及所在簇的簇内数据进行比较，得到进一步的汇聚结果并构造消息如格式(III)，发往 A_j 的上游汇聚节点 A_k。此过程一直持续到最上游的汇聚节点将结果发送到基站。

3.4 基站解密

基站收到汇聚结果后，先进行相应的 MAC 验证，然后对来自 N_i 的数据，基站用与 N_i 的配对密钥 S_i 计算 m_i：

$$(m_i\oplus f^l(r_i))\oplus f^l(s_i\oplus r_i\oplus s_i)=m_i\oplus f^l(r_i)\oplus f^l(r_i)=m_i \tag{4}$$

完成数据解密后，基站可对数据进行进一步分析与处理。

4 安全性分析

安全性分析主要针对 2.3 节定义的攻击者模型下本文方案的安全性分析，并与 ESPDA 及 SEDA 方案作比较。

4.1 抗明文/密文攻击分析

本文以明文/密文攻击下各方案的密钥安全性衡量其抗明文/密文攻击的能力。

1)STDA 方案中共有 3 类密钥：N_i 与基站的配对密钥 S_i；N_i 与 A_i 的配对密钥 k_i；N_i 进行数据加密的密钥 r_i。各密钥在明文/密文攻击下的安全性如下：

k_i 只存在于 MAC 中，敌手在未捕获 N_i 时，意图通过 MAC 获得 k_i 的难度等同于单向函数求逆，因此 k_i 在明文/密文攻击下是安全的。

r_i 为选取的随机数，每次数据加密使用不同的随机数生成 $f(r_i)$。敌手从密文中获取的与 r_i 有关的信息有：$m_i\oplus f(r_i)$、$f(f(r_i))$和 $s_i\oplus r_i$。

敌手发动密文攻击时，从 $f(f(r_i))$中获取 $f(r_i)$及从 $f(r_i)$中获取 r_i 的难度都等同于单向函数求逆；敌手在未知 s_i 的情况下企图从 $s_i\oplus r_i$ 中成功猜测 r_i 的概率 $p=2^{-l}$。当安全参数 l 足够大时，p 是一个可以忽略的量；敌手发起明文攻击时，可从 $m_i\oplus f(r_i)$中获取 $f(r_i)$，但从 $f(r_i)$中获取 r_i 的难度等同于单向函数求逆；由于 N_i 每次使用不同的随机数，当前获取的 $f(r_i)$无助于敌手获取下次掩盖数据的 $f(r_i$’$)$；在未知 k_i 的情况下，即使 $f(r_i)$暴露，敌手构造正确的 MAC 依然是困难的。因此，$f(r_i)$暴露不会对密钥安全性造成实质性危害。r_i 在明文/密文攻击下是安全的。

S_i 是 N_i 与基站的配对密钥，$S_i\oplus r_i$ 可看作一个

一次一密方案，其中，每个明文分组都是 S_i，每次使用不同的随机密钥加密。对于一次一密方案，不论明文有何统计规律，都是信息论安全的。因此在敌手未捕获 N_i 时，无法从消息中获取 S_i，S_i 在明文/密文攻击下是安全的。

因此，STDA 方案是明文/密文攻击安全的。

2) ESPDA 方案中，每个节点 N_i 预装入与基站的共享密钥 *Ki* 及全网统一的用于解密基站广播及簇头广播的密钥 *k*。每次数据收集前，基站广播用 *k* 加密本次数据收集的密钥，簇头在簇内广播用 *k* 加密的用于生产模式码的随机种子，加密算法采用 Blowfish[8]。在节点安全的条件下，密钥在明文/密文攻击下是安全的。

3) SEDA 方案中基站为每个节点分配与基站共享的密钥 k_i，节点使用随机数 r_i 进行数据掩盖，该方案在随机预言机模型下证明了 k_i 与 r_i 是 IND-CPA 及 IND-CCA 安全的。

4.2　抗妥协攻击分析

一般来说，传感器网络中节点妥协难以避免，因此设网络中有多个传感器节点妥协，敌手可获得妥协节点的所有秘密信息。本文以未妥协节点的密钥安全性衡量各方案的抗妥协攻击能力。

1) 对于 STDA 方案，各节点与基站的配对密钥是由基站生成的随机数，因此从被妥协节点无法获取未妥协节点与基站的密钥；节点与簇头的配对密钥的抗妥协攻击能力依赖于所用的配对密钥建立方案；r_i 是节点用于掩盖数据的随机数，各个节点生成的随机数是相互独立的，显然，不论敌手捕获多少节点，均无法得到未妥协节点的 r_i。

即使簇头 A_i 被妥协，敌手可得到的信息仅有簇头与基站及簇头与簇内节点的配对密钥。对于任一未妥协节点 N_i，若 N_i 以 A_i 为簇头，则 S_i 及 r_i 是安全的且 A_i 被妥协并不暴露 N_i 与其他邻居节点的配对密钥；若 N_i 所在簇的簇头安全，则 S_i、k_i 以及 r_i 都是安全的。

2) 对于 ESPDA 方案，每个节点预装入 2 个密钥：与基站的配对密钥 *Ki* 和全网共享密钥 *k*，因每个节点的 *Ki* 由基站随机生成，敌手无法从妥协节点获得未妥协节点与基站的配对密钥；但 *k* 为全网共享密钥，单点妥协就会暴露 *k*。敌手可利用 *k* 假冒基站或簇头发布虚假消息，对安全通信构成极大威胁。

3) 对于 SEDA 方案，在簇头安全的前提下，敌手无法从妥协节点获取未妥协节点与基站的配对密钥及掩盖数据的随机数。但只要一个簇头妥协，敌手只需捕获任一非簇头节点，就可从簇头的密钥序列 *L* 中获得所有节点与基站的配对密钥，对网络的安全通信造成极大危害。

4.3　抗主动攻击分析

1) 对于 STDA 方案，簇头每次将汇聚数据发往上游汇聚节点后，将更新自身的序号，通过简单的序号比较可过滤重放消息；敌手假冒节点 N_i 向簇头发送包含正确序号的虚假消息，在不知道 N_i 与簇头配对密钥的情况下，敌手可以伪造正确 MAC 的概率是一个可忽略的量，簇头通过散列计算可将该消息过滤；即使敌手妥协了部分节点，可伪造多个含有正确的 *c* 及 MAC 的虚假消息。当任一虚假消息被簇头 A_i 接受后，A_i 将更新邻居信息表中与 N_i 对应的序号 *c*’为 *c*，后续序号为 *c* 的虚假消息无法通过数据汇聚的第 2 步验证而被丢弃，即使敌手使用序号 *c*+1,也因无法通过数据汇聚的第 1 步验证而被丢弃。因此，在一次数据收集中，即使 N_i 被妥协并发送多个虚假消息，也只能有一个虚假消息被簇头接受，其他虚假消息将由簇头通过序号比较或散列计算过滤，STDA 方案可以以较低的能耗有效抵抗敌手发动的主动攻击。

2) 对于 ESPDA 方案，Cam 等仅指出在节点向基站传送加密数据时使用 MAC，未说明基站广播 $E_k[k_b]$及簇内广播 $E_k[S]$是否附加 MAC。若广播中没有 MAC，即使网络中没有节点妥协，节点解密敌手可以利用发布的虚假消息得到 k_b’或 *S*’，但并不知道该消息是否来自基站或簇头。如抗妥协攻击中所述，传感器网络中节点妥协难以避免，这意味着对于 ESPDA 方案，全网共享密钥 *k* 的暴露难以避免，敌手可利用 *k* 发布虚假消息干扰基站对汇聚结果的判断。ESPDA 方案抗主动攻击能力差主要源自广播消息缺乏 MAC 及全网共享密钥在妥协攻击下的脆弱性。

3) 对于 SEDA 方案，节点向簇头发送的消息中既没有新鲜性参数，不能抵抗重放攻击；也没有对消息附加 MAC，不能抵抗主动攻击。例如敌手假冒 N_i 发送消息<N_i，$A||B$>，*A*、*B* 为敌手伪造的与合法消息对应部分等长的随机数据，则 *A*、*B* 可以表示为

$$B= k_i \oplus r_i', \quad A=m_i' \oplus r_i' \oplus f(r_i')$$

不论簇头或基站均无法判断这个消息的发送方身份是敌手还是 N_i。SEDA 方案在主动攻击下安全性差。

4.4　抗 DoS 攻击分析

敌手假冒身份发动 DoS 攻击，意图大幅度增加节点的计算能耗或通信能耗，减少传感器网络生命周期。

1) 对于本文方案，如抗主动攻击部分所述，在 N_i 安全情况下，敌手假冒 N_i 发送的消息无法通过簇头的 MAC 认证(散列计算)而被过滤，不会带来大的通信开销及计算开销；即使 N_i 被妥协并发送多个消息，在一次数据收集过程中只有一个消息被簇头接受，其他消息将被簇头通过序号比较或 MAC 认证过滤掉，不会带来大的通信开销，STDA 方案可有效抵抗 DoS 攻击。

2) 对于 ESPDA 方案及 SEDA 方案，两者都不能很好地抵抗主动攻击，敌手可伪造基站广播或簇头广播使网内节点收集并发送数据，极大地增加节点的数据传输量，降低传感器网络的生命期，ESPDA 及 SEDA 方案的抗 DoS 攻击能力较差。

4.5　确定安全参数

由密钥安全性分析可知，敌手没有对节点 N_i 进行物理捕获的情况下，获取 S_i、k_i 及 r_i 的概率等同于对密钥的随机猜测。敌手测试一个密钥最少需经过一次散列求值，设 $f()$ 与 $g()$ 的运算量等同 SHA−1，SHA−1 运算一次约需在信息单元间进行约 1 740 个基本运算，即使配置专用的 FPGA 芯片，仍需约 80 个时钟周期。设安全参数 l 为密钥长度(bit)，敌手使用 m 台 4GHz 且配有 FPGA 的 PC 对密钥空间穷举搜索所需时间 T 约为

$$T = 2^l \times 80/(m \times 4 \times 10^9 \times 86\,400 \times 365) \qquad (5)$$

则穷举 64bit 密钥空间所需时间 T 如图 3 所示。

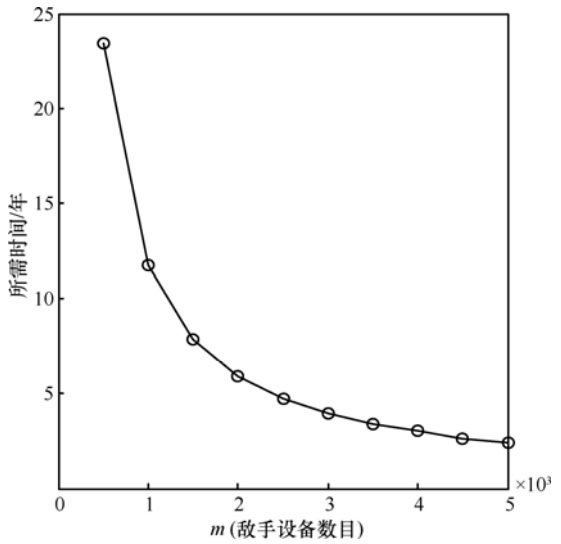


图 3　穷举 64bit 密钥空间所需时间

由图 3 可知，即使敌手使用 5 000 台 4GHz 且配有 FPGA 的 PC，仍需 2.34 年完成对 64bit 密钥空间的穷举。敌手增加用于攻击的设备数目可缩短需要的穷举时间，但将大幅度增加攻击成本。一般的传感器(如 MICA2)使用电池供电，其生命期约为 1~2 年，因此本文选择安全参数为 64bit。

综上所述，3 种方案的安全性如表 2 所示。

表 2　　安全性比较

方案	抗明文/密文攻击	抗妥协攻击	抗主动攻击	抗 DoS 攻击
STDA	是	是	是	是
ESPDA	是	部分	否	否
SEDA	是	取决于簇头	否	否

5　开销分析

本小节分析 STDA 方案的存储开销、计算开销及通信开销，并与 ESPDA 方案及 SEDA 方案作比较。

1) 存储开销分析

设节点 ID 长度为 2byte，密钥长度为 8byte，序号为 2byte，单向函数存储开销为 m_1，密钥材料存储开销为 m_2。

在本文方案中，每个节点需预装入 ID、与基站的配对密钥、初始序号、密钥材料及单向函数 $f()$、$g()$。

在 ESPDA 方案中，每个节点需预装入 ID、与基站的配对密钥、全网共享密钥。设 ESPDA 方案使用单向函数 $h()$进行 MAC 计算，则各节点还需装入 $h()$。

在 SEDA 方案中，每个节点需预装入 ID、与基站的配对密钥及单向函数 $f()$，另外，每个簇头需预装密钥序列 L，L 大小等同于$(N-1)$个密钥所需的空间，N 为网络中节点总数。显然，当网络规模较大时，密钥序列 L 所需的存储开销是传感器节点无法承受的。

3 种方案的存储开销如表 3 所示。

表 3　　存储开销比较

方案	存储开销
STDA	$(2m_1+m_2+12)$byte
ESPDA	(m_1+18)byte
SEDA	$(m_1+10$ (簇内节点)) byte；$(m_1+N\times 8+2$(簇头)) byte

要指出的是建立邻居节点间的配对密钥是节点间通信安全的基础，虽然在 ESPDA 及 SEDA 方案中并未明确指出节点需预装建立配对密钥所需的密钥材料，但这部分内容在传感器网络安全通信

协议中是不可或缺的，特别在多跳转发时，需依靠邻居节点间的配对密钥来提供单跳通信的安全。所以，本文方案的存储开销是传感器节点可以接受的。

2) 计算开销分析

ESPDA 方案对数据进行了分区，每个数据区间对应一个模式码，显然，ESPDA 方案通过降低数据精度来增加冗余数据，从而达到提高汇聚效率的目地。不同的数据精度要求及各节点数据达到簇头的顺序将影响簇头在汇聚过程中对数据的比较次数。为公平起见，设 3 种方案中要求的数据精度相同且各节点的数据到达簇头的顺序相同，即一次簇内数据汇聚过程中，簇头进行数据比较的次数相同，为 h 次。设簇内平均节点数为 n，本文以一次数据汇聚中簇内普通节点及簇头的计算复杂度衡量各方案的计算开销。3 种方案的计算复杂度如表 4 所示。

表 4　　计算复杂度

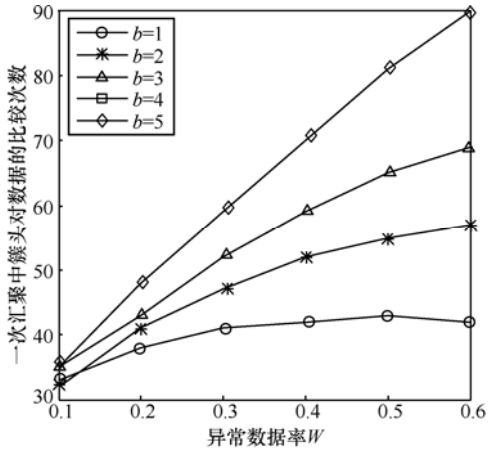
方案	ENC		DEC		HASH		⊕	
	A_i	N_i	A_i	N_i	A_i	N_i	A_i	N_i
STDA	0	0	0	0	$h+n$	3	$3\times h$	3
ESPDA	1	1	1	2	d	d+1	h	1
SEDA	0	0	0	0	h	1	$(3+q)\times h$	3

表中 ESPDA 方案所需的 d 次散列计算用来产生与 d 个数据区间对应的模式码，若节点需进行数据发送则需多一次散列计算得到 MAC；SEDA 方案中的 q 为根据密钥序列计算 $k_i \oplus k_j$ 所需的平均异或操作次数。

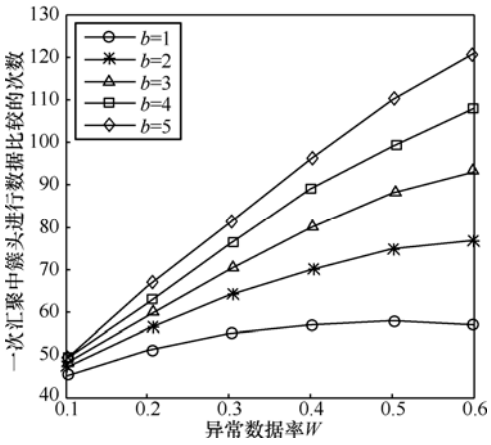
设测量数据范围为 5~94；ESPDA 方案将数据区间划分为[5,14]，[15,24]，…，[85,94]并生成 9 个对应的模式码；SEDA 方案及 STDA 方案可通过对测量结果进行简单的四舍五入得到与 ESPDA 方案相同的数据量。设[15,24]为正常数据区间，其他区间为异常数据区间；异常数据率 W 为测量结果不在正常区间的节点数所占比例；异常数据在 b 个异常数据区间内随机均匀分布；则各方案中簇头进行簇内数据汇聚所需的比较次数 h 如图 4 所示，簇头的计算能耗如图 5 所示。

显然，增加 b、W 或簇内节点数，都增大了簇内的异常数据量，从而增加了簇头的数据比较次数。

设 3 种方案中所采用的单向函数或 MAC 计算与 SHA-1 具有相同的计算开销，原始数据长度 8byte。据 Gura 等[9]的研究，MICA2 传感器上(8bit，ATmega128L，8MHz，电压 3V，活动电流 8mAh)[10]，SHA-1 的能耗约 5.9μJ×L，L 为输入长度(byte)。据 Cam 等[4]的研究，在 SmartDust (8bit，4MHz)上 Blowfish 输出 32byte 密文耗时约为 2 444ms，则估算在 MICA2 上平均输出 1byte 密文耗时约为：2 444/(32×2)=38ms，能耗约为：0.038s×3V×8mAh =0.9mJ。在忽略异或操作能耗的情况下，一次数据汇聚中 N_i 的计算能耗如表 5 所示。



(a) n=30 时簇头比较次数



(b) n=40 时簇头比较次数

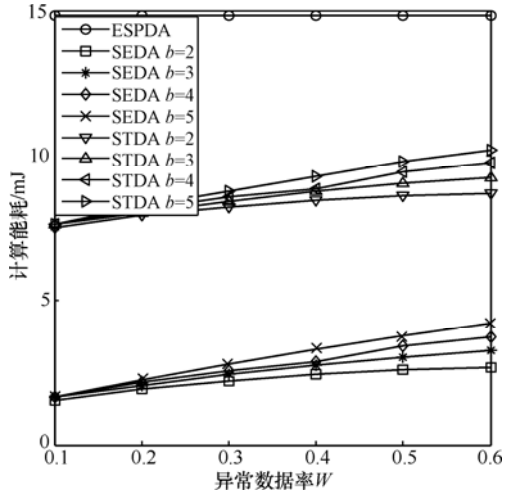
图 4　一次汇聚中簇头进行的簇内数据比较次数

表 5　　N_i 的计算能耗

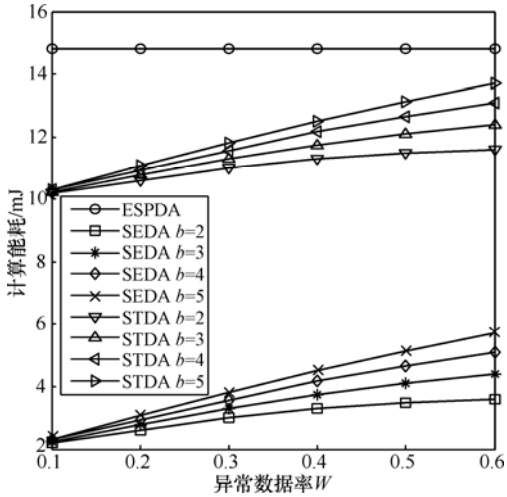
方案	计算能耗
SEDA	5.9×8/1 000=0.047mJ
ESPDA	3×0.9×8+0.047×9+0.047×2=22.12mJ
STDA	0.047×2+0.047×34/8=0.29mJ

其中，ESPDA 方案的 3 部分能耗分别为：2 个 Blowfish 解密及 1 个 Blowfish 加密、生成模式码的 9 个散列计算、MAC 计算；STDA 方案的 2 部分能耗分别为：生成 $f(r_i)$、$f(f(r_i))$的 2 个散列计算、对 34byte 数据(如 3.2 节中式(2)所示)计算 MAC。

从图 5 及表 4 可知，SEDA 方案中簇头在一次汇聚中仅需 h 次散列运算，能耗较优；STDA 方案比 SEDA 方案多出 h 次散列及 n 次 MAC 计算，簇头的计算能耗高于 SEDA 方案，但散列及 MAC 的计算能耗相对较低，STDA 方案的计算开销是传感器节点可接受的；ESPDA 方案的计算开销主要来自 Blowfish 算法的加解密运算。显然，随着簇内节点数、异常数据率以及异常区间数目的增加，SEDA 方案及 STDA 方案的计算开销将达到并超过 ESPDA 方案，但在一般情况下(如文中所设簇内节点 40、异常区间小于 6、异常数据率不大于 0.6)，三者的计算开销总体上属于同一层次(单位为 mj)。如在 n=30，b=3，异常数据率为 0.3 时，ESPDA、STDA 及 SEDA 这 3 种方案中簇头的计算开销分别为 14.82mj、8.4mj 及 2.42mj。STDA 方案的综合计算能耗高于 SEDA 方案，但低于 ESPDA 方案。



(a) n=30 汇聚能耗



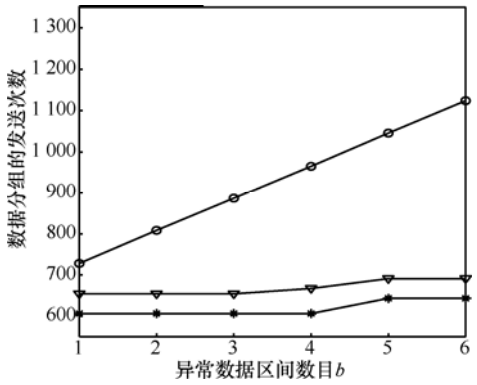
(b) n=40 汇聚能耗

图 5　簇头进行一次数据汇聚的能耗

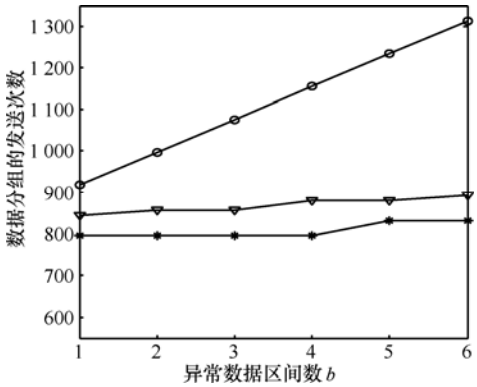
3) 通信开销分析

3 种方案都使用分簇传感器网络，其中，ESPDA 方案只有一层汇聚节点，异常数据经多跳转发单播到基站；SEDA 及 STDA 方案为多层汇聚方案，汇聚结果通过汇聚树到基站。本文在 2.2 节的网络模型(图 1)中比较各方案的通信开销，对应的汇聚树如图 2 所示。

本文使用 NS2 对 3 种方案在 2.2 节中网络环境下的通信开销进行模拟。SEDA 与 STDA 方案使用图 5 的汇聚树将汇聚结果传送到基站，父亲节点与孩子节点通过网关节点 2 hop 通信。ESPDA 方案未说明使用何种路由协议将消息单播到基站，文中以图 5 的数据汇聚树作为 ESPDA 方案的路由方案。设 3 种方案均采用 802.15.4 标准对消息进行封装，该标准允许最多 102 byte 的可变载荷，总长度最大为 128byte。对于 3 种方案，一次数据汇聚中总的数据分组发送次数如图 6 所示。



(a) n=30 数据传输量

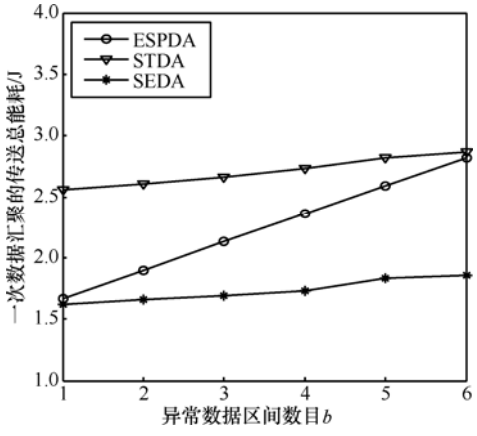


(b) n=40 数据传输量

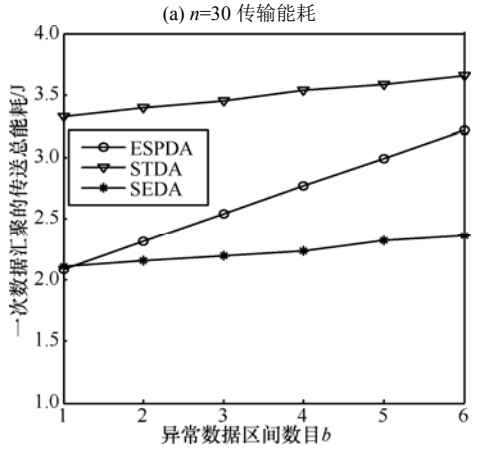
图 6　一次数据汇聚中的数据分组发送次数

图 6 中，ESPDA 方案的数据分组传输量最大，因为 ESPDA 方案只采用了一层汇聚，由簇头指定的节点通过单播将数据发到基站。STDA 方案的数据分组传输量略大于 SEDA 方案，原因有 2 方面，

首先，簇头发送的汇聚结果中包含了与数据对应的节点 ID 列表，当该列表较大时，需要多个数据分组才能完成汇聚结果的发送；其次，STDA 方案的消息长度大于 SEDA 方案的消息长度，对于相同数量的数据，STDA 方案中的汇聚节点可能需要更多的数据分组进行封装。SEDA 方案的数据分组传输量最小。据 Gura 等[9]的研究，MICA2 节点发送与接收一个字节的能耗分别为 59.2μJ、28.6μJ。各方案中一次数据汇聚所需的发送能耗如图 7 所示。



(a) n=30 传输能耗



(b) n=40 传输能耗

图 7　传输能耗比较

由图 7 可知，SEDA 方案的传输能耗较小，ESPDA 方案次之，STDA 方案的传输能耗最大。如在 n=30，b=5 时，SEDA、ESPDA、STDA 方案进行一次数据汇聚的综合传输能耗分别为 1.83J、2.6J 及 2.82J，SEDA 及 ESPDA 方案的传输能耗约为 STDA 方案的 64.9%、92.2%。ESPDA 方案的传输能耗大于 SEDA 方案的主要原因在于 ESPDA 方案采用一层汇聚导致数据分组发送量较多；STDA 方案的传输能耗大的主要原因在于 2 方面：1)STDA 方案的消息长度大于 SEDA 方案及 ESPDA 方案的消息长度，导致较大的传输开销。但附加 MAC 提供了更高的安全性，且 STDA 方案解决了 SEDA 方案中簇头存储开销过大的问题，提供了更好的可扩展性；2)簇头仅对数据进行汇聚，而保留了具有相同数据的 ID，从而增加了消息长度及数据分组发送量。但保留 ID 列表解决了传统的数据汇聚方案中数据丢失问题，有利于基站了解全网的数据分布。

6　结束语

数据汇聚是减少传感器网络冗余数据传输的重要技术手段。由于无线链路的开放性，恶意环境下无线传感器网络的安全性显得尤为重要，如在战场监控中，战场信息需要加密传输。由于敌手可以捕获节点并注入需要的代码使这些节点完成敌手需要的操作，因此加密信息在到达远端服务器前应尽量避免解密操作以提供高安全性，即需要充分考虑数据隐私保护。数据汇聚方案应综合考虑汇聚效率、数据的安全性、隐私性等多方面因素。

本文提出的无线传感器网络数据汇聚方案在不对加密数据解密的情况下完成数据汇聚，且比现有的提供数据隐私保护的数据汇聚方案具有更高的安全性；同时，本文方案的数据汇聚结果可为基站提供各个数据在全网的分布情况，更有利于恶意环境下的信息收集。

为提供高的安全性与全网数据分布信息，本文方案中的消息中附加了较多内容（如 MAC 与节点 ID 列表），消息长度大于相关方案，因此通信能耗略高。如何在提高安全性的基础上尽量减少通信能耗将是下一步的工作目标。

参考文献：

[1] INTANAGONWIWAT C, GOVINDAN R, ESTRIN D, *et al*. Directed diffusion for wireless sensor networking[J]. IEEE/ ACM Transactions on Networking, 2003, 11(1):2-16.

[2] BARTOSZ P, DAWN S, ADRIAN P. SIA:secure information aggregation in sensor networks[A]. Proceedings of ACM SenSys Conference[C]. Los Angeles, USA, 2003. 255-265.

[3] WAGNER D. Resilient aggregation in sensor networks[A]. Proceedings of the 2nd ACM Workshop on Security of Ad Hoc and Sensor Networks[C]. Washington, DC, USA, 2004.78-87.

[4] CAM H, OZDEMIR S. Energy-efficient security protocol for wireless sensor networks[A]. Proceedings of IEEE VTC Fall 2003 Conference[C]. New York, USA, 2003. 2981-2984.

（下转第 70 页）

LI X Y, GUI X L. Cognitive model of dynamic trust forecasting[J]. Journal of Software, 2010, 21(1):163-176.

[14] FÉLIX G M, GREGORIO M P. Security threats scenarios in trust and reputation models for distributed systems[J]. Computers and Security, 2009, 28(7): 545-556.

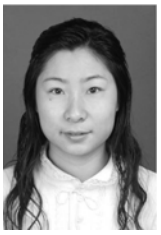
[15] 鲍宇,曾国荪,曾连荪.P2P 网络中防止欺骗行为的一种信任度计算方法[J]. 通信学报,2008, 29(10):215-222.
BAO Y, ZENG G S, ZENG L S. Reputation computation based on new metric in P2P network[J]. Journal on Communications, 2008, 29(10): 215-222.

[16] 苗光胜,冯登国,苏璞睿.P2P 信任模型中基于行为相似度的共谋团体识别模型[J].通信学报,2009, 30(8):9-20.
MIAO G S, FENG D G, SU P R. Colluding clique detector based on activity similarity in P2P trust model[J]. Journal on Communications, 2009, 30(8):9-20.

作者简介：

李峰（1978-），男，山东德州人，东北大学讲师，主要研究方向为信任管理和信任建模技术。

申利民（1962-），男，黑龙江佳木斯人，博士，燕山大学教授、博士生导师，主要研究方向为软件工程和可信计算。



司亚利（1981-），女，黑龙江齐齐哈尔人，燕山大学讲师，主要研究方向为物联网与信息安全。

牛景春（1977-），男，河北秦皇岛人，燕山大学博士生，主要研究方向为信任管理和信任建模技术。

（上接第 59 页）

[5] ACHARYA M, GIRAO J. Secure comparison of encrypted data in wireless sensor networks[A]. 3rd International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks[C]. Trentino, Italy, 2005.47-53.

[6] HUANG S I, SHIUHPYNG S, TYGAR J D. Secure encrypted-data aggregation for wireless sensor network[J]. Springer Wireless Networks, 2010, 5(16):915-927.

[7] CHAN H, PERRIG A. ACE: an emergent algorithm for highly uniform cluster formation[J]. LNCS, 2004,2(2920):154-171.

[8] SCHNEIER B. Fast software encryption[A]. Cambridge Security Workshop Proceedings[C]. Springer-Verlag, 1994.191-204.

[9] WANDER A, GURA N, EBERLE H, *et al*. Energy analysis of public-key cryptography on small wireless devices[A]. Proc of PerCom'05[C]. Kauai Island, Hawaii, USA, 2005.324-328.

[10] MICA. datasheet[EB/OL]. http://www.xbow.com/Pro-ducts/ Product_pdf_files/Wireless_pdf/MICA2_Datasheet.pdf/, 2006.

作者简介：

郭江鸿（1975-），男，山西长治人，西安电子科技大学博士生，主要研究方向为无线移动安全、网络安全。

马建峰（1963-），男，陕西西安人，博士，西安电子科技大学教授、博士生导师，主要研究方向为计算机安全、密码学、移动与无线网络安全。