计算机科学
COMPUTER SCIENCE

基于可逆数字水印的无线传感器网络可恢复数据聚合协议

高光勇, 韩婷婷, 夏志华

引用本文

高光勇, 韩婷婷, 夏志华.基于可逆数字水印的无线传感器网络可恢复数据聚合协议[J]. 计算机科学, 2023, 50(8): 333-341.

GAO Guangyong, HAN Tingting, XIA Zhihua. Recoverable Data Aggregation Protocol for Wireless Sensor Networks Based on Reversible DigitalWatermarking [J]. Computer Science, 2023, 50(8): 333-341.

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.220800089

基于可逆数字水印的无线传感器网络可恢复数据聚合协议

高光勇[1,2,3]　韩婷婷[1,3]　夏志华[4]
1 南京信息工程大学数字取证教育部工程研究中心　南京 210044
2 九江学院计算机与大数据科学学院　江西 九江 332005
3 南京信息工程大学计算机学院、软件学院、网络空间安全学院　南京 210044
4 暨南大学网络空间安全学院　广州 510632
(gaoguangyong@163.com)

摘　要　针对无线传感器网络数据安全认证协议的高能耗与传感器节点资源受限的对立问题,文中提出了一种基于可逆数字水印的聚合协议。一方面,在感知节点处,将水印嵌入到感知数据中,对水印数据进行基于椭圆曲线的同态加密,以此保证数据在传输过程中的私密性;在簇头节点处,对接收的数据只进行聚合和转发操作,以此减小网络通信开销;在基站处,通过提取水印对数据进行完整性认证。另一方面,该协议提出了一种基于环的聚合树,从而降低节点的传输能耗,延长网络生命周期。通过理论分析证明了所提协议将水印技术与数据聚合技术进行了更好的结合,具有较好的安全性和较低的计算开销,且能实现感知数据的完整性认证。此外,与同类算法的对比实验表明该协议在通信开销和时延方面都具有一定的优势。

关键词:可逆数字水印;完整性验证;数据聚合;同态加密
中图法分类号　TP393

Recoverable Data Aggregation Protocol for Wireless Sensor Networks Based on Reversible Digital Watermarking

GAO Guangyong[1,2,3], HAN Tingting[1,3] and XIA Zhihua[4]
1 Engineering Research Center of Digital Forensics, Ministry of Education, Nanjing University of Information Science & Technology, Nanjing 210044, China
2 School of Computer and Big Data Science, Jiujiang University, Jiujiang, Jiangxi 332005, China
3 School of Computer Science, Nanjing University of Information Science & Technology, Nanjing 210044, China
4 College of Cyber Security, Jinan University, Guangzhou 510632, China

Abstract　Aiming at the opposition between the high energy consumption of data security authentication protocol and the resource limitation of sensor nodes in wireless sensor networks, this paper proposes an aggregation protocol based on reversible digital watermarking. On the one hand, at the sensing node, the watermarking is embedded into the sensing data, and the elliptic curve is used to encrypt the watermarked data homomorphically, so as to ensure the privacy of the data in the transmission process. At the cluster head node, the received data are only performed aggregation and forwarding operations, so as to reduce network communication overhead. At the base station, the watermark is extracted to authenticate the integrity of data. On the other hand, this scheme proposes an aggregation tree protocol based on clustering protocol, which can reduce the transmission energy consumption of nodes and prolong the network lifetime. Theoretical analysis proves that the proposed protocol combines watermarking technology and data aggregation technology better, has good security and lower computation cost, and can realize the integrity authentication of perception data. In addition, experimental results show that, compared with the latest similar algorithms, the proposed protocol has certain advantages in communication cost and delay.

Keywords　Reversible digital watermarking, Integrity verification, Data aggregation, Homomorphic encryption

到稿日期:2022-08-09　返修日期:2022-11-22
基金项目:江西省自然科学基金重点项目(20192ACBL20031);国家自然科学基金(61662039,62122032,U1936118);国家重点研发计划(2020YFB1005600);南京信息工程大学人才引进启动基金(2019r070)
This work was supported by the Key Program of the Natural Science Foundation of Jiangxi Province, China(20192ACBL20031), National Natural Science Foundation of China(61662039,62122032,U1936118), National Key Research and Development Program of China(2020YFB1005600) and Startup Foundation for Introducing Talent of Nanjing University of Information Science and Technology(NUIST)(2019r070).
通信作者:夏志华(xia_zhihua@163.com)

1 引言

无线传感器网络在网络覆盖的区域内，可以利用节点采集相关信息，对数据进行处理和传输[1-3]。此类网络具有成本低廉、结构简单、体积小和组建方式自由的特点。在局部损坏的情况下，无线传感器网络其余部分依然可以正常工作，因而具有较高的应用价值。近年来，无线传感器网络已被广泛应用于各领域[4]，然而其通信线路不能做到私密可控，因此传输数据的安全得不到保证。同时无线传感器网络的感知节点存在计算能力差、资源有限等问题，故而在考虑数据安全的同时，还需兼顾无线传感器网络的低功耗问题[5-8]。

在无线传感器网络中，所有节点感知的实时数据会被发往基站，这会导致过大的传输数据流，从而增加了数据的处理难度。为了降低无线传感器网络的数据传输量，数据聚合技术[9-10]被提出并得到广泛应用。文献[11-12]提出了基于簇的私有数据聚合方案和分片混合聚合方案。在没有丢包的情况下，两种方案都可以获得精确的聚合结果，并且保证私有数据不被泄露给其他节点。但基于簇的私有数据聚合方案需要簇内节点频繁进行通信，且采用的多项式算法和逆矩阵计算会消耗大量资源，增加了节点能耗。在此基础上，文献[13-16]提出的安全聚合方案中，对分片技术的应用进行了优化，同时保证了数据的私密性，但是分片数据依然会进行频繁通信，从而产生大量通信开销。文献[17]利用簇成员监视簇头，验证簇头是否被俘获以及聚合结果是否改变，从而进一步保证数据的隐私性以及基站接收结果的完整性，但是该方案依旧存在计算方法复杂和通信频繁的问题。为了降低节点能耗，文献[18]进行了连续的数据聚合，通过选用安全通道来保证数据隐私性，并对数据进行过滤以减少对数据的操作，从而有效降低能耗。文献[19]提出了基于授权和验证的数据聚合模型来实现数据的安全传输，从而降低计算开销和通信开销。

传统的聚合方案无法在基站将聚合数据恢复成感知数据，因此聚合技术的应用受到了很大的限制。近年来，研究者们提出了可支持数据恢复且能进行数据完整性验证的聚合方案。文献[20]提出了一种基于数字签名的可恢复的聚合方案(Recoverable Concealed Data Aggregation，RCDA)，该方案可以在基站接收聚合数据后，将其恢复成每个感知数据，但是该算法的数据完整性验证依赖于数字签名，数字签名的生成要进行多次 Hash 函数操作。Yang 等[21]提出了一种同时保障隐私性与完整性的无线传感器网络可恢复数据聚合方案(Recoverable Privacy-preserving Integrity-assured Data Aggregation，RPIDA)，该方案利用同态加密技术保证数据的私密性，同时使用根据每个数据生成的相应消息验证码(Message Authentication Code，MAC)来实现数据完整性的验证。之后，Shen 等提出了一种基于身份的聚合签名方案[22]，并设计了验证器。根据聚合签名的优点，该方案同样能保证数据的完整性。文献[23]提出了一种利用可再生哈希链进行安全数据聚合的有效方案，该方案中的所有网络节点所存储的密钥都是一致的，且在每一轮生成 MAC 的过程中会进行密钥更新操作，从而保证了数据的机密性和完整性。文献[24]提出了一种基于信任的数据聚合协议(Data Validation and Integrity Verification for Trust Based Data Aggregation Protocol，DVIVTDAP)，每个节点根据信任特征求出信任值，形成汇聚树，然后用同态 MAC 标签对加密数据进行分片和签名处理，最后将处理后的数据发送给簇头进行聚合。以上方案虽然能将聚合数据恢复并进行完整性验证，但是都需要使用标签进行验证，这会产生额外开销，并且计算数字签名或消息验证码需要进行多次 Hash 函数操作，计算开销大。

可逆数字水印技术又称无损数字水印技术，在继承传统数字水印技术优点的基础上，可以将嵌入水印后的数据无误差地恢复为原始数据。近年来，由于可逆数字水印技术计算简单，相比传统加密技术，基于水印技术的信息隐藏的计算能耗更少，且不产生额外开销，已被应用于无线传感器网络[25-26]。在 Shi 等提出的方案中[27]，设计的同步点被用来对节点采集的数据流进行动态分组，同时每两组数据构成一个验证组，验证组中的第一分组负责生成水印，然后该水印被嵌入到第二分组中。验证组中任何一个分组数据被篡改，都会被有效检测到。该方案采用可逆数字水印技术，因此能在基站将接收的数据无损地恢复为原始感知数据。但如果方案中出现伪同步点，则会导致分组紊乱，从而导致即使数据流未被攻击，原始数据也无法恢复。Alromih 等[28]提出了一种针对物联网应用的随机水印过滤方案，该方案利用原始感知数据生成水印，然后将水印随机嵌入到有效数据的不同位置。嵌入水印后的数据不仅具有私密性，而且能抵御大部分攻击，但是该方案对原始数据的大小有一定要求，因此算法的应用受到限制。以上的水印方案虽然能保证数据隐私性并进行完整性验证，但是算法依然会出现虚警率过高或者应用依然有限制的问题，并且上述方案未使用数据聚合技术，无法减少数据传输量。

文献[29]提出了一种可恢复的数据聚合协议(Reversible Digital Watermarking based Protocol for Data Integrity，RDWPDI)，该协议运用可逆数字水印技术和基于椭圆曲线的同态加密技术，在感知节点处将读取的原始数据分片并嵌入水印，同时对原始数据进行椭圆曲线加密，然后将嵌入水印后和加密后的数据都发往簇头节点进行聚合。该方案可以同时保证数据完整性与机密性，但是在该方案中同样的数据进行操作后需要被发送两次，增加了通信开销。

为了将水印技术与数据融合技术更好地结合，且能同时发挥两种技术的优势，在数据传输过程中保证数据隐私性的同时使用聚合技术降低通信开销，并且能在基站中仅通过嵌入水印后的感知数据本身进行数据完整性验证，不依赖于额外标签或多次发送数据验证，本文提出了一种基于可逆数字水印的无线传感器网络可恢复数据聚合协议(Recoverable Data Aggregation Protocol for Wireless Sensor Networks Based on Reversible Digital Watermarking，RDAPRDW)。该协议将可逆数字水印技术和椭圆曲线加密技术相结合，将水印嵌入到感知数据中，再进行基于椭圆曲线的同态加密，最后将数据发送到簇头节点进行聚合。实验结果表明，对感知数据进行水印嵌入和加密操作可以保证数据的安全性，并且在基站能无损恢复。同时，该协议使用可逆数字水印技术进行完整性认证，与以往聚合协议中常用的 MAC 等技术相比，其计算开销更小，因此该协议的综合性能相比以往的聚合协议研究得到了进一步提高。该协议的主要贡献如下：

(1)在聚合协议中利用嵌入水印的感知数据本身进行数据完整性验证，无需额外标签或重复发送数据。目前用于验证无线传感器网络数据完整性的大部分聚合协议都是使用MAC或数字签名，在计算MAC或数字签名时，Hash函数资源的消耗远大于嵌入水印的资源消耗，因此本协议降低了计算开销。本协议通过提取数据中的水印进行验证，无需额外标签，降低了通信开销。

(2)能确定恶意节点。该协议利用簇头节点 ID_x 生成水印序列，原始感知数据根据每轮分簇后分配的簇内节点 ID_i 生成一个衍生数据，并按照簇内节点 ID_i 的序号大小排序，依次选取水印嵌入到原始感知数据和其衍生数据中。在基站恢复数据时可以检测到个体数据是否被攻击，若验证数据被攻击，可以根据该数据的节点 ID 确定恶意节点。

2 协议介绍

2.1 协议数据传输模型及流程

在本协议中，先用DEEC[30]协议对所有的感知节点进行动态分簇，确定簇头和簇成员，簇内数据的传输方式是单跳传输，即数据由簇成员节点直接传给簇头节点。如果感知节点到基站的距离小于到所有簇头的距离，则感知节点不加入任何分簇，而是直接将数据发往基站。图1是无线传感器网络中分簇后的簇成员节点的数据传输模型图，所有的簇成员节点将数据进行操作后，把加密数据发往簇头。本协议会进行多轮分簇，聚合树的构造也随着分簇的变化而变化。

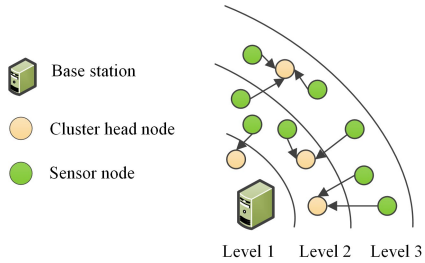

图1 簇成员节点的数据传输模型图

Fig. 1 Data transmission model diagram of cluster member nodes

分簇完成后，利用簇头节点构造一棵聚合树，基站作为树的根节点。假设将节点的分布划分为 q 环，每个环内的节点在聚合树的同一深度。$Pa(q)$ 代表第 q 环节点的父节点所在环，每一环中的簇头节点的父节点都在它的上一环，即分布在第 q 环中的簇头节点的父节点在第 $q-1$ 环，其表达式为：

$$Pa(q)=q-1 \tag{1}$$

簇头节点先选取上一环中距离自己近的节点作为父节点，如果簇头节点到上一环中的两个节点的距离相同，则选取剩余能量更大的节点作为父节点。如果簇头节点到选取的父节点距离大于到基站的距离，则簇头节点直接选择基站作为父节点。图2是无线传感器网络中由簇头构成的聚合树的数据传输模型图，簇头节点对簇成员节点传输来的数据进行聚合后，发送给它的父节点，父节点只负责转发数据，不进行其他操作。

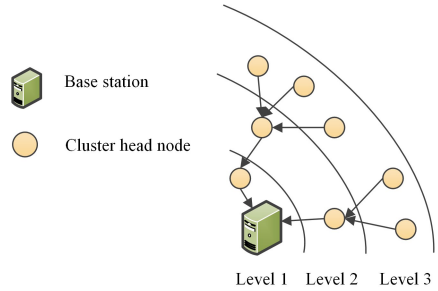

图2 聚合树的数据传输模型图

Fig. 2 Data transmission model of aggregation tree

协议流程如下：(1)利用DEEC协议对所有感知节点进行分簇，并构造聚合树；(2)系统参数初始化；(3)在感知节点处生成与原始感知数据 S_i 相关的衍生数据 S_i^*，将水印嵌入到 S_i，S_i^* 这两个数据中，然后对两个水印数据进行聚合并编码得到数据 m_i；(4)对数据 m_i 进行基于椭圆曲线的同态加密，然后将密文发送给簇头节点；(5)簇头节点对接收到的所有加密数据进行聚合操作后将其发往父节点或者基站；(6)基站解密数据，提取水印，进行完整性验证。具体流程如图3所示。

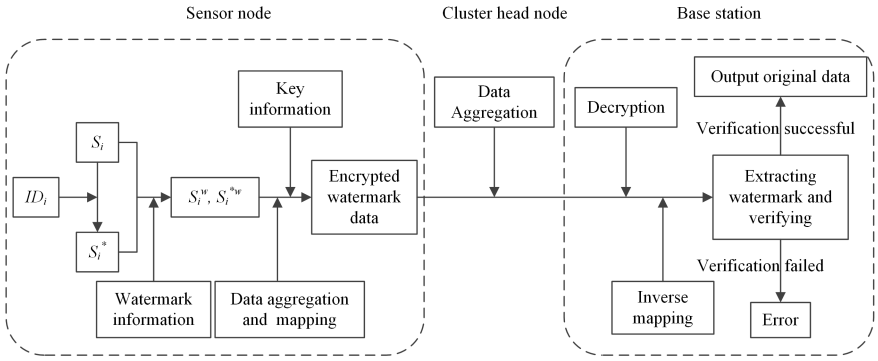

图3 协议流程图

Fig. 3 Protocol flow chart

2.2 系统初始化设置

整个网络先进行分簇，由簇头对每个簇内成员按序分发一个簇内节点 ID_i，基站通过簇头向整个网络广播可靠的聚合请求信息，得到该轮分簇所有感知节点中最大的数据 S_{max} 以及最远簇头到基站的距离，根据距离确定环的划分，进一步为构造聚合树做准备。本协议使用椭圆曲线加密(Elliptic Curve Eigamal，ECEG)[20]对数据进行操作，以保证数据在传输过程中的安全性。构造有限域 F_p 上的一条合适的椭圆曲线 E，p 是一个大素数，并在椭圆曲线上选定一点 P 作为基点，n 是基点的素数阶，这几个参数构造

出椭圆曲线的参数组 T 定义为：

$$T=(E,P,p,n) \tag{2}$$

基站随机选择一个 k 作为私钥，其中 $k\in[1,n-1]$ 且 $k\in Z$，并生成公钥 $Y=kP$，得到公私密钥对 (k,Y)。基站将椭圆曲线的参数和生成的公钥发送给每个感知节点，使感知节点预先存入参数组和基站的公钥。本协议的符号及其解释如表 1 所列。

表 1 符号及解释

Table 1 Symbols and corresponding interpretations

符号	解释	符号	解释
S_i	原始感知数据	m_i	编码后的数据
S_i^*	感知数据的衍生数据	$\|$	连接符
ID_x	簇头节点的 ID	0^β	β 位 0
ID_i	分配的簇内节点 ID	W	生成的二进制水印序列
S_i^w	嵌入水印后的原始感知数据	w	1 bit 水印
S_i^{*w}	嵌入水印后的衍生数据	l	本次所有待编码数据最大值的二进制长度

2.3 可逆数字水印的生成及嵌入

感知节点发送感知数据前，根据式(3)得到与原始感知数据 S_i 相关的衍生数据 S_i^*，其中 ID_i 是每轮分簇后分配给该节点的簇内 ID。S_i^* 的计算式为：

$$S_i^*=floor\left(\frac{S_i}{ID_i}\right) \tag{3}$$

本协议将可逆数字水印嵌入到原始感知数据 S_i 和其衍生数据 S_i^* 中，二进制水印序列 W 在感知节点中根据其所在分簇的簇头节点 ID_x 生成。按照簇内的节点 ID_i 的排序，从 W 中的对应位置取 1 bit 水印 w 嵌入到每个感知节点的原始数据和其衍生数据中。若 W 长度小于簇内节点个数，则循环嵌入。将原始数据和衍生数据嵌入可逆水印后所得的数据称为水印数据。为保证水印隐秘不可察觉，水印序列 W 的计算式如式(4)所示：

$$W=ID_x\oplus(rv(10-v)) \tag{4}$$

其中，r 取值为 4；v 代表第几轮分簇，当 $v>10$ 时，则取其个数位。

本协议中使用的是基于奇偶不变性的可逆数字水印技术，该可逆数字水印技术的具体描述如下。

假设有数据对 (x,y)，根据式(5)先计算数据对两个值的差值 d 为：

$$d=x-y \tag{5}$$

将二进制水印 w 根据式(6)和式(7)分别嵌入到 x,y 中，得到嵌入水印后的数据对 (x',y')，即：

$$x'=x+floor(d/2)+w \tag{6}$$

$$y'=y-floor(d/2)-w \tag{7}$$

提取水印时，首先对原始数据对的差值 d 加 $2y$ 后，可得：

$$d+2y=x+y \tag{8}$$

同理，令 $d'=x'-y'$，对嵌入水印后像素对的差值 d' 加上 $2y'$ 后可得：

$$d'+2y'=x'+y' \tag{9}$$

由于根据式(6)、式(7)可推 $x'+y'==x+y$，因此可知 $d+2y==d'+2y'$。由于 $2y$ 与 $2y'$ 都是偶数，因此可推断出，d 与 d' 的奇偶性相同，故：

$$LSB(d)=LSB(d') \tag{10}$$

$$d=2floor(d/2)+LSB(d) \tag{11}$$

将式(5)－式(7)、式(10)以及式(11)代入到 $d'+LSB(d')$ 中，可得式 $d'+LSB(d')=2\times d+2\times w$，从而推得：

$$d=\frac{d'+LSB(d')}{2}-w \tag{12}$$

在提取水印时，根据接收端得到的数据对 (x',y')，可求得 d'。因为 $w\in\{0,1\}$，且 d 与 d' 的奇偶性相同，从而可根据式(12)求出水印 w 和 d，然后根据式(6)、式(7)求解原始数据对 (x,y)。本协议用原始数据 S_i 和衍生数据 S_i^* 构造数据对，使用基于奇偶不变性的可逆数字水印技术将由 ID_x 生成的水印嵌入到其中，得到嵌入水印后的数据 S_i^w 和 S_i^{*w}。当数据传输到基站后，由于水印嵌入技术可逆，因此可根据提取的水印验证数据完整性，并将数据无损恢复。

2.4 数据编码聚合

在感知节点处，需要对数据进行两次编码，第一次编码是根据式(13)对水印数据进行操作，其中 t_{ij} 是编码后的数据，$data_{ij}$ 是待编码数据，$\beta=l(j-1)$，l 是本轮 $S_{\max}$ 的二进制长度，对原始数据 S_i^w 进行编码时，j 取 1，编码衍生数据 S_i^{*w} 时，j 取 2。t_{ij} 的计算式为：

$$t_{ij}=data_{ij}\|0^\beta \tag{13}$$

当 S_i^w 和 S_i^{*w} 是待编码数据时，编码后的数据则分别用 t_{i1} 和 t_{i2} 表示，然后将数据 t_{i1} 和 t_{i2} 聚合为一个数据。具体操作实例如下。

假设有数据 S_i^w 和 S_i^{*w} 分别取值为 15 和 5。假设本轮 l 取 4，对两数据进行编码可得 $t_{i1}=(1111)_2$，其中 $\beta=4\times(1-1)=0$；同理可得 $t_{i2}=(01010000)_2$，将两个数据聚合可得 $t_i=(01011111)_2$。

在水印数据编码聚合得到数据 t_i 后，对 t_i 根据式(14)进行第二次编码操作，目的是为了保证数据在簇头节点能进行数据聚合。

$$m_i=t_i\|0^\alpha,i=1,2,3,\cdots \tag{14}$$

其中，m_i 是编码后的数据，t_i 是待编码数据，由于一次编码后 t_i 的长度是 l，因此 $\alpha=2l(i-1)$，i 按照簇内节点 ID_i 进行选取。

实例如下：假设 $t_i=(01011111)_2$ 是本轮第二个待编码数据，则 i 取 2。根据式(14)计算 $\alpha=2\times4\times(2-1)=8$，因此对数据 t_i 编码后得到的数据 $m_i=(0101111100000000)_2$。

2.5 数据加密及聚合

同态加密的思想源于私密同态，可以直接对加密后的数据进行操作，故而可以在保证数据机密性的同时对数据进行操作。对明文先进行环上的加法和乘法的运算后再进行加密操作与先加密后对密文进行相应的运算这两种操作的结果是等价的。

假设有两个明文 a 和 b，$En(\cdot)$ 是加密运算，如果存在一种有效算法 $\odot$，其运算满足：

$$En(a)\odot En(b)=En(a\odot b) \tag{15}$$

则称 $En(\cdot)$ 是运算 $\odot$ 下的同态加密。当 $\odot$ 代表加法时，称该加密为加法同态加密；当 $\odot$ 代表乘法时，称该加密为乘法同态加密。

椭圆曲线上的全体点构成一个加法群，点与点之间可以

进行加法运算，具有加法同态的性质。本文协议采用基于椭圆曲线的同态加密技术，目的是在保证传输数据隐私性的同时，还能进行聚合操作。

在感知节点中，对编码后的数据 m_i 进行加密，具体步骤如下：

(1)根据式(16)将明文数据映射成椭圆曲线上的一点，即：

$$M_i = map(m_i) \tag{16}$$

(2) 传感器节点选择一个随机数 $\eta_i \in [0, n-1]$，用于对 M_i 进行加密；

(3)根据式(17)对数据 M_i 进行加密，得到密文 C_i。

$$C_i = (R_i, S_i) = (\eta_i P, M_i + \eta_i Y) \tag{17}$$

加密数据后，将其发送给簇头节点，簇头节点在接收到所有簇内节点发送的消息 C_i 后，根据簇内分配的节点 ID_i 对数据排序，再对簇内节点传来的 $\zeta-1$ 个密文进行聚合并转发，其中 ζ 是每个簇的节点总数，聚合密文 C 为：

$$C = \sum_{i=1}^{\zeta-1} C_i = (\sum_{i=1}^{\zeta-1} R_i, \sum_{i=1}^{\zeta-1} S_i) = (\sum_{i=1}^{\zeta-1} \eta_i P, \sum_{i=1}^{\zeta-1} M_i + \sum_{i=1}^{\zeta-1} \eta_i Y) \tag{18}$$

2.6 解密及水印验证

基站在接收到聚合密文 C 后，会对数据先进行解密，然后利用反映射得到聚合数据，再使用聚合数据的恢复公式，计算出聚合中的每个数据，从而提取水印进行验证，具体计算过程如下：

(1) 根据式(19)对聚合密文 C 进行解密得到聚合明文 M，其中 k 是基站选择的私钥。

$$\begin{aligned} M &= -kR + S \\ &= \sum_{i=1}^{\zeta-1} M_i + Y \sum_{i=1}^{\zeta-1} \eta_i - kP \sum_{i=1}^{\zeta-1} \eta_i \\ &= \sum_{i=1}^{\zeta-1} M_i + Y \sum_{i=1}^{\zeta-1} \eta_i - Y \sum_{i=1}^{\zeta-1} \eta_i \end{aligned} \tag{19}$$

(2) 根据反映射式(20)计算出聚合数据 m：

$$m = rmap(M) = m_1 + m_2 + \cdots + m_{\zeta-1} \tag{20}$$

(3)根据聚合数据恢复计算式(21)，计算出每一个聚合数据 t_i 为：

$$t_i = m[2(i-1)l, 2il-1] \tag{21}$$

(4)t_i 是对水印数据进行编码聚合后的数据，根据式(22)计算出聚合成 t_i 的两个数据，j 取 1 时，恢复的数据为 S_i^w；j 取 2 时，恢复的数据为 S_i^{*w}。聚合数据恢复计算式如下：

$$t_{ij} = t_i[(j-1)l, jl-1] \tag{22}$$

(5)根据式(12)提取嵌入到原始数据和衍生数据中的水印，并计算原始数据 S_i 与衍生数据 S_i^*。

(6)为验证数据的完整性，根据接收到的簇内节点 ID_i 确定数据 S_i 在水印序列中对应的水印 w，并将其与步骤(5)计算所得的水印进行比较。如果水印匹配成功，并且步骤(5)计算得到的原始数据与衍生数据以及簇内节点 ID_i 满足式(3)，则认定数据验证成功。

为了保证本协议中数据完整性验证的准确性，每次对两个相邻数据 t_i 和 t_{i+1} 进行验证，若两个数据都验证成功，且节点 ID_i，ID_{i+1} 及对应的水印是相邻的，则认定无误；否则，判定数据 t_i 被篡改，将 t_i 对应的原始数据舍弃，再验证数据 t_{i+1} 和 t_{i+2}，以此类推，继续验证。

3 安全性及计算量分析

3.1 安全性分析

命题 1 在本协议中可逆数字水印技术和加密技术相辅相成，可以较好地保证数据隐私性、阻止未授权聚合、进行数据完整性验证、抵抗簇头节点俘获攻击以及检测恶意节点。

证明：

(1)保证数据隐私性。由于在感知节点中对原始数据进行了水印嵌入及椭圆曲线加密，因此加密后的数据与原始数据不同，故假设攻击者即使拦截到传输数据，也得不到有效数据。即使簇头节点被攻击者俘获，但由于簇头节点内部不存储密钥信息，没有基站的私钥，因此攻击者无法解密数据。因此，本协议保证了数据的机密性，可有效抵御窃听攻击。

(2)阻止未授权聚合。未授权聚合指攻击者在未俘获任何节点、不知道任何有用信息的情况下能执行聚合操作。本协议使用的是基于椭圆曲线的同态加密技术，在进行数据聚合时需要使用到椭圆曲线的点加和点乘函数，然而攻击者无法得知椭圆曲线的构成参数，就无法实现未授权聚合，因此，本协议能够抵抗未授权聚合攻击。

(3)数据完整性验证。本协议在传输过程中一直是以密文的形式传输，攻击者不知道曲线构成参数，无法篡改出符合要求的密文。由于本协议的完整性验证依赖于提取的簇内节点 ID_i 与水印，即使有恶意节点能够插入、删除或篡改聚合密文，本协议在基站处依然能对接收的数据进行完整性认证。l 是所有待编码数据最大值的二进制长度，在操作时会根据 l 将原始数据与衍生数据编码聚合，故而被篡改的数据与对应的簇内节点 ID_i 满足式(3)的概率为 $1/2^l$。本协议每次同时验证两个相邻的数据，因此两个数据的簇内节点 ID_i 都满足条件的概率为 $1/2^{2l}$，篡改数据的水印也恰好满足条件的概率为 1/2，两个数据的水印满足条件的概率为 1/4，故本协议在簇头节点被俘获的情况下的漏检率为：

$$PR = \frac{1}{2^{2l+2}} \tag{23}$$

因此，本协议能够有效抵抗针对数据完整性的攻击，如篡改、伪造和注入虚假数据。

(4)抵抗簇头节点俘获攻击。若攻击者俘获了簇头节点，即使获取了 Y 和 P，也无法在多项式时间内解出基站私钥，因此无法解密密文。此外，攻击者俘获了簇头节点，可能篡改聚合数据，根据式(23)可知，篡改后的数据满足验证条件的概率极低。综上所述，本协议能够抵抗聚合节点的俘获攻击。

(5)检测恶意节点。在基站根据 2.6 节的步骤对数据进行验证，如果验证失败，则找到验证失败的数据对应的簇内节点 ID_i，根据节点 ID 和簇内分配的节点 ID_i 的映射表判断恶意节点。

本协议将选择 4 种应用于无线传感器网络的聚合方案进行对比分析，分别是 RCDA[20]，RPIDA[21]，RDWPDI[29] 以及 DVIVTDAP[24]。这 4 种方案都无法针对单个数据进行验证，因此如果攻击者发送的虚假数值在合理范围，则这几种方案都无法有效检测恶意节点。此外，文献[24]使用的是对称

加密技术，如果一方的密钥泄露，那么密文就不再安全，因此方案[24]无法有效抵抗节点的俘获攻击。

3.2 计算量分析

命题 2 本协议的簇头节点和簇内节点的计算开销较低。

证明：基站计算资源充足，因此主要比较感知节点和簇头节点的资源能耗。假设 H 代表 Hash 值计算函数，N 代表一次点乘操作，A 代表一次 160 bit 或 271 bit 的点加操作，XOR 代表一次异或操作。

首先考虑在感知节点处的能耗，假设一个分簇内有 ζ 个成员，本协议与 RCDA，RPIDA 以及 RDWPDI 几个方案都采用了相同的加密方式，在感知节点中需要执行两次点乘操作和一次点加操作，因此这几种方案在感知节点处的共有开销为 $2N+1A$。此外，RCDA 要生成数字签名用于完整性验证，需要进行一次 271bit 的点乘操作和一次 Hash 函数操作，因此 RCDA 在感知节点处的开销为 $3N+1A+1H$；RPIDA 在生成消息验证码 MAC 时，需要执行一次 Hash 函数，因此 RPIDA 在感知节点处的开销为 $2N+1A+1H$；RDWPDI 除了加密，还要进行多次异或操作来生成水印用于嵌入，然后对嵌入水印后的数据进行 RC4 加密，因此 RDWPDI 在感知节点处的开销为 $2N+1A+(\zeta-1)XOR+RC4$；而本协议除了加密外只需要进行水印嵌入，是简单的数学运算，能量损耗可忽略不计，因此本方案在感知节点处的开销仅为 $2N+1A$。DVIVTDAP 方案没有分簇，因此将该方案中聚合树的叶子节点作为感知节点进行比较，该方案使用对称加密技术对感知数据进行加密，默认该方案选用 DES 对称加密方法，然后对每一个分片数据生成消息验证码 MAC。如果原始感知数据分片成 K 片，则需要对 K 片数据都生成相应的数字标签，即进行 K 次 Hash 函数操作，因此 DVIVTDAP 在感知节点处的开销为 $DES+KH$。

在考虑簇头节点能耗时，为方便比较，本协议与 RCDA，RPIDA 以及 RDWPDI 的参数 p 取 160bit 的大素数，DVIVTDAP 方案中的密钥也选取 160 bit。本协议与 RCDA，RPIDA 以及 RDWPDI 的簇头对加密数据进行了聚合，因此这几种方案在簇头节点的共有开销为 $2(\zeta-1)A$。此外，在簇头位置，RCDA 方案还需要对签名进行 $\zeta-1$ 次 271bit 的点加操作，因此 RCDA 在簇头节点处的开销为 $3(\zeta-1)A$；RPIDA 需要对每个数据对应的消息验证码 MAC 进行 $\zeta-1$ 次异或操作，因此 RPIDA 在簇头节点处的开销为 $(\zeta-1)(2A+XOR)$；RDWPDI 需要进行一次 RC4 加密，因此 RDWPDI 在簇头节点处的开销为 $2(\zeta-1)+RC4$；而本协议不需要进行额外操作，因此本方案在簇头节点处的开销仅为 $2(\zeta-1)A$。DVIVTDAP 方案不存在簇头，在此将该方案中聚合树的父节点作为簇头节点进行比较，父节点将本身的感知节点加密，再生成对应的 MAC，与子节点传输来的数据与标签进行聚合操作，假设父节点有 ζ 个子节点，则每个父节点需要进行 $2(\zeta-1)$ 次加法聚合，因此 DVIVTDAP 在簇头节点处的开销为 $2(\zeta-1)A+DES+KH$。

根据上述分析，表 2 列出了 5 种方案的计算开销。

表 2 5 种方案的计算开销

Table 2 Calculation costs of five schemes

方案	感知节点	簇头节点
RCDA	$3N+1A+1H$	$3(\zeta-1)A$
RPIDA	$2N+1A+1H$	$(\zeta-1)(2A+XOR)$
RDWPDI	$2N+1A+(\zeta-1)XOR+RC4$	$2(\zeta-1)A+RC4$
DVIVTDAP	$DES+KH$	$2(\zeta-1)A+DES+KH$
RDAPRDW	$2N+1A$	$2(\zeta-1)A$

根据表 2 可以直观地看出，无论是簇头节点还是感知节点，本协议的计算资源消耗都比方案 RCDA，RPIDA 以及 RDWPDI 低。方案 DVIVTDAP 采用的加密方式是对称加密，虽然在感知节点处该方案比本协议能量消耗低，但是对称加密中的密钥管理存在风险，易被拦截，数据安全性更低。在簇头节点处，本协议的计算操作比 DVIVTDAP 更少，因此在簇头节点处，本协议更节约计算资源。

4 仿真实验和结论

4.1 仿真参数设置

本协议在 NS3 的仿真平台上进行了多次仿真实验，并取多次实验结果的平均值作为实验结论，具体的仿真参数如表 3 所列。

表 3 系统参数

Table 3 System parameters

仿真参数选项	参数设定值
目标区域/m	100×100
基站位置/m	(50,50)
节点数/个	100
发送能量/(μJ/bit)	0.3

本协议选择与 RCDA，RPIDA，RDWPDI 以及 DVIVTDAP 这 4 种方案进行对比，因此参数设置都相同，且默认通信链路是无差错的。

在目标区域内，随机生成 100 个感知节点，节点分布如图 4 所示，其中五角代表基站的位置。根据 DEEC 协议将生成的节点分簇，然后用簇头节点构造聚合树。

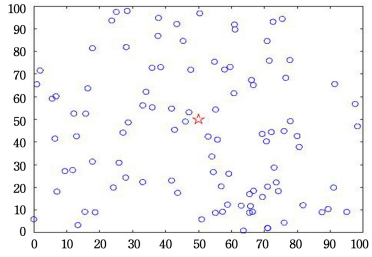

图 4 节点分布图

Fig. 4 Nodes distribution

4.2 确定节点环数划分的实验分析

本文采用 DEEC 协议对 100 个节点进行分簇，簇头的选取考虑了节点剩余能量和网络平均能量，从而均衡网络能耗，避免剩余能量低的节点多次充当簇头节点，进一步延长网络生命周期。在分簇后，构造一棵基于环的聚合树。如果构造树的环数太少，数据传输的平均距离会较长，导致节点能耗变大。如果构造树的环数太多，虽然缩短了数据的平均传输距离，但是会增加中间节点转发数据包的数量，因此根据

节点数量和节点间的平均距离寻找一个合适的环数，可以使传输距离和转发数据包产生的总能耗更小。

通过实验模拟确定最佳的环数值，使网络模型最大程度地节省能耗。在此实验中，固定发送数据包大小为 4 000 bit。图 5 给出了本协议在运行 1000 轮后，本协议传输节约的网络能量消耗。根据实验结果可得，当环数为 3 时，本协议节约的网络能量是最多的，因此本协议后续实验中的环数均设为 3。

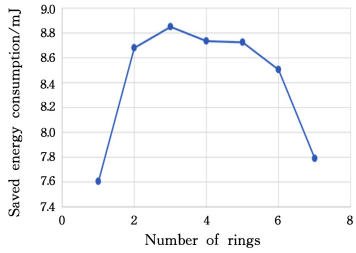

图 5 环数对网络能耗的影响

Fig. 5 Influence of ring number on network energy consumption

4.3 通信能耗分析

本协议在分簇过程中，规定簇内成员节点到簇头距离大于到基站的距离时，可直接将数据发往基站。这样做可有效避免簇头集中在某一区域时，簇成员节点传输数据产生更多能耗。在分簇后，本协议构造聚合树，利用树状拓扑实现多跳，减少数据远距离传输的次数，进一步节约能源。

各方案中数据分组的头部长度、节点 ID 和时间戳等都保持一致，因此只分析发送的消息大小，不考虑其他额外开销。借助点压缩技术，椭圆曲线 E 上的一个点表示为 161 bit，本协议与 RCDA，RPIDA 以及 RDWPDI 的密文是椭圆曲线 E 上的两个点，因此占 322 bit。此外，RCDA 方案生成的数字签名长度为 271 bit，故 RCDA 的感知节点发送消息的大小为 593 bit。RPIDA 生成一个 MAC 标签，其长度为 160 bit，故需要发送 482 bit 的消息。RDWPDI 还需要发送两个分片数据，两个分片数据占 16 bit，因此需要发送的消息长度为 338 bit。本协议无需发送除密文外的标签或数据等，因此各个感知节点需要发送的消息长度仅为 322 bit。DVIVTDAP 使用 DES 加密，密文长度仅为 64 bit，对于每个数据生成一个 MAC 标签，标签长度为 160 bit，因为要将数据分 K 片，所以每个节点需要发送 $224K$ bit，在此取分片个数 K 为 2，故而每个节点需要发送 448 bit。为了便于比较，参考方案 DVIVTDAP 节点每发送或接收 1 bit 数据消耗 0.3 μJ，簇成员节点开销对比如图 6 所示。从实验结果中可以直观地看出，本协议的通信开销低于其他几个方案。

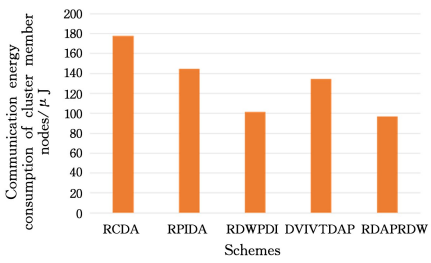

图 6 5 种方案的簇成员节点通信能耗比较

Fig. 6 Comparison of communication energy consumption of cluster member nodes in five schemes

在簇头节点处，RDWPDI 方案除了发送加密信息，还要发送嵌入水印后的分片数据，因此 RDWPDI 方案在簇头节点处的通信开销更高。其余几种方案在聚合后，数据分组长度基本不变，因此通信开销相对于簇成员节点基本不变。簇头节点通信开销对比结果如图 7 所示。

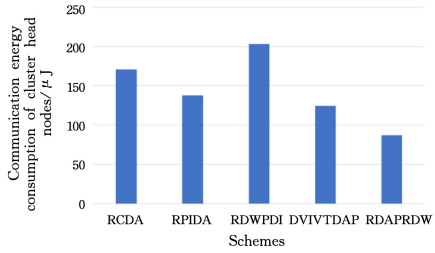

图 7 5 种方案的簇头节点通信能耗比较

Fig. 7 Comparison of communication energy consumption of cluster head nodes in five schemes

4.4 时延比较

无线传感器网络的聚合协议需要考虑聚合时延和发送时延。当节点个数固定时，簇头数量以及数据包大小等都会影响时延，簇头数量过少或数据包太大都会导致时延的增加。本协议以及对比方案中的数据包长度是固定的，因此在实验中仅改变簇头的数量。由于方案 DVIVTDAP 中不进行分簇，因此在图 8 所示的对比实验中，本方案仅与 RCDA，RPIDA 以及 RDWPDI 进行比较，从图中观察到，随着簇头数量的增加，每个簇头需要聚合的数据量减少，时延也随之减小，然而簇头数量过多会使网络中的通信能耗过高。根据实验得到簇头数量占节点总数的 10% 左右时，既可以降低通信能耗，又可以最大程度地减少时延。

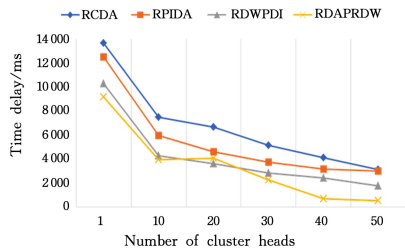

图 8 簇头个数对时延的影响

Fig. 8 Effect of the number of cluster heads on time delay

方案 DVIVTDAP 不进行分簇，因此本方案以节点个数为变量，与该方案进行比较。由上述分析可得，本方案的簇头数量是节点总数的 10% 最佳，故而在比较时，固定簇头节点数量为节点总数的 10%。从图 9 可以观察到，本方案的时延低于方案 DVIVTDAP 的时延。

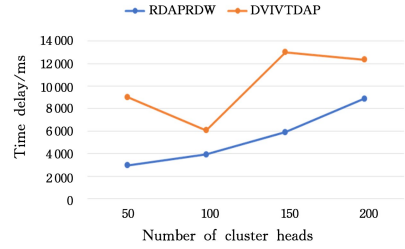

图 9 时延随节点总数变化的比较

Fig. 9 Comparison ofvariation of delay with the total number of nodes

4.5 数据完整性认证

由于拦截攻击无法获得椭圆曲线的构成参数，因此无法篡改出符合要求的密文，故在基站处能检测出篡改数据。如果存在恶意节点发送错误数据，本协议将使用数字水印技术进行数据完整性验证，验证成功的数据就恢复，未验证成功的数据将报错，从而确定恶意节点。数据恢复率指成功恢复的数据占所有数据的比例。

在现实情况中，会有一定的丢包率存在，当同时传输过多的数据包造成网络堵塞、拥堵时，丢包率会升高，因此簇头数会影响水印丢包率，从而影响数据恢复率。同时，篡改率也会影响数据的恢复率。在假设有恶意节点篡改的情况下，数据恢复率随着篡改率和簇头数量的变化如图 10 所示。

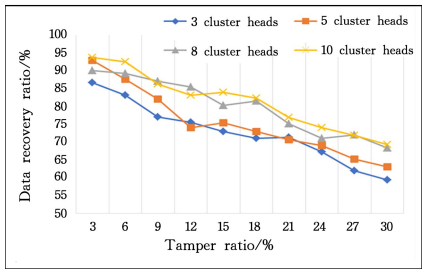

图 10 数据恢复率的变化

Fig. 10 Changes in data recovery rate

结束语 本协议提出了基于可逆数字水印的无线传感器网络可恢复数据聚合协议。该协议中，原始感知数据根据分发的簇内节点 ID_i 生成衍生数据，对原始数据和衍生数据嵌入水印后，再将两个数据编码加密进行传输，保证了数据的高机密性。同时，由于本方案并未使用额外的数字标签等进行数据完整性认证，大大地降低了节点的能源消耗。经过实验验证，本协议能高效完成完整性验证，且运行速度快，资源消耗少，相对于现有的其他基于分组的聚合方案而言，性能得到了进一步的提高。

由于无线传感器网络在现实生活中的应用越来越广泛，对于一些敏感数据的安全性也得到了越来越多人的重视，在保证无线传感器网络数据传输安全的同时，降低无线传感网络的能耗也变得十分重要。在未来的研究中，我们将重点改进数据完整性认证方案的复杂度，在保证数据安全的同时，进一步降低能源消耗。

参 考 文 献

[1] JAN S R U, KHAN R, JAN M A. An energy-efficient data aggregation approach for cluster-based wireless sensor networks [J]. Annals of Telecommunications, 2021, 76(5): 321-329.

[2] MANDEEP K, AMIT M. Data aggregation algorithms for wireless sensor network: A review [J]. Ad Hoc Networks, 2020, 100(3): 102083.

[3] SANKARALINGAM S K, NARMADHA A S. Energy aware decision stump linear programming boosting node classification based data aggregation in WSN[J]. Computer Communications, 2020, 155: 133-142.

[4] ADAM M S, ANIST M H, ALI I. Object tracking sensor networks in smart cities: Taxonomy, architecture, applications, research challenges and future directions [J]. Future Generation Computer Systems, 2020, 107: 909-923.

[5] GOMATHI S, GOPALA K C. Malicious node detection in wireless sensor networks using an efficient secure data aggregation protocol[J]. Wireless Personal Communications, 2020, 113(4): 1775-1790.

[6] ADNAN M, YANG L, AHMAD T, et al. An unequally clustered multi-hop routing protocol based on fuzzy logic for wireless sensor networks[J]. IEEE Access, 2021, 9: 38531-38545.

[7] JURADO-LASSO F F, CLARKE K, CADAVID A N, et al. Energy-aware routing for software-defined multihop wireless sensor networks[J]. IEEE Sensors Journal, 2021, 21(8): 10174-10182.

[8] LIU X, YU J, LI F, et al. Data aggregation in wireless sensor networks: from the perspective of security[J]. IEEE Internet of Things Journal, 2019, 7(7): 6495-6513.

[9] RAWAT P, CHAUHAN S. A Novel Cluster Head Selection and Data Aggregation Protocol for Heterogeneous Wireless Sensor Network [J]. Arabian Journal for Science and Engineering, 2022, 47: 1971-1986.

[10] ZHANG J, HU P, XIE F, et al. An energy efficient and reliable in-network data aggregation scheme for WSN[J]. IEEE Access, 2018, 6: 71857-71870.

[11] HE W, LIU X, NGUYEN H, et al. PDA: Privacy-preserving data aggregation in wireless sensor networks[C]// 26th IEEE International Conference on Computer Communications (IEEE INFOCOM 2007). IEEE, 2007: 2045-2053.

[12] HE W, LIU X, NGUYEN H V, et al. PDA: privacy-preserving data aggregation for information collection[J]. ACM Transactions on Sensor Networks (TOSN), 2011, 8(1): 1-22.

[13] FANG W, WEN X Z, XU J, et al. CSDA: a novel cluster-based secure data aggregation scheme for WSNs[J]. Cluster Computing, 2019, 22(3): 5233-5244.

[14] HUA P, LIU X, YU J, et al. Energy-efficient adaptive slice-based secure data aggregation scheme in WSN [J]. Procedia Computer Science, 2018, 129: 188-193.

[15] ZHANG X, LIU X, YU J, et al. Energy-efficient privacy preserving data aggregation protocols based on slicing[C]// 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). IEEE, 2019: 546-551.

[16] ZHOU L, GE C, HU S, et al. Energy-efficient and privacy-preserving data aggregation algorithm for wireless sensor networks [J]. IEEE Internet of Things Journal, 2019, 7(5): 3948-3957.

[17] HE W, LIU X, NGUYEN H, et al. A cluster-based protocol to enforce integrity and preserve privacy in data aggregation[C]// 2009 29th IEEE International Conference on Distributed Computing Systems Workshops. IEEE, 2009: 14-19.

[18] WANG T, QIN X, DING Y, et al. Privacy-preserving and energy-efficient continuous data aggregation algorithm in wireless sensor networks[J]. Wireless Personal Communications, 2018, 98(1): 665-684.

[19] NELS S N, SINGH J. Security-aware authorization and verifica-

tion based data aggregation model for wireless sensor networks [J]. International Journal of Numerical Modelling: Electronic Networks, Devices and Fields, 2021, 34(3): e2844.

[20] CHEN C M, LIN Y H, LIN Y C, et al. RCDA: Recoverable concealed data aggregation for data integrity in wireless sensor networks[J]. IEEE Transactions on parallel and distributed systems, 2011, 23(4): 727-734.

[21] YANG L, DING C, WU M. RPIDA: Recoverable Privacy-preserving Integrity-assured Data Aggregation Scheme for Wireless Sensor Networks[J]. KSII Transactions on Internet and Information Systems(TIIS), 2015, 9(12): 5189-5208.

[22] SHEN L, MA J, LIU X, et al. A secure and efficient id-based aggregate signature scheme for wireless sensor networks[J]. IEEE Internet of Things Journal, 2016, 4(2): 546-554.

[23] MIN W, RUIXIANG C, SHUNBIN H. A Secure Data Aggregation Approach in Hierarchical Wireless Sensor Networks[C]// the 10th International Conference. ACM, 2016: 1-7.

[24] ANITA D D, ROSLIN S E. Data validation and integrity verification for trust based data aggregation protocol in WSN[J]. Microprocessors and Microsystems, 2021, 80(1): 1-6.

[25] KUMAR S, SINGH B K, PUNDIR S, et al. Role of Digital Watermarking in Wireless Sensor Network[J]. Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science), 2022, 15(2): 215-228.

[26] WAZIRALI R, AHMAD R, AL-AMAYREH A, et al. Secure watermarking schemes and their approaches in the iot technology: an overview[J]. Electronics, 2021, 10(14): 1744.

[27] SHI X, XIAO D. A reversible watermarking authentication scheme for wireless sensor networks[J]. Information Sciences, 2013, 240: 173-183.

[28] ALROMIH A, AL-RODHAAN M, TIAN Y. A Randomized Watermarking Technique for Detecting Malicious Data Injection Attacks in Heterogeneous Wireless Sensor Networks for Internet of Things Applications[J]. Sensors, 2018, 18(12): 4346.

[29] JIANG W X, ZHANG Z X, WU J J. Reversible digital watermarking-based protocol for data integrity in wireless sensor network[J]. Journal of Communication, 2018, 39(3): 118-127.

[30] QING L, ZHU Q X, WANG M W. Design of a distributed energy-efficient clustering algorithm for heterogeneous wireless sensor networks[J]. Computer Communications, 2006, 29(12): 2230-2237.

GAO Guangyong, born in 1973, Ph. D, professor, is a senior member of China Computer Federation. His main research interests include reversible data hiding, computer networks security, multimedia information security, and digital image processing.

XIA Zhihua, born in 1983, Ph.D, professor, Ph.D supervisor, is a senior member of China Computer Federation. His main research interests include digital forensic and encrypted image proces-sing.

(责任编辑:喻藜)