ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.20, Supplement, December 2009, pp.239−249

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

无线传感器网络基于云团认证的虚假数据过滤机制*

彭 舸[1], 林亚平[1,2+], 易叶青[1]

[1](湖南大学 计算机科学与通信学院,湖南 长沙 410082)

[2](湖南大学 软件学院,湖南 长沙 410082)

False Data Filtering Mechanism Based on Cloud-Built Authentication Model in Wireless Sensor Networks

PENG Ge[1], LIN Ya-Ping[1,2+], YI Ye-Qing[1]

[1](School of Computer and Communication, Hu'nan University, Changsha 410082, China)

[2](School of Software, Hu'nan University, Changsha 410082, China)

+ Corresponding author: hnu_ger@hotmail.com

Peng G, Lin YP, Yi YQ. False data filtering mechanism based on cloud-built authentication model in wireless sensor networks. *Journal of Software*, 2009,20(Suppl.):239−249. http://www.jos.org.cn/1000-9825/ 09028.htm

Abstract: Most of recent false data filtering mechanisms in WSN added *t*-MAC (Message Authentication Code) for the data packets, these mechanisms are usually restricted within the *t*-threshold safe limits, and do not support dynamic routing. Based on the idea of a virtual witnesses cluster, adopting the perturbation-based polynomial technology, this paper proposes a authentication algorithm that made a number of nodes within a cloud cooperate to generate the certification polynomial, adopting the perturbation-based polynomial technology and increasing the difficulty of an attack. On this basis, the proposed false data filtering mechanism can verify the validity of data immediately, and support dynamic routing. Theoretical analysis and simulation experiments show that the new method is not limited by the *t*-threshold and save as more energy as the transmission jump increasing. Compared with the other mechanisms, the method enhances the anti-trapping ability. It is more suitable for the network with low credibility and the long-distance transmission application.

Key words: wireless sensor network; false data filtering; cloud-built authentication; polynomial

摘 要: 无线传感器网络中已有的虚假数据过滤机制一般是在数据包后附加 *t* 个消息认证码(MAC),这类机制通常受到 *t* 门限值的限制,且难以支持动态路由.基于虚拟证人簇的思想,采用混淆多项式技术,用云团内部多个节点协作生成的方式构造认证多项式,加大了攻击难度;在此基础上提出的虚假数据过滤机制,能够即时验证数据的有效性,并且支持动态路由.理论分析和仿真实验结果表明,新算法不受 *t* 门限值的限制,随着传输跳数的增加,其节能效果更为明显,与已有的虚假数据过滤机制相比,抗俘获能力增强,更适用于可信度较低的网络以及远距离传输场景的应用.

关键词: 无线传感器网络;虚假数据过滤;云团认证;多项式

* Supported by the National High-Tech Research and Development Plan of China under Grant No.2006AA01Z227 (国家高技术研究发展计划(863)); the Natural Science Foundation of Hu'nan Province of China under Grant No.09JJ6097 (湖南省自然科学基金); the Project Supported by Scientific Research Fund of Hu'nan Provincial Education Department of China under Grant No.06B047 (湖南省教育厅优秀青年科研项目)

Received 2009-05-01; Accepted 2009-07-23

随着无线传感器网络(简称WSN)被越来越多地应用于军事、商业、检测等数据敏感领域,其安全问题引起人们极大的关注[1−4].由于无线传感器网络工作环境特殊,且节点能力有限,容易受到虚假数据攻击(false data attack),即攻击者通过获得的密钥信息和妥协节点捏造虚假事件、篡改正在传送的数据包或者发送重复数据包等.这种攻击不仅会使用户接收到错误的报告,而且浪费WSN有限的网络资源,造成严重后果[5,6].由于传感器网络的主要任务是采集数据,因此,如何在"以数据为中心"的无线传感器网络中,识别和过滤虚假数据,保证收集到的数据具有新鲜性、完整性、真实性是至关重要的.

已有的能够识别和过滤虚假数据的数据认证机制[7−10]一般都是借鉴数字签名的思想,在发送的数据包后额外附加t个MAC信息来实现虚假数据的识别和过滤.然而,这类方法的安全性受到t门限值的限制,即一旦攻击者俘获超过t个节点,就可以联合利用预置在节点内的秘密信息伪造虚假数据包,已有的这些安全措施将难以防范[5].另外,无线传感器网络通常采用睡眠调度算法,节点也会由于被俘或者能量耗尽等物理原因而失效,这常常会引起路由变更[6],然而上述机制[7,10]通常假设在每一轮数据采集的过程中,数据的传输路径是固定不变的.因此,需要研究一种不受t门限值限制并且支持动态路由的数据认证机制.

为此,本文采用虚拟证人簇的思想,提出了一种基于云团认证的虚假数据过滤机制,能够突破 t 门限值的限制.另外,将整个网络按云团划分,节点根据所属云团预置密钥信息,确保在每个簇中都能找到有效的中转节点,使得该认证机制能够支持动态路由.

本文第 1 节概述相关工作.第 2 节给出基于云团认证的虚假数据识别和过滤算法.第 3 节在理论上对认证机制的安全性能和系统开销等方面进行定性分析.第 4 节进行相应的仿真实验.第 5 节对本文的工作进行总结.

1 相关工作

针对WSN中的虚假数据攻击,Zhu等人率先提出了IHA[7].该方案在发送的数据包后附加t个MAC来实施认证,并通过建立关联路径在数据包传输过程中进行交叉逐步认证,以尽早地对虚假数据进行过滤,达到节约网络资源的目的.由于每个数据包后附加了t个MAC,攻击者要想伪造虚假数据就必须获得t个不同密钥集的密钥,从而增加了攻击难度.不足之处在于:关联路径的建立是在网络的初始阶段完成,一旦节点失效,路径维护将带来巨大的通信开销.

几乎同时,Ye等人提出了SEF[8,9]的虚假数据过滤机制,该方案采用基于概率的方法,通过在每个节点中存储多个密钥来实现认证,无须建立和维护关联路径.不足之处在于:这种机制的抗俘获能力差,容易由于少数节点被俘获而泄露较多的密钥信息,最坏情况下,整个网络只能容忍t−1 个妥协节点.

Li等人认为攻击者不仅可以在网络中注入虚假数据,同时也可以篡改附加的MAC信息,从而使其他的中转节点将本来正确的数据误认为是虚假数据包而进行删除,于是基于簇组织和选票机制,提出了PVFS[10],以同时抵制两种攻击.该方案的中转节点以一定的概率对数据进行验证,采用Sloom Filter记录真假选票数,真选票数达到某一阈值的数据包无须再认证,从一定程度上降低了总体开销.但PVFS算法并没有从根本上解决之前的问题,相反地,攻击者可以篡改已经达到阈值、无须再次验证的数据包,而逃脱认证机制的过滤.

Ma等人认为上述一些方法受t门限值的限制,提出一种不受门限值限制的虚假数据过滤机制[11],该方案假设存在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,虽然不受t门限值的限制,但该方案的假设过高.此外,文献[12–15]等认为上述的虚假数据过滤算法都是基于对称密钥机制,其安全性不够,于是,基于公开密钥技术提出一些虚假数据过滤算法.但公钥密钥机制对计算和存储的要求较高,对于资源有限的WSN来说,大多数学者认为不适宜采用公钥认证机制.

Zhang 等人在文献[16]中,提出了一种轻量级的,可直接认证的数据认证技术.该方案采用多项式技术对消息进行即时认证,能够容忍大量的妥协节点;与基于公钥数字签名技术的机制相比,降低了计算开销.该方案在验证数据完整性方面体现出了良好的应用价值,但对于源节点数据的真实性却没有充分考虑:如果源节点是一个被俘获的节点,那么攻击者就可以通过源节点发送虚假数据,而系统难以通过认证识别这类虚假数据.

2 基于云团认证的数据过滤机制

本节基于虚拟证人簇的思想,提出一种基于混淆多项式的云团协作认证码生成算法,在此基础上,提出一种基于云团认证的虚假数据过滤机制.

2.1 基本假设和认证信息的预置

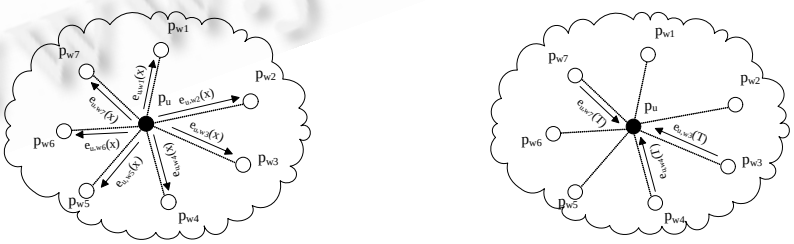
2.1.1 基本假设

假设无线传感器网络被部署在类似于战场的敌对环境中,用于监测敌方的坦克和军队等,采用Mica2 系列的节点,每个节点都有足够的缓存空间用来存放数据、密钥等.节点一般是通过飞行器投放的,假定同一簇的节点是一起投放的,并且它们的物理位置彼此靠近,大部分节点能直接通信,因此,在同一个簇内节点所采集的数据往往具有较高时空相关性[17].假设在节点投放之前,簇内随机生成k个云团(cloud cluster),每个节点属并且只属于一个云团.假设网络刚部署后有一段时间是安全的,节点将在这段时间内将完成相应的初始化操作.邻居节点之间通过建立对偶密钥加密节点之间的交换信息,如文献[18,19].本文假定基站(sink)部署在一个相对安全的位置[6−16]难以被攻击者俘获,为了保证基站的安全,基站距离监测区需有一定的距离,数据需要经过多跳(hop)才能从监测区传回基站;传输路径可以按GPRS[20]或GEAR[21],即在相邻且距基站最近簇中选择下一跳的传输节点,如果相邻的簇到基站的距离一样,则采用右手法则[22].

2.1.2 认证信息的预置

网络部署之前,基站在有限域 $F\left(2^L\right)$ 上,随机选择认证 Hash 函数 $H(.)$,k 个云团认证多项式 $K_{c_i}(x)\ i\in(1,2,...,k)$,$k$ 个秘密数 θ_{c_i},以及系统多项式 $f(x,y,z)$.其中 $K_{c_i}(x)$ 是次数为 d_w 的单变量多项式,θ_{c_i} 是长度为 r 的随机数,$f(x,y,z)=\sum_{0\le i\le d_x,0\le j\le d_y,0\le k\le d_z}A_{i,j,k}x^iy^jz^k$,其中 $A_{i,j,k}$ 是多项式系数,系统参数 d_x,d_y,d_z 分别是 x,y,z 的度数,该三元多项式可理解为 f(发送节点 ID,认证节点 ID,报告认证信息).

节点将预置 3 方面的信息:(1) 网络唯一的节点标识 ID;(2) 每个节点根据自身 ID,结合系统认证多项式 $f(x,y,z)$ 生成各自的消息认证函数和验证函数,例如:节点 u 生成消息认证函数 $auth_u(y,z)=f(u,x,z)\oplus r_{u,0}$ 和消息验证函数 $ver f_u(x,z)=f(x,u,z)\oplus r_{u,1}$,其中 $r_{u,0}$ 和 $r_{u,1}$ 是长度为 r 的随机数.(3) 节点 $p_u\in C_i^\gamma$ 根据所属的云团序号 i,预置与基站共享的云团认证多项式 $K_{c_i}(x)$ 和秘密数 θ_{c_i},γ 为簇序号.在网络部署之后的初始化过程中,节点 p_u 在有限域 $F\left(2^L\right)$ 上随机选择一个二元多项式 $e_u(x,y)=\sum_{0\le i\le d_w,0\le j\le\mu}A_{i,j}x^iy^j$,其中 $2\le\mu<n-2$ 是与参与协作认证的节点个数相关的系统参数.节点 p_u 根据邻居节点的 ID 为云团内部所有的邻居节点 $p_{w_1}\ (i=0,1,...,n-1)$ 计算多项式分量 $e_{u,w_i}(x)=e_u(x,w_i)$,如图 1(a)所示,节点 p_{w_1} 保留该分量.随后,节点 p_u 计算 $K_{c_i,u}(x)=e_u(x,u)\oplus K_{c_i}(x)\oplus\sigma_u$,并将 $K_{c_i}(x)$ 和 $e_u(x,y)$ 删除,仅保留 $K_{c_i,u}(x)$.



(a) 云核节点分配多项式 (b) 云核节点收集多项式分量

Fig.1 Collaboration certification within the cloud (n=8, μ =2)

图 1 云团内部协作认证,取 n=8, μ =2

2.2 基于混淆多项式的云团协作认证码生成算法

本小节提出了一种基于混淆多项式的云团协作认证码生成算法.通过云团内部多个节点协作认证,攻击者若想伪造一个认证信息,需要俘获云团内部的多个节点,加大了攻击难度;由于采用混淆多项式技术,引入了随机因子使得该算法具有后向安全性,增强了节点的抗俘获能力,因此该算法能够突破 t 门限值的限制,从而提高了系统的安全性能.

2.2.1 混淆多项式

在描述云团协作认证码生成算法之前,先介绍一下混淆多项式技术的基本原理[24,25].

$K(x)$ 为有限域 $F(2^L)$ 上,一个 d_w 次的单变量多项式,如果攻击者得到 d_w+1 个不同的变量及其所对应的函数值,例如: $\langle x_1,K(x_1)\rangle,\langle x_2,K(x_2)\rangle,...,\langle x_{d_w+1},K(x_{d_w+1})\rangle$,则可以通过联立线性方程组,解出 $K(x)$.假设 $K^r(x)=K(x)\oplus\sigma_t$,其中 σ_t 是长度为 r 的随机数.对于某个确定的 T 而言, $K^r(T)$ 和 $K(T)$ 的高 $L-r$ 位相同,低 r 位不同,因此取 $K_{c_i}(x)$ 的高 $L-r$ 位为云团认证码.即使攻击者得到某个时间段内的 d_w+1 个不同的变量及其所对应的函数值 $\langle i,K^r(i)\rangle$,也只能求出 $K^r(x)$,得到当前时段的认证码,而无法推出以后时间段的认证码.即本文所说的"后向安全性".

2.2.2 云团协作认证码生成算法

定义 1(云核节点). 某时间段内,发起云团内部协作认证的节点称为云核节点 p_{nuclear} ,云核节点随机指定.

下面,我们以图 1(b)为例,描述云团内部节点协作认证的过程:

Step 1. 云核节点 $p_u\in C_i^\gamma$ 在云团内部广播数值 T,该值是由源节点形成聚合数据的实时时间决定的(在第 2.3.1 中将有介绍).

Step 2. 节点 $p_{w_j}\in C_i^\gamma$ 计算 $e_{u,w_i}(T)$,并将 $\{w_i,e_{u,w_i}(T)\}$ 发送给 p_u .

Step 3. 当 p_u 收集到 $\mu+1$ 个分量, $\{\langle w_i,e_{u,w_i}(T)\rangle,i=0,1,...,\mu\}$,可以通过计算解线性方程组 $\sum_{j=0}^{\mu}B_j(w_i)^j=e_{u,w_i}(T),i=0,1,\ldots,\mu$,重构 μ 次多项式 $e_u(T,y)$,如图 1(b), p_u 收集到 $\{w_3,e_{u,w3}(T)\}$, $\{w_4,e_{u,w4}(T)\}$ 和 $\{w_7,e_{u,w7}(T)\}$.

Step 4. p_u 计算 $K_{c_i,u}^r(x)=e_u(x,u)\oplus K_{c_i,u}(x)=K_{c_i}(x)\oplus\sigma_u$,可知 $K_{c_i,u}^r(x)$ 和 $K_{c_i}(x)$ 的高 $L-r$ 位相同,因此,取 $K_{c_i,u}^r(x)$ 的高 $L-r$ 位为云团认证码,记为 $V_{c_i}=K_{c_i,u}^r(x)\bmod 2^r$.

2.3 基于云团认证的虚假数据过滤机制

本节提出了一种基于云团认证的虚假数据过滤机制,图 2 为该机制的原理图.由于数据包需要附加 t 个云团的认可才被认为有效,因此源节点在形成聚合数据以后需要征集 t 个由上述方法生成的云团认证码,作为参数计算出系统认证多项式,形成最终的数据包,然后经过多跳传送到基站.在中转路由的过程中,中转节点可以发起云团内部节点协作,以一定的概率对数据包进行认证,并对虚假数据进行过滤,以达到节省网络资源的目的.基站重构系统认证多项式,能够识别和过滤逃脱了中转节点认证的虚假数据.以下将详细介绍该机制的具体实现.

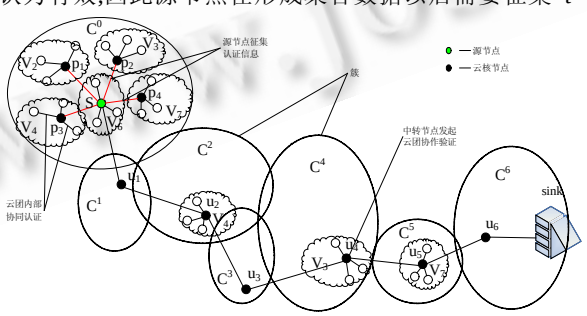


Fig.2 A false data filtering mechanism based on cloud-built authentication model

图 2 基于云团模型的虚假数据过滤机制,取 t=4, μ =2 数据包的生成

定义 2(作证云核节点集). 源节点 $S\in C_s^0$, $1\le s\le k$ 形成聚合数据后,随机选取 $t(t+1<k)$ 个所属云团各不相同的邻居节点的集合为作证云核节点集,记为 $VP=\left\{p_1,p_2,...,p_t,p_i\in C_{v_j}^0,1\le v_j\le k,v_i\ne s\right\}$.

定义 3(作证云团集). 作证云核节点所属云团的集合,记为 $VC=\left\{v_1,v_2,...,v_t,p_i\in VP,p_i\in C_{v_j}^0\right\}$.

根据假设,同一个簇内节点所采集的数据往往具有较高时空相关性,因此,我们让同一个簇内的邻居云团对源节点数据的正确性和真实性做出评判,并生成云团认证码.如图 2 所示,源节点 $S\in C_{v_6}^0$,可以选取 $VP=\left\{p_1,p_2,p_3,p_4\right\}$, $VC=\left\{v_2,v_3,v_4,v_7\right\}$.

数据包的具体生成过程如下:

Step 1. 源节点 S 生成聚合消息 M,将 $\left\{H(M),T\right\}$ 发送给 $p_i\in VP,(i=1,2,...,t)$, T 为当前时间,如果时刻 T_1 比时刻 T_2 早,则有 $T_1<T_2$.

Step 2. 节点 $p_i\in VP$ 将收到的数据与自身的数据 M_i 比较,计算 $Diff=H(M)-H(M_i)$,若 $Diff\le\varepsilon$,则认为聚合数据 M 是真实的,发起云团内部协作认证,生成认证码 $V_{c_{v_i}^0}$,发送给 S;若 $Diff>\varepsilon$ 则说明聚合数据 M 是可疑的,发送一个拒绝消息,其中参数 ε 表示相关性的误差范围,用户可以根据需要进行设置.如果存在 $r(r\le k-1-t)$个证人节点发送的拒绝消息,则源节点将重新选择 r 个认证云核节点,更新 VP 和 VC.若收到了 $r(r>k-1-t)$个拒绝消息,则令源节点 S 停止发送数据,并删除已收集和处理的数据,算法结束.

Step 3. S 计算 $V_{total}=V_{c_{v_1}^0}\oplus V_{c_{v_2}^0}...\oplus V_{c_{v_t}^0}$,将 V_{total} 发送给 $p_i\in VP(i=1,2,...,t)$.

Step 4. $p_i\in VP$ 计算 $V_{c_{v_i}^0}=V_{total}\oplus V_{c_{v_i}^0}+\theta_{c_{v_i}^0}$,发送给 S.

Step 5. 源节点 S 生成最终的数据包 $\left\{S,E_{V_{c_{c_q'}^0}}(M),MAC_{u,M}(y),T,VC,V_{c_{v_1}^0},V_{c_{v_2}^0},...,V_{c_{v_t}^0}\right\}$,其中 $V_{c_{c_q'}^0}$ 为 S 所在云团的实时认证码,由 S 发起云团内部协作生成,系统认证函数为 $MAC_{S,M}(y)=f\left(S,y,H\left(V_{total},E_{V_{c_{c_q'}^0}}(M)\right)\right)\oplus r_{S,0}$.

2.3.1 中转节点的认证

如图 2 所示,数据包从源节点 S 到基站的传输路径为 $u_1,u_2,...,u_6,u_i\in C^i$,由于每个节点属于并且只属于一个云团,当中转节点所属云团序号属于作证云团集,例如 $u_v\in C_{q'}^v,1\le q'\le k$, $q\in VC$ 时,节点 u_v 将发起云团内部协作,求出云团实时认证多项式 $K_{c_{q'}}^r(T)$,继而得到云团认证码 $V_{c_{q'}}$,对数据包进行认证.由于 $|VC|=t$,则 $q'\in VC$ 的概率为 t/k,所以节点 u_v 对数据包的认证概率为 $p_v=t/k$.中转节点接收到数据包以后的处理过程如下:

Step 1. 比较数据包中的 T 与节点中的记录 T',若 $T\le T'$,则认为该数据包为重复数据包,将其删除,并向基站报告,算法结束;若 $T>T'$,则执行下一个步骤.

Step 2. 将 T' 更新为 T 的值,判断所在云团序号 q' 是否属于作证云团集 VC,若属于则执行下一个步骤,若不属于则将数据包直接转发给下一个节点.

Step 3. 假设 $q'\in VC$,如图 2 中的 $\left\{u_2,u_4,u_5\right\}$,以 u_4 为例,已知 $u_4\in C_{v_3}^4$,可以推出 $V_{total}^{u_4}=\left(V_{c_{v_3}^0}-\theta_{c_{v_3}^4}\right)\oplus V_{c_{v_3}^0}$, $V_{c_{v_3}^0}=V_{c_{v_3}^4}=K_{v_3,u_4}^r(T)\bmod 2^r$.

Step 4. 计算 $verf_{u_4}\left(S,H\left(V_{total}^{u_4},E_{V_{c_{c_q'}^0}}(M)\right)\right)-MAC_{S,M}(u_4)=r_{u_4,1}-r_{S,0}$,如果 $\left(r_{u_4,1}-r_{S,0}\right)\in\left\{0,1,...,2^r-1,2^L-\left(2^r-1\right),...,2^L-1\right\}$,则认为该数据包有效,转发给下一个节点;否则,将数据包删除,并向基站报告.

2.3.2 基站的认证

基站拥有所有云团的认证多项式和云团秘密数,因此基站可以全面地校验数据包,进一步识别并过滤逃逸中转节点认证过滤的虚假数据.基站认证与中转节点认证的方式不同,具体认证步骤如下:

Step 1. 比较数据包中的 T 与基站中的记录 T',若 $T \le T'$ 则认为该数据包为重复数据包,将其删除,算法结束;若 $T > T'$,则将 T' 更新为 T 的值,执行下一个步骤.

Step 2. 根据 VC 计算相应的云团认证码,$V_{c_{v_i}^{sink}} = K_{v_i,sink}(T) \bmod 2^r$,其中 $v_i \in VC$.

Step 3. 计算 $V_{total}^{sink} = V_{c_{v_1}^{sink}} \oplus V_{c_{v_2}^{sink}} \ldots \oplus V_{c_{v_t}^{sink}}$, $MAC_{S,M}(y)' = f\left(S,y,H\left(V_{total}^{sink}, E_{V_{c_q^0}}(M)\right)\right)$,则有 $MAC_{S,M}(y)' - MAC_{S,M}(y) = r_{S,0}$,如果 $r_{S,0} \in \{0,1,\ldots,2^r-1, 2^L-(2^r-1),\ldots,2^L-1\}$,则执行下一步;否则,将数据包删除,算法结束.

3 算法分析

本文提出的基于云团认证的虚假数据过滤机制,由源节点结合多个云团认证码,构造出系统认证多项式,在确保数据完整性的同时,提供了真实性、新鲜性认证;采用混淆多项式技术,通过云团内部节点协作认证,突破了 t 门限值的限制,提高了系统的安全性能,而且多项式运算,与基于公钥数字签名技术相比,计算开销较低.另外,本文提出的机制允许同一个节点沿不同的路径传输连续数据包,无需建立和维护关联路径,节省了能耗,增强了机制的可用性.

3.1 安全性分析

本节通过计算俘获云团认证多项式所需的时间复杂度和攻破系统认证机制的概率上界,分析了系统的抗俘获能力;并且,用虚假数据传输距离的期望值描绘了算法对虚假数据的过滤能力.

3.1.1 攻击时间复杂度和概率分析

基于云团认证的虚假数据过滤机制具有良好的抵抗攻击的能力.攻击者想要伪造数据包,关键在于获得 t 个有效的实时云团认证码.由于初始化阶段已经将最初的云团多项式 $K_{c_i}(x)$ 删除,认证码生成使用的是混淆后的云团多项式,根据第 2.2.1 节中的分析,即使攻击者获得了 $n \ge d_w+1$ 个 $\left(x_j, K_{c_i}^r(x_j)\right)$,成功破解了当前的云团认证码 $K_{c_i}^r(x)$,也无法获得下一时段有效的云团认证码,所以攻击者只能通过猜测来破解 $K_{c_i}(x)$.另一方面,攻击者通过俘获云团内部节点 p_c 以及该节点的 $\mu+1$ 个邻居节点,协作生成实时混淆多项式 $K_{c_i,u}^r(T') = e_u(T',u) \oplus K_{c_i,u}(T')$,也可以获得有效的云团认证码.下面,我们分别对这两种攻击进行分析.

由于 $K_{C_i}(x) = \sum_{j=0}^{d_w} D_j x^j$,破解 $K_{C_i}(x)$ 即求出系数向量 D_j.攻击者可以选取 d_w+1 个不同的随机数 $\sigma_{x_j} \in \{0,1,\ldots,2^{r-1}\}(j=0,1,\ldots,d_w)$,结合 d_w+1 个 $\left(x_j, K_{c_i}^r(x_j)\right)$,联立线性方程组: $\sum_{j=0}^{d_w} D_j^r x^j = K_{c_i}^r(x_j) \oplus \sigma_{x_j}$, $j=0,1,\ldots,d_w$.求系数向量 D_j 分两步,首先需要求出 D_j^r,然后除去 σ_{x_j} 对 D_j 的干扰.线性方程组有 d_w+1 个未知数,所以求得 D_j^r 的时间复杂度为 $\Omega\left(2^{d_w+1}\right)$;由于 σ_{x_j} 是长度为 r 的随机数,已知 D_j^r 和相应的 σ_{x_j},要确定 D_j,其时间复杂度为 $\Omega(2^r)$;所以总的时间复杂度为 $\Omega\left(2^{r\times(d_w+1)}\right)$.如果选取常用参数 $r=25$, $d_w=5$,则时间复杂度为 2^{125}.

如果攻击者得到的是 $n \ge d_w+1$ 个 $\left(x_j, V_{C_i}(x_j)\right)$,则攻破 $K_{C_i}(x)$ 的复杂度同样为 $\Omega\left(2^{r\times(d_w+1)}\right)$.因为攻击者可以选取 d_w+1 个不同的随机数 $\sigma_{x_j} \in \{0,1,\ldots,2^{r-1}\}, (j=0,1,\ldots,d_w)$,与 d_w+1 个 $\left(x_j, V_{C_i}(x_j)\right)$ 串联,联立线性方程组: $\sum_{j=0}^{d_w} D_j^r x^j = V_{C_i}(x_j) \| \sigma_{x_j}$, $x=0,1,\ldots,d_w$.计算出 D_j^r,再除去 σ_{x_j} 对 D_j 的干扰,其时间复杂度为 $\Omega\left(2^{r\times(d_w+1)}\right)$.

如果攻击者通过俘获云团内部节点 p_c 以及该节点的 $\mu+1$ 个邻居节点,就可以得到

$\left\{\left\langle i,e^r_{p_c,i}(T)\right\rangle, i=0,1,...,\mu\right\}$,联立方程组解出 $e_{p_c}(x)$,计算 $K^r_{c_i,p_c}(T)=e_{p_c}(T)\oplus K_{c_i,p_c}(T)'$,从而计算出连续有效的云团认证码 $V_{c_i}=K^r_{c_i,p_c}(T)\bmod 2^r$.但攻击者需要在同一个云团中俘获 $\mu+2$ 个节点.假设整个簇有 N_c 个节点,若节点被俘获的概率是相同的,对于随机性俘获攻击,俘获的 $\mu+2$ 个节点恰好在云团 c_i 的 n 个节点中,其概率为 $P_r=\dfrac{C_n^{\mu+2}}{C_{N_c}^{\mu+2}}=\dfrac{n!\left(N_c-(\mu+2)\right)!}{N_c!\left(n-(\mu+2)\right)!}$.

所谓系统认证机制失效,是指在持续一段时间内,攻击者发布的任意数据,都能逃脱中转节点的过滤,并且被基站接受.因此攻击者必须得到 $t+1$ 个有效地云团认证码,其中包括源节点的云团认证码,即攻击者至少要俘获 $(t+1)(\mu+2)$ 个节点.最有效的攻击是攻击者俘获的 $(t+1)(\mu+2)$ 个节点,恰好分属于 $t+1$ 个云团,且每个云团中有 $\mu+2$ 个,假设这种情况存在的概率为 P_s,则 P_s 为随机性俘获攻击中系统被攻破的概率上界.首先选择 $t+1$ 个序号各异的云团,其概率为 $m^{t+1}C_k^{t+1}$,在某个云团内部俘获 $\mu+2$ 个节点的概率为 $C_n^{\mu+2}$,若整个网络有 N 个节点,则俘获 $(t+1)(\mu+2)$ 个节点的概率为 $C_N^{(t+1)(\mu+2)}$,所以,

$$P_s=\frac{m^{t+1}C_k^{t+1}C_n^{\mu+2}}{C_N^{(t+1)(\mu+2)}}=\frac{m^{t+1}k!n!\left((t+1)(\mu+2)\right)!\left(N-(t+1)(\mu+2)\right)!}{(t+1)!(\mu+2)!\left(k-(t+1)\right)!\left(n-(\mu+1)\right)!N!}.$$

攻击者也可以采用强力攻击,$MAC_{u,m}(y)$ 最多可能有 $(d_x+1)(d_z+1)$ 项,其恰好构造出 $MAC_{u,m}(y)$ 的概率为 $1/L^{(d_x+1)(d_z+1)}$,这个概率基本上是可以忽略的.

以上分析从俘获某个云团认证多项式的时间复杂度和攻破系统认证机制的概率上界两方面,说明了无论是外部窃取还是内部俘获,要想攻破本文提出的虚假数据过滤机制是困难的.在同等情况下,当参数 t 和 μ 增大时,P_v 和 P_s 减小,攻击难度增大,系统安全性更高,但随着参数 t 和 μ 的增大,参与认证的节点数目增加,通讯开销也随之加大,所以,需要合理选取系统参数,关于这个问题将在第 4 节进一步分析和说明.

3.1.2 过滤能力分析

攻击者通过俘获节点发送虚假数据时,不仅干扰用户决策,同样也会浪费大量的通信能量,所以,要尽早地识别并且过滤掉虚假数据,以节省有限的网络资源.

定理 1. 假设虚假数据数据经过第 h 个中转节点被过滤掉,本文所提出的认证机制下,虚假数据传输距离的期望值 $E(h)=1/p_v\times\left(1-\dfrac{1}{2^{L-r-1}}\right)$.

证明:如果攻击者修改消息密文 $E_{V_{c_q^0}}(M)'\neq E_{V_{c_q^0}}(M)$,中转节点进行认证时,要判断 $verf_{u_v}\left(S,H\left(V_{total}\,\|\,E_{V_{c_q^0}}(M)'\right)\right)$,与 $MAC_{u,m}(u_v)$ 或者 $MAC_{u,m}(u_v)'$ 的差是否属于 $\left\{0,1,...,2^r-1,2^L-\left(2^r-1\right),...,2^L-1\right\}$,作为变量的消息密文为虚假数据时,该差值可能是 $\left\{0,1,...,2^L-1\right\}$ 之间的任意数,所以虚假数据通过验证的概率为 $\dfrac{2^{r+1}}{2^L}=\dfrac{1}{2^{L-r-1}}$.由于中转节点对数据包进行认证的概率为 p_v,则虚假数据被过滤的概率为 $p_{out}=p_v\times\left(1-\dfrac{1}{2^{L-r-1}}\right)$.数据经过第 h 个中转节点被过滤掉的概率为 $p_h=p_{out}\times(1-p_{out})^h$,由此可得:

$$E(h)=\lim_{k\to\infty}\left(\sum_i^k i\cdot(1-p_{out})^{i-1}\cdot p_{out}\right)=\lim_{k\to\infty}\left(p_{out}\cdot\sum_i^k i\cdot(1-p_{out})^{i-1}\right)=1/p_{out}.$$ □

由于采用基于概率的认证,该认证机制不能保证在中转路由阶段过滤掉所有的虚假数据,但随着中转跳数的增加,过滤概率也会增加,也就是说单个节点的过滤能力是有限的,但整个中转路由的节点协作过滤能力则是较强的.而且,基站通过重构认证多项式,几乎能够识别和过滤掉所有已经逃脱中转节点过滤的虚假数据.另一方面,由于认证多项式以与实时时间相关的值 T 为参数,所以本文的机制不仅能有效识别和过滤被篡改的数据,还能有效地识别和过滤重复数据包,具有一定的抵抗重放攻击的能力.

3.2 动态路由

本文提出的认证机制采用GPRS或GEAR算法选择数据的传输路径,由转发节点从邻居簇中选择与基站距离最近簇,然后在该簇中随机选取一个节点担任下一跳传输节点.由于中转节点是以一定的概率,在所属云团内部发起协作来生成相应的云团认证码,不需和参与数据包生成的节点进行交互.因此,该方案与 IHA,LEDS 等方案相比,无须建立关联路径;当某中转节点失效时,可以随机选取同一个簇内任意一个有效节点来担任中转节点,该节点将以同样的概率对数据包进行认证.因此,该认证机制无须耗费多余的能量来建立和维护关联路径,能够支持动态路由.

3.3 能耗分析

无线传感器网络资源有限,如何降低能耗是 WSN 应用中所考虑的一个重要问题.为了方便讨论,假定网络部署后形成多个单跳簇,不记成簇开销.假设整个网络的节点个数为 N,分为 m 个簇,每个簇有 k 个云团数,每个云团的平均节点个数为 n,则有 $N=mkn$,则节点 ID 长度为 $ID_{node}=\lceil \log_2 N \rceil$ bits.簇 ID 长度为 $ID_{cluster}=\lceil \log_2 m \rceil$ bits.云团 ID 长度为 $ID_{cloud}=\lceil \log_2 k \rceil$ bits.

3.3.1 通信开销

通信开销主要分为 3 部分:(1) 云团认证码生成阶段,需要交换的数据为 $l_1=L+(\mu+1)\left[(d_w+1)L+ID_{node}\right]$ bits;(2) 簇内数据包生成阶段,$l_2=2L+2t\left[(L-r)+ID_{cloud}\right]+(L-r)$ bits;(3) 网内数据包传输阶段,数据包总长度为 $l_3=ID_{node}+M+(d_y+1)L+L+t\left[(L-r)+ID_{cloud}\right]$ bits,这部分的通信开销还需要结合节点的过滤概率和传输跳数来计算.

3.3.2 计算开销

计算开销主要分为两部分:(1) 在密钥预置阶段,每个节点需要为 $n-1$ 个邻居节点分配一个 d_w 次多项式,并计算出混淆后的云团多项式 $K_{c_i,u}(t)$,这需要 $o\left(nd_w^2\right)$ 次乘法.(2) 在云团认证码生成阶段,当云核节点收集到 $\mu+1$ 个分量 $\left\{\left\langle w_i, e_{u,w_i}(T)\right\rangle, i=0,1,...,\mu\right\}$ 和 $K_{c_i,u}(t)$,需要联立方程组解出 $e_u(T,u)$,计算复杂度为 $o\left(\mu^3\right)$.

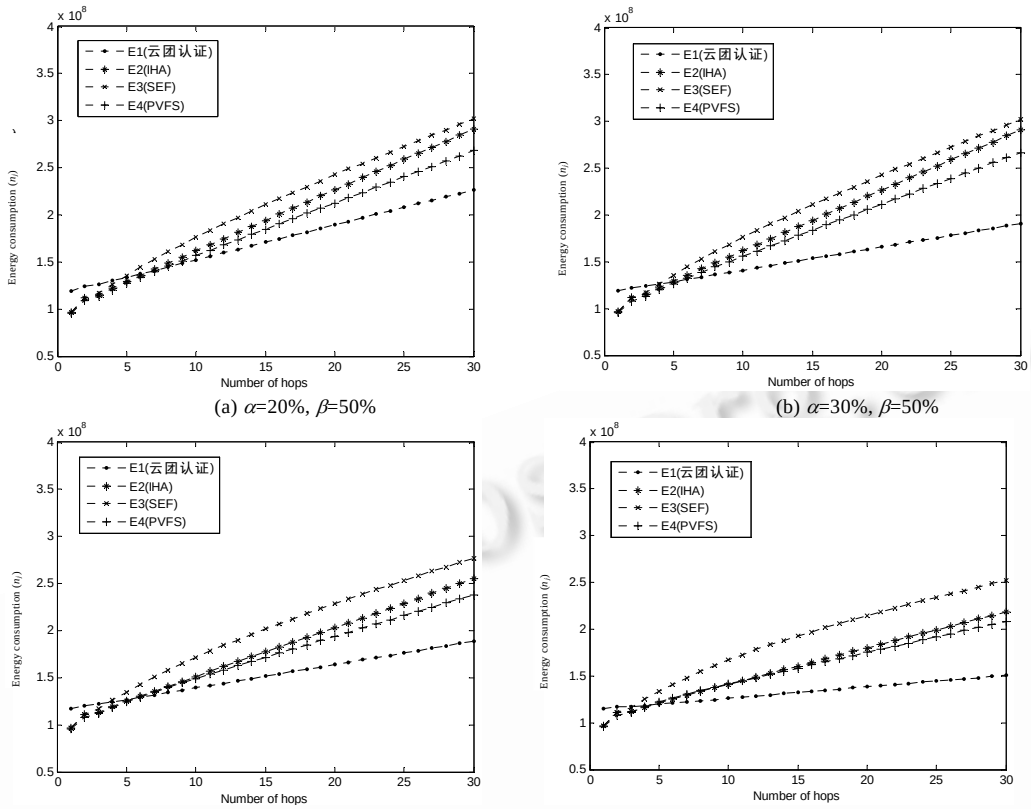
3.3.3 存储开销

每个节点需要存储自身 ID_{node},所属云团 ID_{cloud} 和簇 $ID_{cluster}$,验证函数和验证函数 $(d_y d_z+1)L+(d_x d_z+1)L$ bits,云团秘密数 r,一个 d_w 次多项式 $K_{c_i,u}(t)$,$n-1$ 个邻居节点的分量,所需存储开销为 $n(d_w+1)L$ bits.

4 仿真实验

为了评价和分析本文提出的云团认证过滤机制的性能,本节在 MATLAB 平台上,将该机制与具有代表性的 IHA,SEF 和 PVFS 在能量开销方面进行了模拟仿真;并通过对比实验,讨论了该机制的参数的选取问题.在一个 $300\times300\text{m}^2$ 的检测范围内节点随机分布,参数 $N=5000$, $m=50$, $k=10$.

假设源节点发送 100 个数据包,其中重复包占的百分比为 α,虚假数据包占的百分比为 β,每个数据包需要征集 4 个认证信息,则云团内部参与协作的节点个数为 5;另外按照文献[8]将密钥池的大小设为 5 000,每个节点选取的密钥数为 200,按照文献[10]路径认证的平均概率为 0.65,假定攻击者俘获了 2 个有效的认证密钥.L 取 32 bits,r 取 25 bits,dw 和 dy 均取 5 bits.忽略计算开销,通信开销用文献[25]提出的一阶无线模型(first order radio model)进行估算,其中发送数据的能耗参数 $\varepsilon_{amp}=100\text{pJ/bit/m}^2$,接收数据的能耗参数 $E_{elec}=50\text{nJ/bit}$,节点通信距离 $d=45\text{m}$.



(a) α=20%, β=50%　　(b) α=30%, β=50%

(c) α=20%, β=60%　　(d) α=20%, β=70%

Fig.3　Energy consumption comparison

图 3　本文的算法和文献[7,8,10]的能耗对比

图 3 描绘了本文提出的云团认证过滤机制和 SEF,IHA,PVFS 机制在能耗方面的差异.结果表明,本文提出的机制在前几跳传输时,能耗略高于另外 3 种算法;大约从 4 跳左右开始,比其余算法能耗要低,并且随着跳数的增加,优势逐渐明显.主要原因是:基于云团的认证机制在数据包形成的过程中需要更多的节点参与认证,带来了较大的通信开销.但另一方面,在最终的数据包中,云团认证过滤机制用较短的云团序号来代替较长的节点 ID 来标识认证信息,减少了需要传输的数据长度,这部分能耗随着数据的传输距离增大成线性增长;而且,本文提出的算法能够识别和过滤重复数据包,其他 3 种机制都没有这种功能,比较图 3(a)、图 3(b),在同等情况下,随着重复数据包所占比例的增加,基于云团认证的机制能耗进一步降低,而其他 3 种机制并无变化.由此推断,本文提出的算法具有抗重放功能.另外,由图 3(a)、图 3(c)、图 3(d)比较发现,在同等情况下,随着虚假数据比例的增加,基于云团认证的机制优势更加明显.综上所述,本文提出的认证机制在网络可信度较低,即虚假数据所占比例较大的网络,以及远距离传输场景中比较适用.

参数的选取将直接影响到系统性能,因此,如何选择合适的参数,至关重要.我们通过对比实验,主要讨论选取不同的参数 t 和 μ 对系统性能的影响.

假设源节点发送了 100 个数据包,其中 α =0, β =90%,图 4 描绘了选取不同的参数 t,系统能耗的差异.随着 t 的增大,参与认证的节点数量增加,数据包形成过程中的能耗也相应增加,所以在前几跳中,能耗随 t 值的增大而增加.另一方面,t 增大时,系统的过滤能力增强,虚假数据被更早过滤掉,反而节省了能量,所以随着跳数的增加,t 值较大的曲线增长速度反而比 t 值较小的曲线要慢.因此,我们可以根据具体的应用选取合适的 t 值.

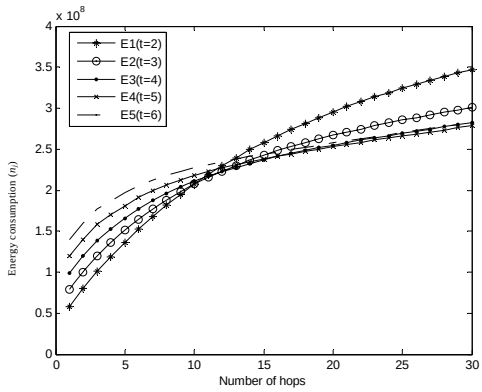


Fig.4 Different system consumption under the different t

图 4 不同参数 t 下的系统能耗

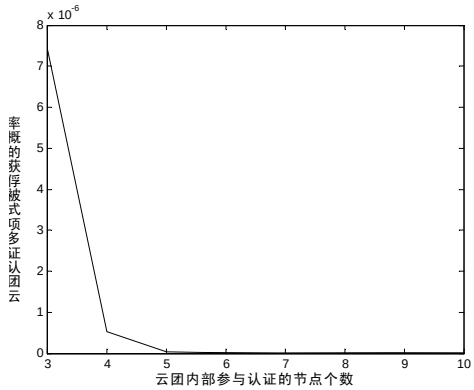


Fig.5 Different anti-trapping ability under the different μ

图 5 不同参数 μ 下的抗俘获能力

图 5 描绘了就某个云团而言,不同的 μ 值对应的云团认证多项式被俘获的概率 P_r.如图所示,当 μ 取 2,即参与认证的节点个数为 4 时,攻击者俘获认证多项式的概率低于 10^{-6},当 μ 取 3,即参与认证的节点个数为 5 时,攻击者俘获认证多项式的概率低于 10^{-7}.而对于整个系统而言, P_s 趋近于 0,因此,本文提出的认证机制能够不受t门限值的限制.

5 结 论

本文提出了一种适用于无线传感网络的云团认证算法,为传感数据的真实性,完整性,新鲜性提供保证.在此基础上,提出了一种虚假数据过滤机制,与已有机制相比,该机制更适用于可信度较低的网络,以及远距离传输场景的应用.主要特点表现在:(1) 采用混淆多项式技术,由云团内部多节点协作认证,与已有基于 MAC 的认证算法相比,不受 t 门限值的限制;(2) 允许源节点沿不同的路径发送连续数据包,支持动态路由.另外,本文提出的认证机制不仅能有效地识别和过滤虚假数据包,还具有一定的抵抗重放攻击的能力.

References:

[1] 任丰原,黄海宁,林闯.无线传感器网络.软件学报,2003,14(7):1282−1291 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/14/1282.htm

[2] 李建中,李金宝,石胜飞.传感器网络及其数据管理的概念、问题与进展.软件学报,2003,14(10):1717−1722 (in Chinese with English abstract). http://www.jos.org.cn/ 1000-9825/14/1717.htm

[3] 裴庆祺,沈玉龙,马建峰.无线传感器网络安全技术综述.通信学报,2007,28(8):113−122 (in Chinese with English abstract).

[4] 崔莉,鞠海玲,苗勇.无线传感器网络研究进展.计算机研究与发展,2005,17(3):163−174 (in Chinese with English abstract).

[5] Li P, Lin YP, Zeng WN. Search on security in sensor networks. Journal of Software, 2006,17(12):2577−2588 (in English with Chinese abstract). http://www.jos.org.cn/ch/reader/view_abstract.aspx?flag=1&file_no=20061217&journal_id=jos

[6] Ye F, Yong G. A dynamic en-route scheme for filtering false data injection in sensor networks. In: Proc. of the IEEE INFOCOM 2006. 2006. 1−12.

[7] Zhu S, Setia S, Jajodia S. An interleaved hop-by-hop authentication scheme for filtering of injected false data in sensor networks. In: Proc. of the IEEE SSP 2004. 2004. 259−271.

[8] Ye F, Luo H, Lu S, *et al*. Statistical en-route filtering of injected false data in sensor networks. In: Proc. of the IEEE INFOCOM 2004. 2004. 2446−2457.

[9] Ye F, Luo H, Lu S, *et al*. Statistical en-route filtering of injected false data in sensor networks. IEEE Journal on Selected Areas in Communication, 2005,23(4):839−850.

[10] Li F, W J. A probabilistic voting-based filtering scheme. In: Proc. of the IEEE IWCMC 2006.2006. 255−265.

[11] Ma M. Resilience of sink filtering scheme in wireless sensor networks. Computer Communications, 2006,30(1):55−65.

[12] Malan DJ, Welsh M, Smith MD. A public-key infrastructure for key distribution in Tinyos based on elliptic curve cryptography. In: Proc. of the IEEE SECON 2004. 2004. 71−80.

[13] Watro R, Kong D, Cuti S, *et al*. Tinypk: Securing sensor networks with public key technology. In: Proc. of the IEEE SASN 2004. 2004. 59−64.

[14] Du W, Wang R, Ning P. An efficient scheme for authenticating public keys in sensor networks. In: Proc. of the IEEE MOBIHOC 2005. 2005. 58−67.

[15] Wang H, Li Q. PDF: A public-key based false data filtering scheme in sensor networks. In: Proc. of the IEEE WASA 2007. 2007. 129−138.

[16] Zhang W, Subramanian N. Lightweight and compromise- resilient message authentication in sensor networks. In: Proc. of the IEEE INFOCOM 2008. 2008. 1418−1426.

[17] 周四望,林亚平,张建明,欧阳竞成,卢新国.传感器网络中基于环模型的小波数据压缩算法.软件学报,2007,18(3):669−680 (in Chinese with English abstract). http://www.jos.org.cn/ch/reader/view_abstract.aspx?flag=1&file_no=20070320&journal_id=jos

[18] Liu D, Ning P. Establishing pairwise keys in distributed sensor networks. In: Proc. of the ACM CCS 2003. 2003. 52−61.

[19] Du W, Deng J. A pairwise key pre-distribution schemes for wireless sensor networks. In: Proc. of the ACM CCS 2003. 2003. 42−51.

[20] Karp B, Kung HT. GPSR: Greedy perimeter stateless routing for wireless networks. In: Proc. of the IEEE MOBIHOC 2000. 2000. 243−254.

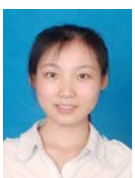
[21] Yu Y, Govindan R, Estrin D. Geographical and energy aware routing: A recursive data dissemination protocol for wireless sensor networks. Technical Report, UCLA-CSD TR-01-0023, UCLA Computer Science Department, 2001.

[22] Ren K, Lou W, Zhang Y. LEDS: Providing location-aware end-to-end data security in wireless sensor networks. In: Proc. of the IEEE INFOCOM 2006. 2006. 585−598.

[23] Zhang W, Tran M, Zhu S, *et al*. A random perturbation-based scheme for pairwise key establishment in sensor network. In: Proc. of the IEEE MOBIHOC 2007. 2007. 90−99.

[24] Zhang W, Cao G. Group rekeying for filtering false data in sensor networks: A predistribution and local collaboration-based approach. In: Proc. of the IEEE INFOCOM 2005. 2005. 503−514.

[25] Heinzelman WR, Chandrakasan A, Balakrishnan H. Energy-Efficient communication protocol for wireless microsensor networks. In: Proc. of the 33rd Annual Hawaii Int'l Conf. on System Sciences. 2000. 3005−3014.



彭舸(1985－),女,湖南湘乡人,硕士生,主要研究领域为无线传感器网络安全.

易叶青(1976－),男,博士生,主要研究领域为数字水印,无线传感器网络安全.

林亚平(1955－),男,博士,教授,博士生导师,主要研究领域为计算机网络,机器学习.