基于单向函数的无线传感器网络安全数据汇聚方案*

郭江鸿[1]，刘志宏[2]，巫喜红[1]

（1. 广东嘉应学院 计算机系，广东 梅州 514015；2. 西安电子科技大学 计算机学院，西安 710071）

摘　要：为有效减少冗余数据传输，增强安全性，提出了一种基于单向函数的传感器网络安全数据汇聚方案，汇聚节点无须解密数据，利用单向函数完成数据完整性检查、身份认证、数据汇聚等，保证了端到端传输的数据私密性。与相关数据汇聚方案相比，除了具有较小的能耗开销，该方案还在抵抗主动攻击、节点妥协攻击及 DoS 攻击等恶意行为方面具有较高的安全性。

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　　**文献标志码**：A　　**文章编号**：1001-3695(2012)08-3067-04

doi:10.3969/j.issn.1001-3695.2012.08.069

Secure data aggregation scheme for wireless sensor networks based on one-way function

GUO Jiang-hong[1], LIU Zhi-hong[2], WU Xi-hong[1]

(1. *Dept. of Computer Science & Technology, Jiaying University, Meizhou Guangdong* 514015, *China*; 2. *Dept. of Computer Science & Technology, Xidian University, Xi' an* 710071, *China*)

Abstract: For reducing redundant transmissions and enhancing security, this paper proposed a secure data aggregating method for wireless sensor networks based on one-way function. The aggregators completed the data integrity checking, source identity authentication and data aggregation without decrypting the cipher-text and ensured the data privacy for end-to-end transmission. Compared with related scheme, besides lower energy consumption, the proposed scheme has higher security in terms of resilient against active attack, node compromise attack and DoS attack.

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

随着微电子技术的发展，无线传感器网络（wireless sensor network, WSN）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。WSN 由大量资源有限的传感器节点组成，通过无线链路进行通信，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。由于传感器节点一般由电池供电且大规模更换节点电池较为困难，因此能耗成为设计传感网各种协议的主要考虑因素之一。数据汇聚技术则是减少冗余数据传输，降低通信能耗的重要手段。同时，由于无线链路的开放性，通信安全也成为传感网的主要研究内容之一。特别是近年来端到端的数据安全引起了广泛关注，即秘密数据的明文仅出现在发送方与目的接收方，中间节点不论转发还是汇聚，都不对加密数据进行解密。

目前传感网数据汇聚技术的研究主要集中在以下几方面：

a）研究重点以减少数据汇聚中通信能耗为目地的路由协议或优化算法[1~4]等，基本不考虑数据安全性。

b）直接对密文进行统计分析。此类方案多用同态加密技术在不解密的情况下由各簇头完成相应的数据分析，如均值、求和等[5,6]；因汇聚节点不知道每个数据的明文无法进行异常数据检查，当存在内部攻击时，各汇聚头的汇聚结果将使基站得到错误的统计数据。

c）分布汇聚、集中统计。在该类方法中，各汇聚头主要排除冗余数据，最后的统计分析由基站完成[7~9]。但大部分该类方法的通信开销不理想且存在数据丢失问题，即多个相同数据只上报一个，但没有上报数据对应的计数值，即由多少个节点拥有该数据，给基站的统计分析带来不便。

本文提出了一种属于分布汇聚、集中统计的基于单向函数的传感器网络安全数据汇聚方案（secure data aggregation scheme for wireless sensor networks based on one-way function, SDAS-OWF）。

1　预备知识

1.1　ESPDA 简介

Cam 等人提出的 ESPDA 方案简介如下：

a）基站为每个节点 N_i 预装入 ID、与基站的配对密钥 k_i，

收稿日期：2011-12-05；**修回日期**：2012-01-09　　**基金项目**：国家自然科学基金资助项目（61173135）；广东省省部产学研重大项目（2011A090200031）

作者简介：郭江鸿（1975-），男，山西长治人，硕士，主要研究方向为无线移动安全、网络与信息安全、无线传感器网络（gjhjyu@gmail.com）；刘志宏（1967-），男，副教授，博士，主要研究方向为网络与信息安全、无线传感器网络；巫喜红（1975-），女，副教授，硕士，主要研究方向为网络与信息安全.

以及公共密钥 k。

b)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 $K=k_b\oplus k_i$ 作为自己本次加密密钥。

c)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。

d)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。

e)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。

1.2 网络模型及攻击者行为

STDA 采用与 ESPDA 方案相同的网络结构——分簇传感器网络(图 1),并作如下假设:a)节点间可通过合适的密钥协议建立配对密钥;b)通过现有的协议构造数据汇聚树,基站为根节点。

因已有诸多文献对建立配对密钥、构建汇聚树等问题进行了研究并取得了系列成果,所以作出上述假设而不对其进行具体讨论。

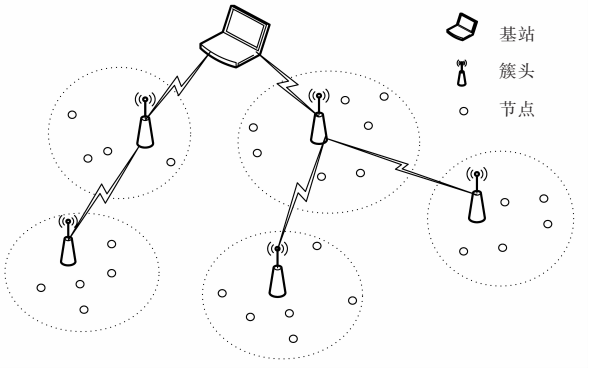

图 1 分簇传感器网络

本文假设攻击者可以窃听、截获所有经过网络的消息,可以发动以下攻击:

a)由于临近的传感器节点可能得到相同的数据,因此假设敌手可发动已知明文攻击;

b)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;

c)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;

d)敌手可以向网络中注入大量虚假数据发起 DoS 攻击。

2 SDAS-OWF 方案简介

SDAS-OWF 方案主要由系统初始化、节点数据处理、汇聚数据处理、基站数据处理等部分组成。

2.1 系统初始化

设网络共有 N 个节点,ID 分别为 $N_1,N_2,\cdots,N_N$,|ID| = 2 Byte,部署前,基站选取 N 个 l bit 的随机密钥 $K_1,K_2,\cdots,K_N$,l 为安全参数,生成种子 s^n 并计算反向 hash 链为

$$s^i = H_l(s^{i+1}) \quad 0\leqslant i<n-1 \tag{1}$$

对任意的 $i\in[1,N]$,服务器将 K_i、s^0、N_i、$H_l(\ )$ 预装入节点 N_i。其中,K_i 为节点 N_i 与基站的配对密钥,$H_l(\ )$ 表示取单向函数 $H(\ )$ 输出的前 l 比特,s^0 用于 μTESLA 广播验证秘密承诺。在传感器网络部署后的安全时段内,节点完成下列工作:

a)建立分簇网络[10]。簇 i 的簇头 A_i 选取种子并按照式(2)生成用于簇内广播的 2 级反向 hash 链,将 ${K_A}^0$ 簇内广播到 n 个簇内节点,作为安全时间之后簇内广播认证的秘密承诺。

$$\begin{aligned}&1\text{ 级} \quad {K_A}^i = {H_l}^1({K_A}^{i+1}) \quad 0\leqslant i<n-1\\ &2\text{ 级} \quad {K_A}^i = {K_A}^{i,0},{K_A}^{i,j} = {H_l}^2({K_A}^{i,j-1}) \quad 1\leqslant j<m\end{aligned} \tag{2}$$

使用 2 级 hash 链主要出于存储开销的考虑,以 $n+m$ 的存储开销提供 $n\times m$ 个秘密承诺。为便于分析,本文采用图 2 中较为理想的分簇网及汇聚树。

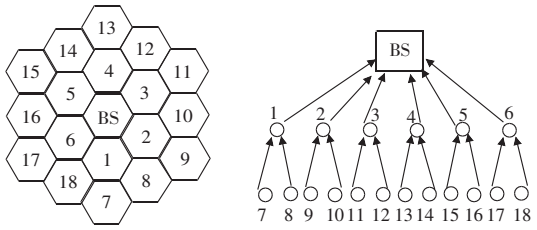

图 2 分簇网拓扑及汇聚树

b)节点与簇头及汇聚树上父簇头与子簇头间根据相应的密钥协商方法建立配对密钥,簇头维护一张孩子节点信息表,如表 1 所示。

表 1 簇内节点信息表(NIB)

ID	配对密钥	标志
N_1	k_1	C
…	…	…
A_1	W_1	C
…	…	…

其中:N_i 为簇内节点;k_i 为节点 N_i 与簇头的配对密钥;C 为序号,初始化为 0;A_1 为 A_i 在汇聚树上的孩子节点且配对密钥为 W_1。

2.2 节点数据处理

基站向网内发布带有新鲜数 r 的数据收集指令[11],收到基站指令后,各节点进行数据测量,但并不马上将数据发送到簇头。设簇头 A_i 测得的数据为 m_A,则 A_i 计算 $H_l(m_A)$,连同数据精度 e 一起进行簇内广播。精度 e 的作用为:设数据区间 D 为[1,10]且节点测量值为整数,若 $e=1$,则 D 划分为 10 个子区间[1],[2],…,[10];如 $e=2$,则 D 划分为 5 个子区间[1,2],…,[9,10],节点按测量值所在子区间取上限为最后结果。

N_i 根据精度要求计算自己的数据哈希值 $H_l(m_i)$ 并与簇头数据 hash 作比较,若异或结果不为 0,则 N_i 使用与基站的配对密钥 K_i 及与簇头的配对密钥 k_i 构造消息 M_i(如式(3))发往簇头:

$$N_i\parallel m_i\oplus H_l(r_i)\parallel r_i\oplus H_l(K_i\oplus r)\parallel H_l(m_i)\parallel \text{MAC} \tag{3}$$

其中:N_i 为节点 ID;m_i 为 N_i 采集的数据;r 为基站广播新鲜数;r_i 为节点生成的随机数;MAC 为消息用节点与簇头配对密钥 k_i 及哈希函数 $H_l(\ )$ 计算的认证码,如式(4):

$$H_l(k_i,r\parallel m_i\oplus H_l(r_i)\parallel r_i\oplus H_l(K_i\oplus r)\parallel H_l(m_i)) \tag{4}$$

2.3 簇头数据处理

从接收到第一个簇内消息开始计算,A_i 在时间 t 内接收数据并进行 MAC 认证,t 由簇内节点数目、采用的冲突碰撞退避算法等因素决定。对通过检查的消息 M_i,若簇内节点信息表

中与发送方对应的标志位 0,则 A_i 去掉 M_i 中的 MAC 并保存为 M_i', 同时更新簇内节点信息表中状态标志位为 1;否则丢弃。

若进行数据汇聚,A_i 建立一个簇 i 的汇聚结果表 ARL_i,表中每一项由 data 及 num 构成。为便于描述,用 A、B、C、D 分别表示 M_i'中用“ ‖ ”连接的四部分,$M_i'\rightarrow *$ 表示 M_i'中对应部分,如:$M_i'\rightarrow D$ 表示 M_i'中的 $H_l(m_i)$。设 A_i 根据自身数据构建的消息为 M_A',对于收到的 j 个消息 $M_i',\cdots,M_j'$,A_i 执行以下步骤:

a) ARL_i 为空则将消息放入 $ARL_i[0]\rightarrow$ data, 同时 $ARL_i[0]\rightarrow$num = 1。

b)若 $M_i'\rightarrow D\oplus ARL_i[s]\rightarrow D$,$0\leqslant s\leqslant$ length(ARL_i) − 1,为 0,则 $ARL_i[s]\rightarrow$num + 1;否则,将 M_i'加入 ARL_i,并设对应 num 为 1; 返回执行下一消息的比较。显然,对于 m_j 与 m_i 有

$$\Pr[H_l(m_i)\oplus H_l(m_j)]=\begin{cases}2^{-l} & m_i\neq m_j\\ 1 & m_i=m_j\end{cases} \tag{5}$$

c)当 j 个消息比较完后,A_i 将数据 M_A'加入 ARL_i,并设对应 num 为 $n-j$,n 为簇内节点数。

d)A_i 将按照消息格式(6)将汇聚结果 ARL_i 发送到汇聚树上的父节点 A_f。

$$A_i \parallel ARL_i \parallel MAC \tag{6}$$

e)重置 NIB 中所有标志位为 0。

其中,MAC 使用与汇聚树上父节点的配对密钥 W_i 计算。同理,A_f 对簇内数据及孩子簇头发送的汇聚结果按照相似的方法进行进一步数据整合,并将整合后的汇聚结果 ARL_f 发送到汇聚树上的上游节点,直到所有数据到达基站。

2.4　基站数据处理

作为整个汇聚树根节点的基站收到汇聚结果后,先对各条消息进行 MAC 验证,对合法的消息进行汇聚,再对最终汇聚结果解密。对于最终汇聚结果中的任一条消息 M_i',基站根据节点 ID 取得对应配对密钥,结合本轮数据收集新鲜数 r 进行解密,如式(7):

$$m_i=(M_i'\rightarrow B)\oplus H_l(H_l(K_i\oplus r)\oplus M_i'\rightarrow C) \tag{7}$$

若 $H_l(m_i)=M_i'\rightarrow D$,则保存数据 m_i 及对应的计数值 num。显然,本文方案同时提供了数据及相应计数值,有效避免了数据丢失问题,基站可根据汇聚结果进行有效的全网数据分析及异常数据检测等操作。

3　安全性分析

本节主要针对 2.2 节定义的攻击者行为进行方案安全性分析,并与 ESPDA 方案作比较。

1)抗选择明文攻击

本文方案消息式(3)中,加密部分为 $M_i\rightarrow B \parallel M_i\rightarrow C$,不失一般性,这部分形式化为

$$m_i\oplus G(r_i) \parallel h(r_i) \tag{8}$$

其中:$G(\,)$为随机数生成器,$h(\,)$为陷门置换。

Bellare 在随机预言机模型下证明了与式(8)相同的形式化加密方案是 IND-CPA 安全的。限于篇幅,本文对此不再证明,具体可参考文献[12]。ESPDA 方案使用 blowfish[13] 加密算法,同样可有效抵抗选择明文攻击。

2)节点妥协攻击

一般来说,传感器网络中节点妥协难以避免,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。

对于 SDAS-OWF,各节点与基站的配对密钥是由基站生成,因此从被妥协节点无法获取未妥协节点与基站的密钥;广播使用反向 hash 链秘密承诺认证。即使簇头 A_i 被妥协,敌手可得到该簇内未妥协节点的 MAC 计算密钥及簇内广播种子,可伪造小范围的簇内广播,但不会泄露未妥协节点的任何加密密钥,也不能在未妥协簇头所在簇伪造广播。

对于 ESPDA,每个节点预装入两个密钥:与基站的配对密钥 K_i 及全网共享密钥 k。K_i 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,对全网安全通信构成极大威胁。

3)抗主动攻击/DoS 攻击

对于 SDAS-OWF,敌手在不知道 N_i 与簇头配对密钥的情况下,伪造正确 MAC 的概率是一个可忽略的量;即使某节点 N_i 被妥协,敌手伪造 k 个含有正确 MAC 的虚假消息,这 k 个消息中任意一个通过簇头验证将使簇头的簇内节点信息表中对应状态标志更新为 1,后续的虚假消息将被丢弃。SDAS-OWF 可以较低的能耗有效抵抗敌手发动的主动攻击及以消耗节点能量为目的注入大量虚假数据的 DoS 攻击。

由于 ESPDA 中单点妥协暴露广播密钥的安全瓶颈,且基站的加密广播没有提供相应认证,所以抗主动攻击能力弱。内部敌手容易假冒基站发布全网广播,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,抗 DoS 攻击能力弱。综上所述,两种方案的安全性比较如表 2。

表 2　安全性比较

	抗选择明文攻击	抗妥协攻击	抗主动攻击	抗 DoS 攻击
SDAS-OWF	强	较强	较强	较强
ESPDA	强	较弱	弱	弱

4　开销分析

本章分析 SDAS-OWF 的存储开销、计算开销及通信开销。

1)存储开销分析

在 ESPDA 中,每个节点需预装入 ID、与基站的配对密钥、全网共享的广播密钥 k 及 Blowfish[12] 加密算法。

在 SDAS-OWF 中,服务器将 K_i、s^0、N_i、H_l 预装入节点,网络部署后,簇头需生成簇内广播所需的 2 级反向 hash 链,以 $n+m$ 的存储开销提供 $n\times m$ 个秘密承诺。如在 $n=m=50$,$l=64$ bit 时,以 800 Byte 存储空间提供 2 500 个秘密承诺。本文方案中节点的存储开销与 ESPDA 相当,簇头的存储开销略高于 ESPDA,但仍是传感器节点可以接受的。同时,簇头生成反向 hash 链用于簇内广播认证,提供了比全网共享广播密钥更高的安全性。

2)计算开销分析

因 ESPDA 只在簇内进行一层汇聚,因此以普通数据节点 N_i 及叶子簇头 A_i 的计算开销为例进行分析。

SDAS-OWF 中,簇头首先计算自己所测数据的 hash 及对应 MAC,然后进行簇内广播;节点接收到簇头广播后验证 MAC,并与自己的数据 hash 进行异或比较;异或结果不为 0 的节点再通过两个 hash 运算构造消息 M_i 发往簇头;簇头对接收的数据进行 MAC 检验,并对合法消息进行汇聚。

ESPDA 中,簇头先用公用密钥加密种子,计算 MAC 并进行簇内广播;簇内节点验证 MAC 并解密种子,生成新的模式码;节点将自己对应模式码附加 ID 与时戳发往簇头;簇头对模式码进行汇聚并对相同模式码的节点选择一个要求发送数据;节点使用与基站的配对密钥加密数据并生成 MAC,通过簇头发送到基站。由于 Cam 等未指出使用何种方法生成模式码,假设其使用单向函数。

设簇内有 n 个节点,在 ESDA 中有 d 个数据节点所测数据的 hash 值与簇头节点的数据 hash 值不同;数据区 D 分为 h 个子区间,对应 h 个模式码。在一次数据汇聚中,两种方案下数据节点与汇聚节点的计算复杂度如表 3 所示。

表 3 计算复杂度

	SDAS-OWF		ESPDA	
	A_i	N_i	A_i	N_i
加密			1	$\leqslant 1$
解密				1
hash	$d+3$	$\leqslant 5$	1	h

设 $n=40$,$h=10$, 冗余度 $=0.4$(簇内有 40% 的节点测得相同数据)时,SDAS-OWF 中簇头作 27 个 hash, 节点最多作 5 个 hash;EDPDA 中簇头作 1 个加密 - 1 个 hash, 节点作 1 个加密、1 个解密、10 个 hash。两种方案均使用低能耗的哈希计算或对称密钥技术,计算开销基本属于同一层次。要指出的是,ESPDA 中簇头没有对节点消息进行认证,存在安全隐患。

3)通信开销分析

本文使用 NS2 在图 2 所示的网络拓扑及汇聚树对两种方案的通信开销进行模拟,汇聚结果通过汇聚树到基站。由于 ESPDA 并未指出使用何种路由算法,因此以汇聚树作为 ESPDA 的路由。

由于簇内节点数据相关性较高,所以设各簇头数据是以基站数据(设为 30)为均值的正态分布随机值,节点数据是以簇头数据为均值的正态分布随机值(设簇内节点数为 40,每个节点平均邻居数为 40),同时对节点数据作取整处理。

本文采用 802.15.4 标准对两种方案的消息进行封装,该标准允许最多 102 Byte 的可变载荷,总长度最大为 128 Byte。在不同数据精度要求下,两种方案中各簇簇内数据包发送量及全网数据传输量如图 3、4 所示。

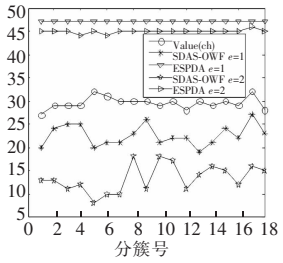

图 3 簇内数据包发送量

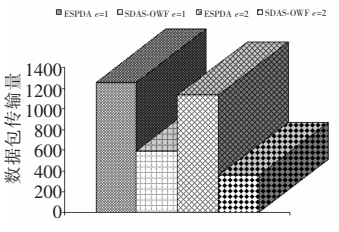

图 4 全网数据包发送量

图 3 中 Value(ch)为各簇头的值;e 为数据精度;SDAS-OWF $e=1$ 表示在 $e=1$ 时,本文方案各簇簇内数据包传输量;ESPDA $e=1$ 表示在 $e=1$ 时,ESPDA 方案各簇簇内数据包传输量。两种方案下一次汇聚过程中全网数据包传输量如图 4 所示。

从图 3、4 可知,EDAS 方案中每个节点都向簇头发送模式码,不同的数据通过簇头单播到基站,没有真正地降低通信负载。SDAS-OWF 方案通过簇内广播有效减少了簇内数据包传输量,同时通过多级汇聚进一步减少了簇头的负载,通信开销明显优于 ESPDA 方案。

5 结束语

数据汇聚是减少传感器通信开销的重要手段,同时,安全的端到端数据传输要求加密信息在到达远端服务器前应尽量避免解密操作。本文提出的方案不仅提供安全的端到端数据传输,而且在抗选择明文攻击、主动攻击、抗协攻击及 DoS 攻击方面比相关方案具有更高的安全性;同时,本文方案以较低的开销解决了汇聚过程中的数据丢失问题,有利于对全网数据的分析统计。

参考文献:

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

[2] AL-KARAKI I N, UI-MUSTAFA R, KAMAL A E. Data aggregation in wireless sensor networks-exact and approximate algorithms[C]//Proc of Workshop on High Performance Switching and Routing. [S. l.]: IEEE Society Press, 2004: 241-245.

[3] AONISHI T, MATSUDA T, MIKAMI S, *et al*. Impact of aggregation efficiency on GIT-routing for wireless sensor networks[C]//Proc of International Conference on Parallel Processing. [S. l.]: IEEE Society Press, 2006: 151-158.

[4] VILLAS L A, GUIDONI D L, ARA′UJO R B, *et al*. A scalable and dynamic data aggregation aware routing protocol for wireless sensor networks[C]//Proc of the 13th ACM International Conference on Modeling, analysis, and simulation of wireless and mobile systems. 2010: 110-117.

[5] PRZYDATEK B, SONG D, PERRIG A. SIA: secure information aggregation in sensor networks[C]//Proc of ACM SenSys Conference. 2003: 255-265.

[6] ACHARYA M, GIRAO J. Secure comparison of encrypted data in wireless sensor networks[C]//Proc of the 3rd International Symposium on Modeling and Optimization in Mobile, Ad hoc, and Wireless Networks. 2005: 47-53.

[7] CAM H, OZDEMIR S. Energy-efficient security protocol for wireless sensor networks[C]//Proc of IEEE VTC Fall Conference. New York: IEEE Society Press, 2003: 2981-2984.

[8] HUANG S I, SHIEH S, TYGAR J D. Secure encrypted-data aggregation for wireless sensor networks[J]. Springer Wireless Networks, 2010, 16(4): 915-927.

[9] 秦晓良,魏琴芳,张双杰. WSNs 中高效且适应性强的安全数据融合[J]. 计算机应用研究,2011,28(11):4299-4302.

[10] CHAN H, PERRIG A. ACE: an emergent algorithm for highly uniform cluster formation[C]//LNCS, vol 2920. 2004: 154-171.

[11] PERRIG A, SZEWCZYK R, TYGAR J, *et al*. SPINS: security protocols for sensor networks[J]. ACM Wireless Network, 2002, 8(5): 521-534.

[12] BELLARE M, ROGAWAY P. Random oracles are practical: a paradigm for designing efficient protocols[C]//Proc of the 1st ACM Conference on Computer and Communications Security. 1993.

[13] SCHNEIER B. Fast software encryption[C]//Proc of Cambridge Security Workshop Proceedings. [S. l.]: Springer-Verlag, 1994: 191-204.