

无线传感器网络中的分簇安全数据汇聚方案

刘 冰, 马 壮, 陈宜栋, 李艳俊

(北京电子科技学院 密码科学与技术系, 北京 100070)

摘 要: 为在无线传感器网络(WSN)的数据收集、处理和传输过程中降低能耗,提高数据的安全性与机密性,提出一种动态分簇安全数据汇聚算法 CDSDA。对基于分簇的 WSN 安全数据汇聚算法 CSDA 进行改进,依据节点所采集数据的重要程度与节点的剩余能量,对数据进行动态切片,在此基础上,进行簇内汇聚和簇间汇聚以得到最终结果。仿真结果表明,与 CSDA 算法相比,CDSDA 算法在通信开销、数据正确率、数据隐私保护等方面具有较好性能。

关键词: 无线传感器网络; 分簇结构; 数据汇聚; 动态切片; 隐私保护

中文引用格式: 刘冰,马壮,陈宜栋,等. 无线传感器网络中的分簇安全数据汇聚方案[J]. 计算机工程,2019,45(9): 136-142.

英文引用格式: LIU Bing, MA Zhuang, CHEN Yidong, et al. Clustering secure data aggregation scheme in wireless sensor networks[J]. Computer Engineering, 2019, 45(9): 136-142.

Clustering Secure Data Aggregation Scheme in Wireless Sensor Networks

LIU Bing, MA Zhuang, CHEN Yidong, LI Yanjun

(Department of Cryptography Science and Technology, Beijing Electronic Science and Technology Institute, Beijing 100070, China)

[Abstract] In order to reduce energy consumption and improve data security and confidentiality in the process of data collection, processing and transmission in Wireless Sensor Networks(WSN), a dynamic clustering secure data aggregation algorithm CDSDA is proposed. The clustering-based WSN secure data aggregation algorithm CSDA is improved. According to the importance of data collected by nodes and the residual energy of nodes, the data is sliced dynamically. On this basis, intra-cluster aggregation and inter-cluster aggregation are carried out to get the final results. Simulation results show that, compared with CSDA algorithm, CDSDA algorithm has better performance in communication overhead, data accuracy, data privacy protection and so on.

[Key words] Wireless Sensor Networks(WSN); clustering structure; data aggregation; dynamic slicing; privacy protection

DOI: 10.19678/j.issn.1000-3428.0051776

0 概述

无线传感器网络(Wireless Sensor Networks, WSN)由若干能够收集各种环境信息的传感器节点组成,目前已广泛应用于军事、救灾、医疗等领域^[1]。由于 WSN 中的多数传感器节点计算能力较低,存储空间较小,且电池容量受限,因此对传感器节点的能量进行有效利用一直备受关注。

通过数据汇聚技术能够有效降低传感器网络的能耗,在该技术下,多个成员节点感知的数据通过应用求和、求均值和取最大值等汇聚函数,最终聚合成一个单独的数据进行传输^[2],进而降低能量开销。

然而,WSN 通常部署在恶劣环境中,且多数传

输数据具有敏感性,经常遭受各种网络攻击。一旦节点遭到攻击,数据就可能暴露给攻击者并被其修改,此时将无法保证 WSN 中传输数据的机密性和完整性。为对传感器节点的能量进行有效利用,并解决数据在传输和汇聚过程中的安全问题,研究人员提出了大量隐私保护数据汇聚方案。

文献[3]提出 CPDA(Cluster-based Private Data Aggregation)和 SMART(Slice-Mix-AggRegaTe)2 种隐私保护数据汇聚方案。在 CPDA 方案中,传感器节点随机形成簇,在每个簇内,利用多项式的代数性质来计算所需的聚合值,同时保证单个节点无法获取其他节点的数据值。SMART 是多级数据汇聚方案,每个簇内的中间集合值将沿着汇聚树的方向进一步

基金项目: 国家自然科学基金(61370188);中央高校基本科研业务费专项资金(2017LG04)。

作者简介: 刘 冰(1976—),男,讲师、博士,主研方向为密码学、信息安全;马 壮、陈宜栋,硕士研究生;李艳俊,副教授。

收稿日期: 2018-06-08 **修回日期:** 2018-08-12 **E-mail:** bing@besti.edu.cn

聚合。SMART 方案共分为 3 步:首先,网络中每个节点 s_i ($i=1,2,\dots,N$) 在 h 跳内随机选择 $J-1$ 个邻居节点构成节点集 s_i ,将感知数据 d_i 随机切分为 J 个数据切片, s_i 为本节点留下其中 1 个数据切片,将 $J-1$ 个数据切片进行加密后随机发送至其他节点 s_j ,用 d_{ij} 表示由节点 s_i 传送到 s_j 的数据切片;然后,每个节点 s_j 对收到的数据进行解密并求和,得到 $r_j = \sum_{i=1}^N d_{ij}$;最后,所有节点 s_i ($i=1,2,\dots,N$) 使用树形路由将计算的 r_j 送至基站,基站对所有 r_j 求和得到汇聚结果 $\sum_{j=1}^N r_j$ 。每个节点通过将其采集的数据进行分片来隐藏其私有数据,并将加密数据分片发送到不同的中间汇聚节点,接收到切片的中间汇聚节点计算中间汇聚值并将它们传输至接收器。

文献[4]建立一种基于切片和混合技术的平衡隐私保护数据汇聚模型 BPDA。该模型基于树状结构网络,首先从节点的度数方面设计平衡数据切分原则,设定一个阈值,通过比较节点的度数与阈值来决定是否进行数据切分。若节点的度数小于阈值,则进行切分并将切分数据传送到相邻节点,同时度数加 1,反之,只接收其他节点的切分数据或不执行任何操作直至本轮结束,同时度数也加 1。然后,在此模型的基础上,分别从能量、度数、能量和度数 3 个方面提出相应算法,得出同时考虑能量与度数进行平衡切片的方法,该方法能够提高隐私保护的效率并平衡能耗。

文献[5]提出一种基于树状网络结构的动态数据切分重组算法 D-SMART,该算法根据传感器节点感知变量之间的偏差程度和平均值大小来确定感知数据的重要程度,并根据该重要程度将感知数据分为 3 个不同的等级,然后进行动态数据切片,越重要的数据将被分割得越多,最后将切片发送至相邻节点。D-SMART 算法根据原始数据的重要程度,通过合理的切片实现了良好的隐私保护性能。该算法改善了 SMART 算法在数据传输和聚合精度方面的不足,优化了树聚合网络的构建方式,同时能够提高数据隐私保护的程度,降低传感器节点的通信成本,延长网络寿命。

文献[6]提出一种新型的节能安全数据汇聚方案 CSDA,该方案是以分簇隐私保护为基础的分簇私有数据汇聚方案,其使用切片重组的方法,根据 WSN 的规模和数量动态调整簇分区的节点数量,同时添加了入侵检测功能。CSDA 能够满足一般的隐私保护要求,不仅具有较好的灵活性和数据汇聚精度,而且降低了通信开销和能耗。但是,CSDA 的切片数量随着簇内节点数量的增加而线性增加,在数据重要性较低时会造成不必要的能量消耗。此外,单一依靠节点数量来决定数据的切片数,无法保证隐私需求较高的数据的安全性。

针对上述问题,本文对 CSDA 方案进行改进,提出一种基于 D-SMART 算法的分簇安全数据汇聚方案 CDSDA。根据所采集数据的重要程度以及节点剩余能量的变化,对数据进行动态切片,在此基础上实现分簇安全数据汇聚,以提高数据隐私保护的程度并降低通信开销。

1 基于 D-SMART 算法的分簇安全数据汇聚

动态切片重组技术指在对节点原始数据进行处理时,通过一定的规则动态地对数据进行切分重组。上述过程中的规则称为动态切片重组技术规则,简称动态规则。在实际应用中,存在多种动态规则,如:根据原始数据的重要程度、节点的剩余能量、簇内节点的数量对原始数据进行动态切分重组等。通过所收集数据的重要程度将数据切分为不同数量的切片,能更大程度地提升有效数据的正确率,同时也更有效地保护了数据隐私;根据节点剩余能量进行动态切分重组的规则能更好地保证网络节点的活力,延长整个网络的寿命;根据簇内节点数量对原始数据进行动态切分重组的规则,能够充分利用簇内分片的优势,有较高的安全性,但其灵活性较低。

1.1 CDSDA 算法的基本思想

在 CSDA 算法中,采取切片数随簇内节点数的增加而线性增加的动态切分重组规则。但该算法存在一定弊端,如:由于其切片数随簇内节点数的增加而增加,若簇内节点足够多,切片数会很大,此时将发送大量的数据切片,增加了节点间的流量最终降低节点的工作寿命。同时,随着数据传输量的增加,数据传输时发生碰撞、延迟、错误的概率将上升,从而影响数据汇聚的结果和效率^[6]。

为保持 CSDA 算法原有的准确性并进一步提高其安全性和节能效果,本文提出一种 WSN 中基于 D-SMART 算法的安全分簇数据汇聚算法 CDSDA。采用类似 D-SMART 算法中的节点剩余能量动态切分重组规则来代替 CSDA 方案中以簇内节点数量为依据的动态切分重组规则。D-SMART 算法根据隐私数据的重要程度将数据分为 3 类,重要的数据将会进行较多切分,不重要的数据将会进行较少切分,该方式能更大程度地提升有效数据的正确率,节约传感器节点的能耗,提高数据隐私保护的效果。CDSDA 算法同时考虑节点剩余能量以对原始数据进行动态切分重组。初始阶段节点能量充足,为充分保证数据汇聚过程中的安全性,可以将数据进行较多切片;当节点能量较少时,为更好地保证汇聚的完整性,可以适当减少切片数。这种动态切分重组的方式,能更好地提高网络节点的活力,延长整个网络的寿命。在 CDSDA 算法中,切片数 J 的确定方法如下:

当 $E_r > 30\%$:

$$J = \begin{cases} 2, & \text{重要程度为低} \\ 4, & \text{重要程度为中} \\ 6, & \text{重要程度为高} \end{cases} \quad (1)$$

当 $E_r \leq 30\%$:

$$J = \begin{cases} 1, & \text{重要程度为低} \\ 2, & \text{重要程度为中} \\ 4, & \text{重要程度为高} \end{cases} \quad (2)$$

其中, E_r 表示节点的剩余能量占总能量的比值。

上述切片数确定方法能满足多种应用场景的需求。如:在以手机为传感器节点、专门的收集监测后台为收集池的应用中,将手机节点收集的数据信息分为 3 类,重要程度低的信息包括“年龄”“身高”“体重”;重要程度中等的信息包括“邮箱”“微信号”“QQ 号”;重要程度高的信息包括“姓名”“身份证号”“手机号”。节点根据自身收集的原始数据进行识别并判断数据信息的敏感程度,再根据具体的动态切分原则将数据进行动态切分。当节点收集的原始数据为“年龄”时,则将该数据动态切分为 2 片;若收集到“邮箱”数据时,则将该数据动态切分为 4 片;若收集到“姓名”数据时,则将该数据动态切分为 6 片。经过切片后的数据被加密传输至簇头,由簇头进行簇内的数据汇聚,然后进行簇间的数据汇聚,将最终汇聚结果传输至收集池。

1.2 CSDA 算法实现

1.2.1 网络模型

WSN 以连通图 $G(V, E)$ 进行建模,其中, V 是传感器节点集合, $|V| = N$ 是传感器节点的数量, E 是连接传感器节点的无线链路集合。每个传感器节点都配备一个无线通信模块,可以与传输范围内的其他传感器节点进行通信^[7]。同时,本文采用分簇结构的网络模型,传感器节点进行网络数据汇聚时被分为 3 类:普通节点(簇内节点),汇聚节点(簇头),收集池。普通节点收集原始数据并根据原始数据的重要程度进行切片,汇聚节点通过合适的函数对切片进行汇聚,收集池负责接收汇聚结果、转发请求、连接整个网络。节点的数据汇聚过程如图 1 所示。

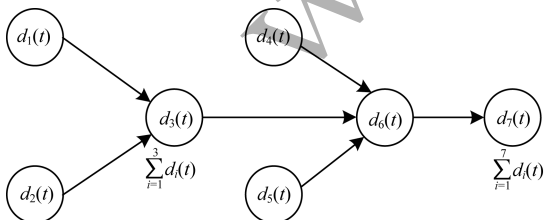


图 1 数据汇聚过程示意图

汇聚函数表示每个节点在时间 t 时的监测数据^[8],其定义为:

$$y(t) = f(d_1(t) + d_2(t) + \dots + d_N(t))$$

汇聚函数通常包括求和、求平均值、求最大值、求最小值和计数等。其中,求和函数较为重要,一些统计函数(如计数、平均值和标准差等)都是基于求和函数,且对于非线性函数,如最大值和最小值,在一定程度上也可以用求和函数进行估计。因此,本文选择求和函数作为汇聚函数。求和函数计算表达式如下:

$$y(t) = \sum_{i=1}^N d_i(t) \quad (3)$$

1.2.2 攻击模型

在 WSN 中,攻击者可以采用多种不同的攻击方式来破坏数据的隐私。本文主要考虑窃听攻击和妥协攻击对 WSN 中数据隐私保护的威胁^[9]。

在窃听攻击中,攻击者可以通过链路层窃听节点与节点之间传输的隐私数据,从而破坏数据的机密性。在妥协攻击中,攻击者通过妥协某一节点或多个节点获取其数据与密钥,进而利用密钥得到其他节点的隐私数据,或通过联合多个受损节点推断邻居节点的隐私数据,最终对整个网络构成极大威胁。

1.2.3 加密密钥分发

为避免遭受窃听攻击和妥协攻击,在节点间传输的数据必须通过一定的密钥进行加密,本文采取与 SMART、CPDA、CSDA 相类似的加密密钥分发机制。密钥分发由 3 个阶段组成^[10]:

1) 密钥预分配:首先,随机生成一个 K 密钥的大型密钥池及其相应的身份标识;然后,网络中的每个节点从密钥池中随机抽取 k 个密钥,构成传感器节点的密钥环。

2) 共享密钥发现:每个传感器节点通过交换发现消息来找出与自己共享公共密钥的邻居节点。如果 2 个相邻节点共享一个公共密钥,则它们之间存在一个安全链路。

3) 路径密钥建立:将路径密钥分配给不能共享公共密钥,但可在共享密钥发现阶段结束时通过 2 个或 2 个以上多跳安全链路连接的相邻传感器节点对。

随机图论理论^[11]指出:设随机图的节点数量为 n ,当 n 很大时,如果整个随机图至少要以概率 P_r 成为连通图,则存在以下极限式:

$$P_r = \lim_{n \rightarrow \infty} P(G(n, p)_{\text{connected}}) = e^{e^{-C}} \quad (4)$$

其中, P 为任意 2 点连通的概率, C 为指定的门限值,且存在关系式:

$$p = \frac{\ln n + C}{n} \quad (5)$$

当给定 WSN 节点数量 n 、期望的整体连通概率 P_r 时,可求得任意 2 点连通的概率 P ,进而求得节点的平均度 d 。综合式(4)、式(5),有:

$$p = (\ln n - \ln(\ln P_r)) / n \quad (6)$$

$$d = p \times (n - 1) \quad (7)$$

安全连通概率 P' 是节点与其通信范围内的邻居节点之间至少存在一个共享密钥的概率,而 p 是密钥共享图中任意 2 点之间的连通概率,使用传感器网络节点的分布密度作为中介,可将上述两者建立联系。设传感器网络的部署密度 n' 为任意节点的邻居节点数量(对于均匀分布, n' 一定,且 $n' \ll n$),根据节点数量很大时平均度 d 存在极限的事实,在上述随机密钥分发机制中,任意一对节点间至少存在一个公钥的概率为:

$$P' = \frac{d}{n' - 1} = 1 - P'' = 1 - \frac{((K - k)!)^2}{(K - 2k)! K!} \quad (8)$$

其中, P'' 为节点间无公钥的概率, K 为密钥分发中心产生的密钥池大小, k 为每个节点密钥环的大小。

采用随机密钥分发机制时,随机选择密钥池中 K 密钥的第三方窃听节点,如果 K 密钥中包含通信节点选定的密钥,则攻击者可以使用该密钥偷听加密消息。攻击者获得密钥的概率为^[12]:

$$P_{\text{overhear}} = \frac{k}{K} \quad (9)$$

1.2.4 安全汇聚算法

CDSDA 算法流程如图 2 所示。

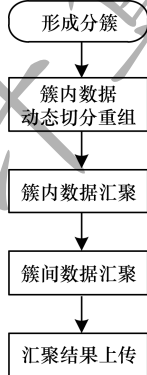


图2 CDSDA 算法流程

CDSDA 算法具体步骤为:

1) 形成分簇。传感器网络按照一定规则选取若干节点作为簇头,其余节点分配至与簇头最近的簇内,完成分簇,然后根据网络情况进行动态调整。为保证数据的汇聚效率,应尽可能使具有高相关性的节点被分配在同一簇内^[13-15]。

2) 簇内数据动态切分重组。首先,簇内节点根据所收集数据的具体内容判别数据的重要程度;然后,根据数据的重要程度进行数据的动态切分;接着,将切片数据用所分配的密钥进行加密,并传输至相邻节点;最后,节点将自身保留的数据切片和接收到相邻节点的数据切片进行混合重组,并传输至簇头。

如图 3、图 4 所示,簇内有 6 个节点,1 代表簇头

节点, A 、 B 、 C 、 D 、 E 代表成员节点,节点 A 和 B 收集的数据是一般数据,故将各自的原始数据切分为 (a_1, a_2) 、 (b_1, b_2) ;节点 C 和 D 收集的数据是重要数据,故将各自的原始数据切分为 (c_1, c_2, c_3, c_4) 、 (d_1, d_2, d_3, d_4) ;节点 E 收集的数据是非常重要的数据,故将其原始数据切分为 $(e_1, e_2, e_3, e_4, e_5, e_6)$ 。在切分完成后,再在簇内将切片进行加密传输并重组,用 M_i 表示节点 i 将来自簇内其他节点的若干切片进行混合重组后得到的新数据包,最后将每个节点混合重组的结果传输至簇头进行簇内汇聚。

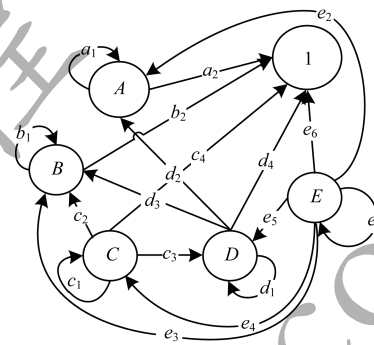


图3 数据切片过程

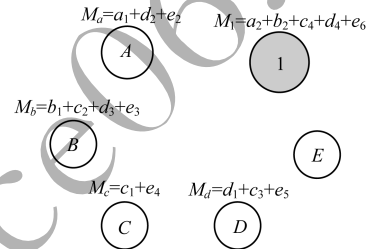


图4 数据重组过程

3) 簇内数据汇聚。每个簇的簇头收到簇内节点传输来的重组数据后,先对重组数据进行解密,接着对切片实现数据汇聚并加密,用于簇间数据汇聚。如图 5 所示,根据式(3),得簇头节点 1 所获的数据汇聚值为 $y_1(t) = \sum_{i=1}^N d_i(t) = d_A(t) + d_B(t) + d_C(t) + d_D(t) + d_E(t)$ 。此时,簇头节点用于数据汇聚,而簇内其他成员节点负责监视簇头节点的操作。

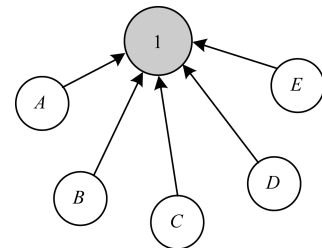


图5 簇内汇聚示意图

4) 簇间数据汇聚。在每个簇头节点对簇内数据进行汇聚后,将其结果根据式(3)进行簇间数据汇聚,且高级簇头汇聚来自低级簇头的的数据。如

图 6 所示,簇头节点 4 的汇聚结果为 $y_4(t) = \sum_{i=1}^N d_i(t) = d_1(t) + d_2(t) + d_3(t)$ 。

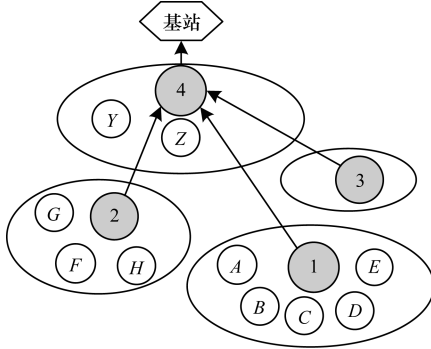


图 6 簇间汇聚示意图

5) 汇聚结果上传。在完成簇间汇聚操作后,对汇聚结果进行加密并传递给基站。基站通过共享密钥解密这些汇聚数据并获得最终的数据汇聚结果。

2 算法分析

CSDSA 算法结合了 CSDA 和 D-SMART 算法,将分簇和以数据重要程度、节点剩余能量为标准的动态切分重组规则相结合,既能基于有效的分簇结构来实现数据的汇聚,又能避免 CSDA 中根据节点数量进行动态切分重组带来的弊端。CSDSA 算法最大的优势在于,其能根据所收集数据的重要程度动态地调整数据切分的数量,从而保证数据切分的合理性。这种合理性体现在数据切分所消耗的能量和数据本身安全性之间的权衡,即在保证数据安全性的同时降低数据冗余和碰撞等现象发生的概率。

本文针对 WSN 进行实验,分析比较 CSDA 算法和 CSDSA 算法在隐私保护、通信开销、数据正确率 3 种性能上的差异。在仿真中,100 个传感器节点随机分布在 1×1 的范围内,通信半径为 0.1,其他网络参数本文暂不考虑。另外,2 种对比算法都在相同的网络拓扑结构上运行。初始网络分簇拓扑如图 7 所示。

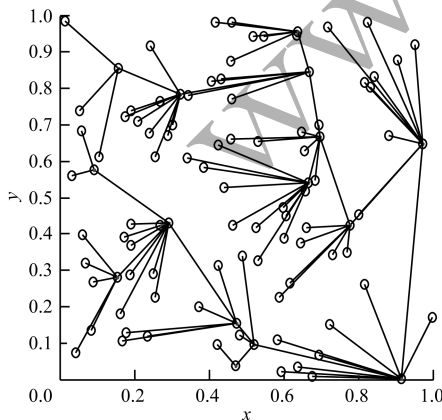


图 7 网络分簇拓扑

2.1 隐私保护分析

在攻击模型中,攻击者必须打破数据切片在节点中传输时的所有出口链接和所有入口链接,才能彻底损害数据的隐私。在 CSDA 算法中,节点 S 向外发送的切片数为 $J-1$ (出口链接数)^[16],但其接收相邻节点的数据切片(入口链接数)不确定,因此,该算法的节点隐私暴露率 $P_{\text{CSDA}}(q)$ 定义为:

$$P_{\text{CSDA}}(q) = q^{J-1} \sum_{k=0}^{d_{\text{in-max}}} p(\text{in-degree} = k) q^k \quad (10)$$

其中, $d_{\text{in-max}}$ 表示节点入口链接数的最大值, $p(\text{in-degree} = k)$ 表示节点 S 入口链接数为 k 时的概率, q 表示节点间链接的破解概率,且 $q \approx P_{\text{overhear}}$ 。

在 CSDSA 算法中,节点切片数随数据重要程度动态变化, J 表示节点生成的最大片数, j 表示由叶节点生成的实际片数。节点 S 向外发送的切片数和接收其他节点的切片数均不确定,故其平均节点隐私暴露率 $P_{\text{CSDSA}}(q, j)$ 定义为:

$$P_{\text{CSDSA}}(q, j) = \sum_{j=2}^J p(\text{out-degree} = j-1) q^{j-1} \cdot \sum_{k=0}^{d_{\text{in-max}}} p(\text{in-degree} = k) q^k \quad (11)$$

其中, $p(\text{out-degree} = j-1)$ 表示节点 S 出口链接数为 $j-1$ 的概率。

根据式 (10) 和式 (11) 得到 CSDA 和 CSDSA 2 种算法的数据隐私保护效果如图 8 所示。

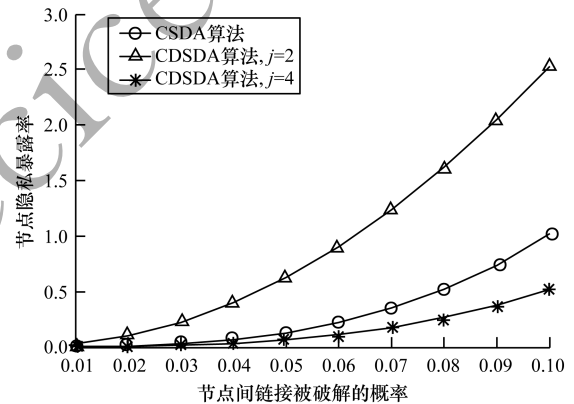


图 8 2 种算法的数据隐私保护效果对比

从图 8 可以看出, CSDA、CSDSA 2 种算法的隐私暴露率均随着 q 值的增加而增加,在相同 q 值时, CSDSA 根据数据重要程度对数据进行切分,将一般数据切分成 2 片,该类数据的隐私暴露率高于 CSDA 算法;将重要数据切分成 4 片,该类数据的隐私暴露率低于 CSDA 算法。因此,在隐私保护性能方面, CSDSA 算法更加注重数据的实际意义,能对更重要的数据进行更好地隐私保护。

2.2 通信开销分析

本文针对相同簇内的通信开销与节点数量、汇聚时间的关系分别进行仿真,比较 CSDA 和 CSDSA

2种算法在通信开销方面的性能。通信开销指节点对数据传输时消耗的能量,可通过节点的切片数量来体现,切片数量越多,需要传输的数据就越多,通信开销就越大。

首先,在节点数为 N (不含簇头)且具有相同结构的簇内,比较CSDA和CSDSA 2种算法的通信开销。CSDA算法采用切分重组技术,且随着簇内节点数的递增,每个节点的切片值也递增,故其通信开销为:

$$C_{\text{CSDA}} = \sum_{i=1}^N j_i, j_i = N \quad (12)$$

CSDSA算法中簇内节点根据数据的重要程度进行合理的动态切分,故其通信开销为:

$$C_{\text{CSDSA}} = \sum_{i=1}^N j_i, j_i = 2, 4, 6 \quad (13)$$

2种算法的通信成本随簇内节点数的变化关系如图9所示,其中,数据重要性以同等概率随机确定。CSDSA算法根据节点剩余能量动态调整切片数量,当 $E_r > 30\%$ 时,视为节点能量较高,采用式(1)进行切片,得曲线CSDSA-H;当 $E_r \leq 30\%$ 时,视为节点能量较低,采用式(2)进行切片,得曲线CSDSA-L。

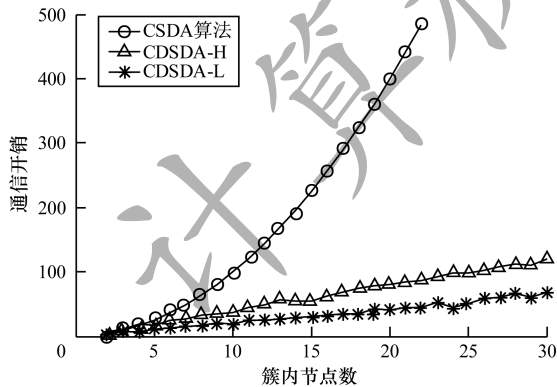


图9 2种算法通信开销与节点数的关系

从图9可以看出,在相同簇的情况下,随着簇内节点数的增加,算法通信开销不断增加。在CSDA算法中,由于簇内每个节点的切片数量均随着节点个数的增加而线性增加,故簇内节点的通信开销随着节点数的增加呈指数增加;而在CSDSA算法中,由于簇内每个节点的切片数量只与所采集数据的重要程度有关,故簇内节点的通信开销随着节点个数的增加而缓慢增加。当簇内节点数量较少时,其通信开销随着收集数据重要程度的变化而变化,在相同情况下,CSDSA算法中的簇内节点对数据的切片数量要多于CSDA算法,故CSDSA算法的通信开销稍高于CSDA算法,但随着节点数量的逐渐增加,CSDSA算法的通信开销逐渐低于CSDA算法,且差距越来越大。

2种算法的通信开销与汇聚时间之间的关系如图10所示。图7中整个网络簇内平均节点数量约

为6个,结合图9可以看出,此时CSDA算法中每个簇内节点的通信开销大于CSDSA算法,故初始阶段CSDA算法整个网络通信开销也大于CSDSA算法。当网络汇聚时间达到 2.0×10^6 ms时,部分节点开始死亡,对于CSDA算法而言,节点收发的切片数量仍然保持不变,故其通信开销开始迅速下降,直到全部节点死亡时通信开销将为零;而对于CSDSA算法而言,此时剩余能量 E_r 小于30%的节点会根据式(2)合理减少对原始数据的切分,因而会延长部分节点以及整个网络的寿命。但是,切片数的减少可能会引起数据传输过程中安全性的降低。

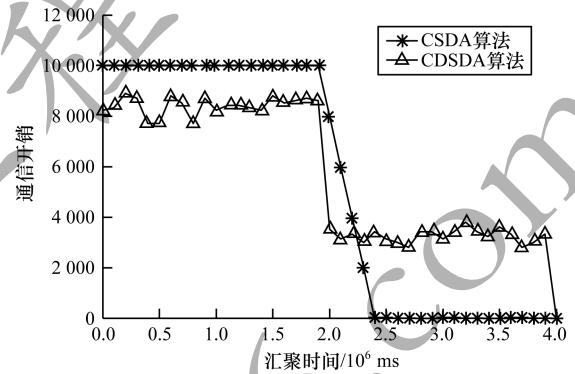


图10 2种算法通信开销与汇聚时间的关系

2.3 有噪信道下的数据正确率分析

数据正确率指收集池端得到的最终正确数据与原始数据之间的比值,用 P_1 表示。在有噪信道中,数据加密传输过程中由于解密算法的雪崩效应,密文传错1 bit,将导致整个分组数据全错。假设传输1 bit数据的正确率为 Z_1 ,则传输 Q bit数据的正确率为 $P_1 = (Z_1)^Q$,其出错率为 $P_2 = 1 - (Z_1)^Q$,即数据分组长度越大,出现错误的几率越大,这表明数据切片数越多,单个数据量越小,传输时正确率就越高。如图11所示,数据正确率随切片数的增加而增加。

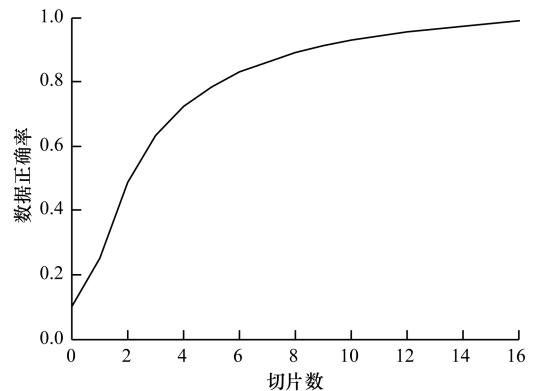


图11 数据正确率与切片数的关系

对于CSDA算法而言,切片数随簇内节点数的增加而增加。当簇内节点数较少时,数据切片数量相对较少,传输出现的错误率较高,数据正确率较

低;当簇内节点数较多时,数据切片数量相对较多,传输出现的错误率较低,数据正确率较高,但此时其通信开销也会大幅增加。

相比而言,CDSDA 算法中的切片数取决于原始数据的重要程度,数据越重要,切片数越多,单个数据量越小,传输出现的错误率就越低,重要数据的正确率就越高;对于一般数据或不重要的数据,切片数较少,其最终获取效率较低。考虑到实际应用,最大程度地获取重要数据,适当放弃一般数据,更加有利于提高有效数据的正确率。因此,在数据正确率方面,CDSDA 算法明显优于 CSDA 算法。

3 结束语

本文提出一种以分簇结构为基础的动态切分安全数据汇聚算法 CDSDA。节点根据原始数据的重要程度对数据进行动态切片,以提高对重要数据的隐私保护程度。仿真结果表明,与 CSDA 算法相比,CDSDA 算法能降低通信开销,提高数据正确率。

参考文献

- [1] BOUDIA O R M, SENOUCI S M, FEHAM M. A novel secure aggregation scheme for wireless sensor networks using stateful public key cryptography [J]. Ad Hoc Networks, 2015, 32 (C): 98-113.
- [2] SHIM K A, PARK C M. A secure data aggregation scheme based on appropriate cryptographic primitives in heterogeneous wireless sensor networks [J]. IEEE Transactions on Parallel and Distributed Systems, 2015, 26(8): 2128-2139.
- [3] HE Wenbo, LIU Xue, NGUYEN H, et al. PDA: privacy-preserving data aggregation in wireless sensor networks [C]//Proceedings of INFOCOM'07. Washington D. C., USA: IEEE Press, 2007: 2045-2053.
- [4] ZHANG Changlun, LI Chao, ZHAO Yi. A balance privacy-preserving data aggregation model in wireless sensor networks [J]. International Journal of Distributed Sensor Networks, 2015 (1): 1-6.
- [5] WANG Jun, CHEN Yu. Research and improvement of wireless sensor network secure data aggregation protocol based on SMART [J]. International Journal of Wireless Information Networks, 2018 (11): 1-9.
- [6] FANG Wei, WEN Xuezi, XU Jiang, et al. CSDA: a novel cluster-based secure data aggregation scheme for WSNs [J]. Cluster Computing, 2017 (4): 1-12.
- [7] MADDEN S, FRANKLIN M J, HELLERSTEIN J M, et al. TAG: a tiny aggregation service for Ad-Hoc sensor networks [C]//Proceedings of the 5th Symposium on Operating Systems Design and Implementation. New York, USA: ACM Press, 2002: 153-167.
- [8] SINHA A, LOBIYAL D K. Prediction models for energy efficient data aggregation in wireless sensor network [J]. Wireless Personal Communications, 2015, 84 (2): 1325-1343.
- [9] 阮志强. 分布式传感器网络数据安全性若干关键技术研究 [D]. 长沙: 湖南大学, 2012.
- [10] ESCHENAUER L, GLIGOR V D. A key-management scheme for distributed sensor networks [C]//Proceedings of the 9th ACM Conference on Computer and Communications Security. New York, USA: ACM Press, 2002: 45-63.
- [11] SPENCER J. The strange logic of random graphs [M]. Berlin, Germany: Springer, 2001.
- [12] SEN J, MAITRA S. An attack on privacy preserving data aggregation protocol for wireless sensor networks [C]//Proceedings of Nordic Conference on Information Security Technology for Applications. Berlin, Germany: Springer, 2011: 205-222.
- [13] BING L. Dynamic clustering of distributed source coding in wireless sensor networks [J]. The Journal of China Universities of Posts and Telecommunications, 2009, 16 (1): 22-26.
- [14] JUNG W S, LIM K W, KO Y B, et al. Efficient clustering-based data aggregation techniques for wireless sensor networks [J]. Wireless Networks, 2011, 17 (5): 1387-1400.
- [15] 任秀丽, 吉鹏硕. WSN 中基于分簇的模糊加权数据融合算法 [J]. 计算机工程, 2018, 44 (3): 109-113, 118.
- [16] 付帅, 姜奇, 马建峰. 一种无线传感器网络隐私保护数据聚合方案 [J]. 计算机研究与发展, 2016, 53 (9): 2030-2038.

编辑 吴云芳