

# 一种改进的水平分布式环境下基于同态加密的隐私保护算法

张燕平 凌捷

(广东工业大学计算机学院 广州 510006)

**摘 要** 提出了一种改进的水平分布式环境下关联规则挖掘的隐私保护算法,该算法应用部分隐藏的随机化回答方法和同态加密技术,引入半可信第三方,将各站点的数据集进行扰乱和隐藏,并将数据的水平格式表示转换成垂直格式表示,通过位运算计算局部支持数,利用 Paillier 算法计算全局支持数。所提算法具有站点之间无须通信、支持数计算效率高、I/O 操作次数少以及传输安全等优点。实验结果表明,所提算法提高了局部支持数的计算效率并减少了 I/O 操作次数。

**关键词** 关联规则挖掘,隐私保护,同态加密,位运算

**中图法分类号** TP393.08

**文献标识码** A

**DOI** 10.11896/j.issn.1002-137X.2017.08.028

## Improved Algorithm for Privacy-preserving Association Rules Mining on Horizontally Distributed Databases

ZHANG Yan-ping LING Jie

(School of Computers, Guangdong University of Technology, Guangzhou 510006, China)

**Abstract** An improved privacy-preserving algorithm based on homomorphic for association rules mining on horizontally partitioned environment was proposed in this paper. The algorithm utilizes the approach of randomized response with partial hiding and homomorphic encryption technology, introduces a semi-trusted third party, disrupts and hides the data sets of each site, convertes the horizontal format into vertical format, calculates the number of local support by bit operation and uses Paillier algorithm to compute the number of global support. The algorithm has some advantages, such as no need for communication between sites, high computational efficiency of support, fewer I/O operations and safe transport. Experimental results show that the algorithm can improve the computational efficiency of local support and reduce the number of I/O operations.

**Keywords** Association rules mining, Privacy-preserving, Homomorphic encryption, Bit operation

## 1 引言

随着网络信息技术的发展,分布式环境下隐私保护数据挖掘<sup>[1]</sup>的应用越来越广泛,比如银行、保险、医疗等领域,目的是在不泄露各方隐私信息的前提下进行数据共享,有效且准确地挖掘出需要的信息。

关联规则挖掘是数据挖掘的重要方法之一,旨在挖掘出大量数据中项集之间有用的关联,其在隐私保护问题上取得了一定的研究进展。分布式环境下的隐私保护关联规则挖掘一般采用安全多方计算策略<sup>[2-3]</sup>,由于其计算复杂,近几年研究热点是在隐私保护技术中加入计算量小且安全的同态加密技术<sup>[4]</sup>。XUAN 等人<sup>[5]</sup>提出使用两个服务站点(一个站点作为初始方,一个站点作为组合方),结合最大频繁项集(MFI)方法生成候选项集,使用 Paillier 同态加密算法计算全局支持数的算法,其缺点是只有在最大频繁  $k$ -项集最大长度接近频繁  $k$ -项集最大长度的情况下算法的计算效率才较高。

文献[6]使用一个服务站点作为数据中心,采用多参数随机干扰方法与同态加密技术相结合的方法对原始数据进行干扰和隐藏,并利用 Paillier 算法计算项集全局支持度。该方法在挖掘数据之前先对原始数据进行了扰乱和隐藏,使得数据中心和攻击者无法得到原始数据信息。在最大频繁项集最大长度与频繁项集最大长度的差距很大的情况下,其计算效率比文献[5]高,但仍然有待改进。

Rana 等人<sup>[7]</sup>提出了基于 FP-Growth 的挖掘算法,该算法通过构造 FP-tree 语法树以提高频繁项集支持数的计算效率,使用 Paillier 算法安全地传输支持数,其缺点是构造 FP-tree 树会占用很大的内存空间。

Kaosar 等人<sup>[8]</sup>提出了一种基于完全同态加密的双方关联规则挖掘方案,该方案通过产生可重复使用的电路及公钥来提高计算效率并降低通信代价。

上述方法有效地保障了个人隐私数据的安全性,但计算效率都偏低。本文提出一种改进的水平分布环境下基于同态

到稿日期:2016-06-01 返修日期:2016-08-12 本文受广东省科技计划项目(2014A010103029, 2015B010128014, 2015B090906015, 2016B010107002, 2016B090918039)资助。

张燕平(1992-),女,硕士生,主要研究方向为网络与信息安全, E-mail: 957743268@qq.com; 凌捷(1964-),男,博士,教授, CCF 会员,主要研究方向为网络与信息安全。

加密的隐私保护关联规则挖掘(EARPRD)算法,该算法综合了部分隐藏随机干扰技术与同态加密技术的优点,引入半可信第三方作为数据服务中心 DSC,以防止站点之间进行通信;同时利用位运算来快速计算局部支持度,降低了计算代价。利用 Paillier 同态加密算法求得全局支持数,实现了数据的有效保护。理论分析和实验结果表明,本文算法是安全有效的。

## 2 问题形式化

设  $I = \{I_1, I_2, \dots, I_m\}$  为一组项的集合,  $D = \{T_1, T_2, \dots, T_n\}$  为数据库事务  $T$  的集合,简称数据集,其中  $T \subseteq I$ 。设  $X$  为  $I$  中不同项的集合,如果  $X \subseteq T$ ,则称  $T$  包含  $X$ 。如果  $X$  和  $Y$  是  $I$  中的项集,且  $X \cap Y = \emptyset$ ,则蕴含式  $X \Rightarrow Y$  为关联规则。规则  $X \Rightarrow Y$  的支持度指的是包含  $X \cap Y$  的事务占数据集  $D$  的百分比,置信度指的是包含  $X \cap Y$  的事务占包含  $X$  的事务的百分比。支持度(sup)和置信度(conf)的数学公式表示如下:

$$\text{sup}(X \Rightarrow Y) = \frac{\sigma(X \cup Y)}{|D|} \quad (1)$$

$$\text{conf}(X \Rightarrow Y) = \frac{\sigma(X \cup Y)}{\sigma(X)} \quad (2)$$

其中,  $|D|$  表示总事务数,  $\sigma(X \cup Y)$  表示包含  $X$  和  $Y$  的事务总数,  $\sigma(X)$  表示包含  $X$  的事务总数。

设最小支持度阈值为  $\min\_s$ ,最小置信度阈值为  $\min\_c$ 。支持度大于  $\min\_s$  的项集为频繁项集,置信度大于  $\min\_c$  的频繁项集为关联规则。

在水平分布式环境下,假设有  $N$  个数据集  $D_1, D_2, \dots, D_N$ ,分别分布在  $S_1, S_2, \dots, S_N$  这  $N$  个站点中,数据集  $D_i$  的大小为  $|D_i|$ ,所有站点拥有的数据集具有相同的项的集合、不同的事务。设在  $S_i$  站点,事务  $T$  的支持数为  $S_{T_i}$ ,则称  $T$  在站点  $S_i$  中的局部支持数为  $S_{T_i}$ 。目标是借助半可信第三方,在不泄露各站点原始数据和  $S_{T_i}$  的前提下,利用部分隐藏的随机化回答方法扰乱与隐藏数据项集,引入位运算技术,并结合同态加密方法,尽量高效准确地计算  $T$  的全局支持度,从而找出全局频繁项集,得到关联规则。

## 3 改进的水平分布式环境下的隐私保护算法

在水平分布式环境下,设有  $N$  个站点  $S_1, S_2, \dots, S_N$ ,各站点拥有相同的项的集合。本文设计的水平分布式环境下的关联规则挖掘的隐私保护模型结构如图 1 所示。

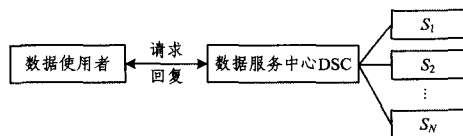


图 1 模型框架

数据服务中心 DSC 的主要任务是负责接收并整合各站点发送过来的局部支持度,求出全局支持度,根据数据使用者的请求信息,产生频繁项集,并得出关联规则。

### 3.1 数据处理

各站点使用从 DSC 发送过来的参数对数据库中的事务进行部分隐藏随机干扰操作。给定随机化参数  $0 \leq p_1, p_2$ ,

$p_3 \leq 1$ ,且满足  $p_1 + p_2 + p_3 = 1$ 。设项的总数为  $k$ ,站点  $S_i$  的原始事务集  $X = \{x_1, x_2, x_3, \dots, x_k\}$ ,进行随机干扰操作后的事务集为  $Y = \{y_1, y_2, y_3, \dots, y_k\}$ ,其可以通过  $Y = F(X)$  计算得到,其中  $y_i = f(x_i)$ ,且项  $y_j$  取值为  $x_j$  的概率为  $p_1$ ,取值为 1 的概率为  $p_2$ ,取值为 0 的概率为  $p_3$ 。由扰乱后数据集的项集支持数可重构出原始数据集的项集支持数。现设  $C^X$  为原始矩阵  $X$  的  $k$ -项集支持度,  $C^Y$  为扰乱后矩阵  $Y$  的  $k$ -项集支持度。

$$M_k C^X = C^Y \quad (3)$$

$M_k = [m_{p,q}]$  是  $k+1$  阶矩阵,  $m_{p,q}$  表示恰好包含  $X$  中  $q$  项的数据集  $D$  中事务经过数据干扰和隐藏后,转变成恰好包含  $X$  中  $p$  项的  $D^Y$  中事务的概率。当  $M_k$  可逆时,有:

$$C^X = M_k^{-1} C^Y \quad (4)$$

数据重构后,挖掘结果与扰乱前的挖掘结果存在一定的误差,文献[12]提出的部分隐藏的随机化回答方法对支持度误差与隐私破坏系数的关系进行了分析及实验,有以下结论:

当  $\frac{1}{3} < p_1 < \frac{1}{\sqrt{2}}$ ,  $p_2 = p_3 = (1 - p_1)/2$  时,部分隐藏的随机干扰方法具有最好的准确性和隐私性。本算法根据该结论产生随机数  $p_1, p_2$  和  $p_3 (p_2 = p_3)$ 。

### 3.2 数据转换

各站点将本地扰乱和隐藏后的数据库扫描一遍,将数据水平格式表示转换为垂直格式表示,使得数据库中的每一行代表数据库的项,每一列代表事务。并将数据库中每一条记录的一个项及其中所存在的所有事务记录的列表构成一个  $Tset$  表,使得任一项集的支持度可通过  $Tset$  进行交集运算求得。下面给出一个数据转换的例子。

表 1 是事务列表,将事务列表中的不同属性进行规范化,表 2 是对表 1 的规范化表示。其中有 10 个属性,用“Bits”来表示这些属性在事务中的存在情况,“Support”表示存在该属性的事务总数。

表 1 事务列表

| TID | Attributes  | TID | Attributes  |
|-----|-------------|-----|-------------|
| 1   | 1 3 4 5 9   | 5   | 5 6 7 9 10  |
| 2   | 2 3 4 6     | 6   | 3 5 6 7 8 9 |
| 3   | 4 6 7 10    | 7   | 1 4 5 9 10  |
| 4   | 1 2 5 6 8 9 | 8   | 2 4 6 9     |

表 2 事务列表规范表示

| Attributes | Representation                     | Bits       | Support |
|------------|------------------------------------|------------|---------|
| 1          | $\{T_1, T_4, T_5\}$                | (10011000) | 3       |
| 2          | $\{T_2, T_4, T_8\}$                | (01010001) | 3       |
| 3          | $\{T_1, T_2, T_6\}$                | (11000100) | 3       |
| 4          | $\{T_1, T_2, T_3, T_7, T_8\}$      | (11100011) | 5       |
| 5          | $\{T_1, T_4, T_5, T_6, T_7\}$      | (10011110) | 5       |
| 6          | $\{T_2, T_3, T_4, T_5, T_6, T_8\}$ | (01111101) | 6       |
| 7          | $\{T_3, T_5, T_6\}$                | (00101100) | 3       |
| 8          | $\{T_4, T_6\}$                     | (00010100) | 2       |
| 9          | $\{T_1, T_4, T_5, T_6, T_8\}$      | (10011101) | 5       |
| 10         | $\{T_3, T_5, T_7\}$                | (00101010) | 3       |

### 3.3 支持度计算及传输

计算扰乱后的项集的支持数时,只需统计“Bits”项中“1”的个数,可通过简单的“位”运算 AND 求得。为简化计算,引

人分治的思想,第一次按 1-bit 长度进行划分,将相邻的两个 1-bit 相加即为每 2-bit 中“1”的个数。然后分别按 2-bit, 4-bit, 8-bit, 16-bit 等长度进行划分,进行相同的操作,计算“1”的个数。通过  $Tset$  进行交集运算计算  $k$ -项集( $k>1$ )的支持度,如  $Tset(I_1) = \{T_1, T_3, T_4, T_5, T_9\}$ ,  $Tset(I_2) = \{T_2, T_3, T_4, T_6\}$ ,  $sup(I_1, I_2) = Tset(I_1) \cap Tset(I_2) = 001100000$ 。统计 1 的个数,即可得出项集  $I_1 I_2$  的支持数。

计算出项集的支持数后,为保证支持数在传输过程中的安全性,使用 Paillier 同态加密算法对其进行加密后再发送给数据服务中心 DSC。

设  $M_i$  为各站点隐藏和扰乱后的事务集对应的矩阵。 $C_k^{M_i}$  为  $M_i$  的候选  $k$ -项集支持数矩阵。各站点并行计算候选  $k$ -项集的局部支持数后,使用 Paillier 算法对其进行加密并构成行矩阵  $E(C_k^{M_i})$ 。

$$\begin{aligned} & E(C_k^{M_1}) * E(C_k^{M_2}) * \dots * E(C_k^{M_k}) \\ & = E(C_k^{M_1}) + E(C_k^{M_2}) + \dots + E(C_k^{M_k}) = E(C_k^M) \end{aligned} \quad (5)$$

### 3.4 算法描述

输入:  $N$  个站点  $S_1, S_2, \dots, S_N$ , 一个数据服务中心 DSC。  
最小支持度  $min\_s$ , 最小置信度  $min\_c$ 。

输出: 关联规则。

第一阶段: 用部分隐藏的随机化回答方法扰乱原始数据, 并转换数据格式。

1) 数据服务中心 DSC 产生密钥( $pk, sk$ )和随机干扰参数  $p_1, p_2$ , 把公钥  $pk$  和参数  $p_1, p_2$  发送到各站点。

2) 各站点根据数据包中的参数  $p_1$  和  $p_2$  并行地对原始数据进行扰乱操作, 并将水平格式数据转换成垂直格式数据。

BEGIN

$k=1$ ;

repeat{

第二阶段: 产生频繁项集。

3) 各站点并行统计扰乱后的候选  $k$ -项集( $k$  的初始值为 1)的局部支持数, 同时利用 Paillier 同态加密算法加密, 构成行矩阵  $E(C_k^{M_i})$  并将其发送给 DSC。

4) DSC 将各站点发送过来的行矩阵进行加法运算操作后解密, 得到扰乱后的候选  $k$ -项集的全局支持数矩阵  $C_k^M$ 。并根据式(3)求出原始数据候选  $k$ -项集的全局支持数矩阵  $C_k^M$ 。

5) DSC 根据支持度公式计算出最小支持数  $min\_sup$ , 将统计出的全局支持数与最小支持数  $min\_sup$  进行比较, 若全局候选集的支持数大于  $min\_sup$ , 则此候选集就是全局频繁项集; 否则, 将其删除。

第三阶段: 产生关联规则。

6) 数据服务中心 DSC 使用关联规则挖掘经典算法 Apriori 改进算法<sup>[13]</sup>生成候选频繁( $k+1$ )-项集。

}until 全局频繁  $k$ -项集的数目等于 0。

7) DSC 计算出频繁项集的置信度, 将频繁项集的置信度与最小置信度  $min\_c$  进行比较, 如果频繁项集的置信度小于最小置信度, 则丢弃该组合; 如果频繁项集的置信度大于最小置信度, 这个组合则被认为是符合要求的关联规则。

## 4 算法的安全性和效率分析

### 4.1 安全性分析

安全性指主要保证各站点仅知道自己的局部支持数, 而无法得到其他站点的局部支持数信息和隐私数据, 并且引入的第三方也不能得到各站点的局部支持数信息和隐私数据。本文算法的安全性分析如下:

(1) 使用引入半可信第三方的模型, 不存在各站点之间的通信, 防止各站点学习其他站点的信息;

(2) 数据服务中心 DSC 将公钥  $pk$  和扰乱参数  $p_1, p_2$  发送给各站点, 发送的内容无隐私信息, 故无安全问题;

(3) 各站点利用 Paillier 算法对计算出的候选项集支持数加密后发送给 DSC, 即使攻击者截取了该数据包, 由于没有私钥  $sk$ , 也不能得到数据包的真正信息;

(4) 重构项集支持数矩阵时, 得到的是原始数据的全局支持数矩阵, 并没有泄漏各站点的局部数据信息。

### 4.2 效率分析

文献[6]提出的 ARPRD 算法在计算项集支持数时, 使用的是 Apriori 经典算法, 而本文使用的是 Apriori 改进算法, 其优势具体体现在以下几个方面:

(1) 在数据垂直表示的数据库中计算项集支持数时只需遍历  $M$  次( $M$  为项的总数), 而 ARPRD 算法则需遍历每一个事务, 遍历次数会达到最大数据长度。一般情况下, 数据库长度是项数的上千倍。

(2) 在计算  $k$ -项集  $I_i$  的支持数时, 本算法采用的是位运算“AND”操作, 并利用分治思想快速求得, 而 ARPRD 算法则是遍历  $C_k$  中所有的  $I_i$ , 如果  $I_i$  在该记录事务中存在, 则  $sup(I_i)++$ 。

(3) 在计算  $k$ -项集( $k>1$ )支持数时, 本算法只需求出  $k$ -项集中项集的交集中“1”出现的个数, 而 ARPRD 算法每次产生新的频繁项集时都需要扫描事务数据库, 大大降低了计算效率。

## 5 实验及结果分析

### 5.1 实验环境

本文使用 C++ 语言实现上述算法, 在一台计算机上模拟多个站点, 使用远程过程方法调用 RMI for C++(一个专门针对 C++ 语言的远程方法调用框架)实现站点之间的通信。

实验中使用的数据为 IBM Almaden 数据集生成器生成合成的事务数据集, 共生成两个数据集, 数据集 A 生成参数设置为 D100k,  $I_4, T_{10}, N_{20}$ , 即总事务为 100000 个, 频繁项集的平均长度是 4, 每个事务平均包含 10 个项, 属性个数共有 20 个。数据集 B 生成参数设置为 D3k,  $I_6, T_{10}, N_{20}$ 。设定支持度为 40%, 置信度为 60%。

### 5.2 实验结果对比

为验证本文提出的 EARPRD 算法的计算效率, 实验分成两部分: 1) 对比只有两个站点参与挖掘的情况下, EARPRD 算法与 ARPRD 算法、文献[14]的 PPDAR 算法在同一数据集、不同事务数的条件下的执行时间。将数据集随机分布在

站点  $S_1, S_2$  上, 测试事务数分别为 2500, 25000, 50000, 100000 个。2) 在增加多个站点后, 比较不同数据集、相同事务数下 EARPRD 算法、ARPRD 算法、文献[5]的 EPPDAR 算法的执行时间。通过手工将生成的两个数据集进行修改, 使得数据集 A(DA) 的最大频繁  $k$ -项集的最大长度接近于频繁  $k$ -项集的最大长度, 使得数据集 B(DB) 的最大频繁  $k$ -项集的最大长度与频繁  $k$ -项集的最大长度差距很大, 测试事务数为 5000 个。

图 2 示出了本文提出的 EARPRD 算法与 ARPRD 算法和 PPDAR 算法在两个站点不同总事务数下算法执行时间的对比情况。图 3 和图 4 示出了不同事务集在站点逐渐增加的情况下, EARPRD 算法与 ARPRD 算法、EPPDAR 算法执行时间的对比情况。从图 2—图 4 中可以看出, 在不同站点、不同事务集下, EARPRD 算法的执行时间都比 ARPRD 算法和 EPPDAR 算法的执行时间少。主要原因是 EPPDAR 算法采用 MFI 方法来求得数据集中的最大频繁项集, 该方法的主要缺点是只有数据集中最大频繁  $k$ -项集的最大长度接近于频繁项集的最大长度时计算效率才比较高; 另一方面, 各站点需要将本地最大频繁  $k$ -项集使用 RSA 加密后再发送给组合器, 计算代价较大。ARPRD 算法中各站点使用 Paillier 算法只对频繁项集的支持数进行加密, 大大减少了计算代价。而本文提出的 EARPRD 算法利用了 ARPRD 算法的优点, 并在其基础上将数据库水平格式数据表示转换成垂直格式数据表示, 减少了扫描数据库时的遍历次数, 并在计算支持数时使用位运算和分治思想, 简化了运算, 使计算效率得到了提高。

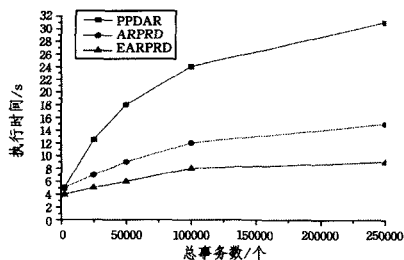


图 2 不同总事务集下算法执行时间的对比

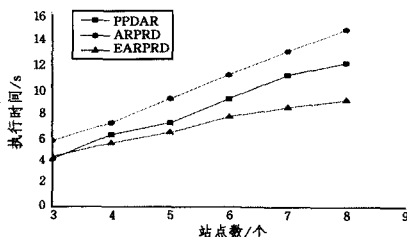


图 3 DA 算法执行时间的对比

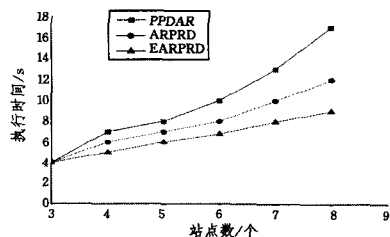


图 4 DB 算法执行时间的对比

**结束语** 本文提出了一种改进的适合水平分布式环境的隐私保护算法。该算法基于同态加密技术, 先利用部分隐藏随机干扰方法对原始数据进行扰乱和隐藏, 保护原始数据; 对数据集中的数据结构进行转换, 减少了事务集遍历次数; 将位运算与分治思想结合, 提高了局部支持数计算效率; 并应用 Paillier 同态加密技术安全地计算全局支持数。实验结果表明, 本文算法与 ARPRD 算法相比具有更高的执行效率。

## 参考文献

- [1] LIU Y H, ZHANG T Y, JIN X L, et al. Personal Privacy Protection in the Era of Big Data[J]. Journal of Computer Research and Development, 2015, 52(1): 229-247. (in Chinese)  
刘雅辉, 张铁赢, 靳小龙, 等. 大数据时代的个人隐私保护[J]. 计算机研究与发展, 2015, 52(1): 229-247.
- [2] KANTARCIOGLU M, CLIFTION C. Privacy-preserving distributed mining of association rules on horizontally partitioned data[J]. IEEE Trans. on Knowledge and Data Engineering, 2004, 16(9): 1026-1037.
- [3] LIU F, XUE A R, WANG W. Hybrid algorithm for privacy preserving association rules mining[J]. Application Research of Computers, 2012, 29(3): 1107-1110. (in Chinese)  
刘峰, 薛安荣, 王伟. 一种隐私保护关联规则挖掘的混合算法[J]. 计算机应用研究, 2012, 29(3): 1107-1110.
- [4] QIAN P, WU M. Survey of privacy preserving data mining methods based on homomorphic encryption[J]. Application Research of Computers, 2011, 28(5): 1614-1617. (in Chinese)  
钱萍, 吴蒙. 同态加密隐私保护数据挖掘方法综述[J]. 计算机应用研究, 2011, 28(5): 1614-1617.
- [5] XUAN C N, LE H B, CAO T A. An Enhanced Scheme for Privacy-Preserving Association Rules Mining on Horizontally Distributed Databases[C]// 2012 IEEE RIVF International Conference on Computing and Communication Technologies, Research, Innovation, and Vision for the Future (RIVF). IEEE, 2012: 1-4.
- [6] CHEN Y C. Research on Privacy Preserving Algorithms for Association Rules Mining in Distributed Environment[D]. Nanjing: Guangxi University, 2013. (in Chinese)  
陈玉婵. 面向关联规则挖掘的分布式隐私保护算法研究[D]. 南宁: 广西大学, 2013.
- [7] RANA S, THILAGAM P S. Hierarchical Homomorphic Encryption Based Privacy Preserving Distributed Association Rule Mining[C]// International Conference on Information Technology. IEEE, 2014: 379-385.
- [8] KAOSAR M G, PAULET R, YI X. Fully homomorphic encryption based two-party association rule mining[J]. Data & Knowledge Engineering, 2012, 76-78(2): 1-15.
- [9] RIVEST R. On databanks and privacy homomorphism[J]. Foundations of Secure Computation, 1978, 4(11): 169-179.
- [10] FENG D G, ZHANG M, LI H. Big Data Security and Privacy Protection[J]. Chinese Journal of Computers, 2014, 37(1): 246-258. (in Chinese)  
冯登国, 张敏, 李昊. 大数据安全与隐私保护[J]. 计算机学报,

- 2014,37(1):246-258.
- [11] HUANG L S, TIAN M M, HUANG H. Preserving Privacy in Big Data: A Survey from the Cryptographic Perspective[J]. Journal of Software, 2015, 26(4): 945-959. (in Chinese)  
黄刘生, 田苗苗, 黄河. 大数据隐私保护密码技术研究综述[J]. 软件学报, 2015, 26(4): 945-959.
- [12] ZHANG P, DONG Y H, TANG S H, et al. An Effective Method for Privacy Preserving Association Rule Mining[J]. Journal of Software, 2006, 17(8): 1764-1774. (in Chinese)  
张鹏, 童云海, 唐世渭, 等. 一种有效的隐私保护关联规则挖掘方法[J]. 软件学报, 2006, 17(8): 1764-1774.
- [13] FU S, ZHOU H J. The Research and Improvement of Apriori-Algorithm for Mining Association Rules[J]. Microelectronics & Computer, 2013(9): 110-114. (in Chinese)  
付沙, 周航军. 关联规则挖掘 Apriori 算法的研究与改进[J]. 微电子学与计算机, 2013(9): 110-114.
- [14] HSSEIN M, EL-SISI A, ISMAIL N. Fast Cryptographic Privacy Preserving Association Rules Mining on Distributed Homogeneous Data Base [C] // Proceedings of the 12th International Conference on Knowledge-Based Intelligent Information and Engineering Systems, Part II. Springer-Verlag, 2008: 607-616.
- [15] ZHANG Y, WANG H G, SHAO Z Z, et al. Frequent itemsets mining based on Apriori-bit[J]. Application Research of Computers, 2013, 30(9): 2610-2612. (in Chinese)  
张岳, 王洪国, 邵增珍, 等. 基于先验位运算的频繁项集挖掘[J]. 计算机应用研究, 2013, 30(9): 2610-2612.
- [16] WAHAB A, OMAR, HACHAMI, et al. DARM: a privacy-preserving approach for distributed association rules mining on horizontally-partitioned data[C] // International Database Engineering & Applications Symposium. ACM, 2014.
- [17] YI X, RAO F Y, BERTINO E, et al. Privacy-Preserving Association Rule Mining in Cloud Computing[C] // Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security. ACM, 2015: 439-450.
- (上接第 133 页)
- 朱世才, 王海涛, 吴连才, 等. 基于 SMP 的分簇 WSN 生存性评估模型[J]. 传感技术学报, 2014, 27(3): 383-387.
- [9] YI Z, DOHI T, OKAMURA H. Survivability modeling and analysis for a power-aware wireless ad hoc network[C] // International Congress on Ultra Modern Telecommunications and Control Systems and Workshops. IEEE, 2012: 813-819.
- [10] XIONG S M, WANG L M, ZHAN Y Z. Quantitative evaluation of topology intrusion tolerance in wireless sensor networks based on semi-Markov process[J]. Journal on Communications, 2010, 31(7): 24-32. (in Chinese)  
熊书明, 王良民, 詹永照. 基于 SMP 的无线传感器网络拓扑入侵定量评估[J]. 通信学报, 2010, 31(7): 24-32.
- [11] SHEN S G, HUANG L J, FAN E. Survivability evaluation for WSNs under malware infection[J]. Chinese Journal of Sensors and Actuators, 2016, 29(7): 1083-1089. (in Chinese)  
沈士根, 黄龙军, 范恩. 受恶意程序传染的 WSNs 可生存性评估[J]. 传感技术学报, 2016, 29(7): 1083-1089.
- [12] LI J P, WANG X K. WSN reliability evaluation based on fuzzy neural network[J]. Journal of Computer Applications, 2016, 36(A02): 69-72. (in Chinese)  
李建平, 王晓凯. 基于模糊神经网络的无线传感器网络可靠性评估[J]. 计算机应用, 2016, 36(A02): 69-72.
- [13] YUE Y, LI J, FAN H, et al. An efficient reliability evaluation method for industrial wireless sensor networks[J]. Journal of Southeast University, 2016, 32(2): 195-200.
- [14] YI Z, DOHI T. Survivability Analysis for a Wireless Ad Hoc Network Based on Semi-Markov Model[J]. IEICE Transactions on Information & Systems, 2012, E95. D (12): 2844-2851.
- [15] CHANG C J, ZHU C H, WANG H G, et al. Survivability Evaluation of Cluster-Based Wireless Sensor Network under DoS Attacks[C] // Wireless Communications, Networking and Mobile Computing. 2009.
- [16] STAVROU E, PITSLILDES A. Situation aware intrusion recovery policy in WSNs[C] // International Conference on Cyber Situational Awareness, Data Analytics and Assessment. 2016: 1-8.
- [17] VIRMANI D, KAUR S, JAIN S. Secure and Fault Tolerant Dynamic Cluster Head Selection Method for Wireless Sensor Networks [J]. Procedia Computer Science, 2015, 46(2): 989-996.
- [18] NEWELL A, YAO H, RYKER A, et al. Node-Capture Resilient Key Establishment in Sensor Networks: Design Space and New Protocols [J]. Acm Computing Surveys, 2015, 47(2): 1-34.
- [19] HAMMOUDEH M, NEWMAN R. Adaptive routing in wireless sensor networks: QoS optimisation for enhanced application performance [J]. Information Fusion, 2015, 22(71): 3-15.
- [20] WARD J R, YOUNIS M. Base Station anonymity distributed Self-assessment in Wireless Sensor Networks[C] // IEEE International Conference on Intelligence and Security Informatics. IEEE, 2015.
- [21] KIM T Y, LEED R. Communication method, base station, communication system, and mobile terminal [J]. Journal of Organizational Behavior, 2015, 30(7): 1001-1018.
- [22] ZHANG Q, YU T, NING P. A Framework for Identifying Compromised Nodes in Wireless Sensor Networks [J]. ACM Transactions on Information & System Security, 2008, 11(3): 1-10.
- [23] SINGH V P, JAIN S, SINGHAI J. Hello Flood Attack and its Countermeasures in Wireless Sensor Networks [J]. International Journal of Computer Science Issues, 2010, 7(3): 23-24.
- [24] KARLOF C, WAGNER D. Secure routing in wireless sensor networks: attacks and countermeasures[C] // IEEE International Workshop on Sensor Network Protocols and Applications, 2003. IEEE, 2003: 293-315.
- [25] GOYAL S, BHATIA T, VERMA A K. Wormhole and Sybil attack in WSN: A review[C] // International Conference on Computing for Sustainable Global Development. IEEE, 2015.
- [26] ZHANG Y, FAN K F, ZHANG S B. AOA based trust evaluation scheme for Sybil attack detection in WSN[J]. Application Research of Computers, 2010, 27(5): 1847-1844. (in Chinese)  
张艳, 范科峰, 张素兵. 一种基于 AOA 信任评估的无线传感器网络 Sybil 攻击检测新方法[J]. 计算机应用研究, 2010, 27(5): 1847-1849.