文章编号: 1000-5641(2017)05-0087-14

面向智能电表隐私保护的电量请求方案

田秀霞[1,2], 李丽莎[3], 赵传强[3], 田福粮[4], 宋 谦[4]

(1. 上海电力学院 计算机科学与技术学院, 上海 200090;
2. 华东师范大学 数据科学与工程研究院, 上海 200062;
3. 国网浙江玉环市供电公司, 浙江 台州 317600;
4. 上海电力学院 电子与信息工程学院, 上海 200090)

摘要: 运通过有效融合Shamir(t,n) 门限密钥共享方案和 Laplace 噪音干扰算法提出了一种面向智能电表隐私保护的电量请求方案, 实现电力公司分时电价计费的同时保护用户隐私. 定量分析了安全性并确定了最优门限值 t 的选择、测试分析了时间效率、验证分析了 Laplace 噪音干扰的ε-差分隐私保护效果并作了方案的可行性比较. 实验结果表明, 提出的方案具有有效性和可行性.

关键词: 智能电表; 隐私保护; 密钥共享; Laplace 噪音
中图分类号: TP309 **文献标志码**: A **DOI:** 10.3969/j.issn.1000-5641.2017.05.009

Smart meter: Privacy-preserving power request scheme

TIAN Xiu-xia[1,2], LI Li-sha[3], ZHAO Chuan-qiang[3],
TIAN Fu-liang[4], SONG Qian[4]

(1. *College of Computer Science and Technology, Shanghai University of Electric Power, Shanghai* 200090, *China*;
2. *School of Data Science and Engineering, East China Normal University, Shanghai* 200062, *China*;
3. *State Grid Zhejiang Yuhuan Power Supply Company, Taizhou Zhejiang* 317600, *China*;
4. *College of Electronic and Information Engineering, Shanghai University of Electric Power, Shanghai* 200090, *China*)

Abstract: A privacy-preserving power request scheme was proposed. The proposed scheme combined Shamir (t,n) threshold secret sharing scheme with Laplace noise perturbation algorithm effectively to achieve paying TOU billing as well as protecting user privacy. Experiments were performed from four aspects: analyzing the security quantitatively and determining the optimal threshold t, giving the experiment on efficiency test, verifying the ε-differential privacy by introducing the Laplace noise perturbation and

收稿日期: 2017-06-20
基金项目: 国家自然科学基金重点项目(61532021); 国家自然科学基金面上项目(61772327); 上海市科学技术委员会地方能力建设项目(15110500700)
第一作者: 田秀霞, 女, 教授, 硕士生导师, 研究方向为数据库安全、隐私保护、基于密码学的访问控制. E-mail: xxtian@shiep.edu.cn.

conducting the scheme feasibility comparison. Experimental results show that the proposed scheme is effective and feasible.

Key words: smart meter; privacy-preserving; secret sharing; Laplace noise

0 引 言

智能电网根据用户侧智能电表实时请求的电量调整电力公司的电量供应, 有效可靠的将电量从电力公司传输到用户侧, 不仅满足用户用电需求而且避免多余发电, 但同时导致了智能电表用户的隐私保护问题. 智能电表实时请求的电量数据会泄露用户隐私信息. 攻击者一旦掌握智能电表实时请求的电量数据及其身份ID, 就可以从实时电量数据推测出该智能电表用户用电模式, 进而获得该用户生活习性等隐私信息, 这将给用户造成难以估量的损失. 电力公司是 semi-honest 实体, 它可能根据实时电量数据窥探用户隐私, 但同时它需要知道用户请求的电量(总电量-均一电价或实时电量-分时电价)计收电费. 针对这个问题, 近年来研究者们提出了许多解决方法, 一般说来, 这些方法依据两大思路: 一是电力公司只知晓智能电表身份 ID 和智能电表请求的总(一个月或两个月)电量, 不知晓智能电表请求的实时电量[1-5], 这里需说明的是总电量无法泄露用户的隐私信息; 二是电力公司只知晓智能电表请求的实时电量, 不知晓智能电表身份 ID[6-8], 这样电力公司所掌握的隐私信息无法定位到某个具体用户. 但大多数解决方法[1-7,10-17]无法在保护用户隐私的同时实现分时电价计费. 老式的电费计算方式, 即均一电价计费, 无论在用电高峰期还是用电低谷期电价是一样的, 这样不利于用电客户主动性地避开用电高峰期节约用电, 所导致的问题就是用电高峰期的用电负荷持续居高不下, 用电低谷期的用电负荷依旧低迷, 给电网设备及发电厂设备造成强烈的, 甚至毁灭性的冲击, 严重影响了电网的稳定性. 实施分时电价计费, 有利于鼓励用电客户合理安排用电时间, 减少用户电费, 削峰填谷, 提高电力资源的利用效率; 有利于电网企业降低电网投资成本和运行成本, 保障电网的安全稳定运行; 有利于社会减少或延缓电力投资, 促进社会资源的合理配置. 分时电价计费的实施随着智能电表的普及部署已成为必然发展趋势. 尤其在我国电力紧缺日益严重的情况下, 分时电价计费格外重要.

本文基于第二个思路, 有效融合 Shamir(t,n) 门限密钥共享方案和 Laplace 噪音干扰算法, 提出一种面向智能电表隐私保护的电量请求方案. 利用 Shamir(t,n) 门限密钥共享方案使智能电表匿名自己身份 ID 并生成匿名购买凭证. 智能电表凭借购买凭证实时向电力公司请求电量, 实现电力公司分时电价计费的同时保护用户隐私不被电力公司(内部攻击者)窥探, 但在一定条件下, 电力公司可以定位智能电表用户以实现电费的可追溯性. 基于 ε-差分隐私的 Laplace 噪音干扰算法对传输前的电量数据进行干扰, 进一步防止智能电表请求的实时电量数据被外部攻击者窃听或截取而泄露用户隐私.

本文的贡献如下

(1) 基于 Shamir(t,n) 门限密钥共享方案实现电力公司分时电价计费以及电费可追溯性的同时保护用户隐私不被电力公司窥探.

(2) 基于 ε-差分隐私的 Laplace 噪音干扰算法干扰传输前的电量数据, 进一步防止电量数据被窃取而泄露用户隐私.

(3) 定量分析安全性并确定最优门限值 t 的选择、测试分析时间效率以及验证分析 Laplace 噪音干扰的 ε-差分隐私保护效果.

本文的结构如下: 第 1 节描述相关工作; 第 2 节介绍预备知识; 第 3 节描述系统模型; 第 4 节进行方案的详细描述; 安全分析和实验分别在第 5 节、第 6 节介绍; 最后对全文加以总结, 并指出进一步的工作方向.

1　相关工作

针对第一个思路, 文献 [1-5] 作出了努力. 文献 [1] 中密钥分发中心汇总智能电表请求的实时电量, 然后把总电量发送给电力公司, 密钥分发中心无法获知智能电表身份而电力公司能够解密电量信息获得智能电表身份 ID, 但该方案的漏洞是无法防止第三方密钥分发中心与 semi-honest 电力公司共谋, 若共谋成功则用户隐私完全暴漏. 文献 [2] 将智能电表作为聚合节点构建电力网络虚拟聚合树, 利用同态加密算法[3]的分布式增量数据加密将子节点的数据传递给父节点, 最终数据传到电力公司(聚合中心), 但电力公司无法计算某个具体用户的用电量. 文献 [4-5] 均采用充电电池保护用户的负荷信息. 充电电池和电力公司可以独自或同时向智能电器供电, 这样智能电表数据不能直接反映用户的用电信息, semi-honest 电力公司无法窥探用户隐私.

针对第二个思路, 文献 [6-8] 作出了努力. 文献 [6] 中智能电表使用匿名凭证请求电量, 不同匿名凭证上标有不同数值的电量, 需要预先生成大量的定值凭证. 文献 [7] 假设每个智能电表有两个 ID, 其中一个 ID 用于匿名传输含有用户隐私的信息(如实时电量), 但第三方(如电表制造商)掌握关键性信息, 即两个 ID 间的关系, 造成了隐患. 文献 [8] 利用 Zero-Knowledge 协议[9], 智能电表只需向电力公司提交 secret 证明自己的身份, 但该方案有两个缺点: 一是智能电表要承担繁琐的电费计算; 二是智能电表需长期保存关键性信息 secret, 即伪随机标签 $\{r_i\}(i=1,2,\cdots,N)$ 和密钥 $\{k_j\}(j=1,2,\cdots,m)$, 一旦泄露, 用户隐私随之暴露. 文献 [10-17] 均没有考虑电力公司的 semi-honest 性.

综上所述, 大多数方案[1-7,10-17] 在保护用户隐私的同时均只能进行均一电价计费, 不能实现分时电价计费. 本文基于第二个思路, 利用 Shamir(t,n) 门限密钥共享方案实现电力公司分时电价计费以及电费可追溯性的同时保护用户隐私不被电力公司窥探. 同时, 基于 ε-差分隐私的 Laplace 噪音干扰算法干扰实时电量数据, 在传输中增强用户隐私保护[26]. 文献 [19] 将智能电表模型化为高斯数据产生源, 利用编码器对输入的负荷电量进行扰动, 其缺点是电力公司收到的电量并非是真实的负荷电量. 针对类似问题, 我们利用 Diffie-Hellman(D-H) 协议[29]去噪, 实现电力公司真实电量的计算. 另外, 为提高效率, 利用高斯随机变量生成 Laplace 噪音[18].

2　预备知识

2.1　Shamir(t,n)门限密钥共享方案

基于多项式插值的 (t,n) 门限密钥共享方案[20]由 A. Shamir 在 1979 年提出, 由于该方案可实现多个个体同时实现密钥的开启, 目前已被应用到网上交易、电子银行及网上服务等领域. 门限密钥共享方案基本思想是一个密钥 K 被分为 n 个分钥. 特定数目(如 t 或更多)的分钥可以根据多项式插值的方法(或其它有效方法)恢复密钥 K. 下面是对分钥的定义.

定义 1　(分钥[21]) 一个一元多项式$f(x)=(K+a_1x+a_2x^2+\cdots+a_tx^{t-1})\bmod Q$, 其中 $a_1,a_2,\cdots,K\in F_Q$, Q 为大质数, F_Q 为一个有限域, K 为密钥. 分钥为多项式 $f(x)$ 曲线上的点, 即 $(x,y)(x\neq 0)$.

从定义可以看出, n 个分钥为多项式 $f(x)$ 曲线上 n 个点, 即 $(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$, 其中 $x_1,x_2,\cdots,x_n$ 为非零已知量; 任意 $t(t\leqslant n)$ 个$(x_{i1},y_{i1}),(x_{i2},y_{i2}),\cdots,(x_{it},y_{it})$ 可以重构多项式 $f(x)$ 恢复密钥 K.

2.2 差分隐私(Differential Privacy)

差分隐私是一个严格的可证明的算法, 在输入数据有差异的情况下, 输出结果保持较高的相似性, 更重要的是该算法不改变输入数据的属性. 差分隐私保护模型最初被应用在数据库领域, 目的是保护数据库中个体隐私信息, 而后被广泛应用在统计学、数据挖掘等领域, 在资源共享、大数据盛行的背景下, 差分隐私保护技术的研究成为热点.

为了更加清晰地说明差分隐私在本文中的运用, 我们根据文献 [18] 中差分隐私的定义来具体地说明差分隐私在本文中的定义. 一般, 智能电表每隔一定时间(如 15 s)请求一次电量, 设每次请求的电量为 $m_j(j=1,2,3,\cdots,l)$, 其中 l 为观察者取定的参考次数. 差分算法, 是改变任意一次请求电量 m_j 其总输出没有显著差别. 下面是具体说明. 设总电量 $m=m_1\cup m_2\cup m_3\cdots\cup m_l$, $\mathrm{nbrs}(m)$ 表示从数据 m 增加或减去一次请求的电量数据, 即 $\mathrm{nbrs}(m)=m\cup m_j(j\notin\{1,2,3,\cdots,l\})$ 或 $\mathrm{nbrs}(m)=m-m_j(j\in\{1,2,3,\cdots,l\})$.

定义 2 (ε-差分隐私[25])$A(m)$ 表示算法A 对应输入数据 m 的输出. 若对于所有 m 下式成立, 则 A 满足 ε-差分隐私, 则对于输入数据 m 算法 A 具有 ε-差分隐私保护水平, 其中 ε 表示隐私保护水平:

$$Pr=[A(m)=x]\leqslant \mathrm{e}^{\varepsilon}Pr[A(m')=x]$$

其中 $m'\in\mathrm{nbrs}(m)$, x 是输出响应, Pr 是算法 A 的随机概率分布.

2.3 Laplace干扰算法

Laplace 算法主要是对一个矩阵进行五点的差分操作, 在数学和信号分析等领域有广泛的应用, 由于该算法和差分隐私保护模型相结合可实现不同水平的隐私保护效果, 也被逐渐被应用到隐私保护领域.

文献 [24] 提出用 Laplace 干扰算法 (Laplace Perturbation Algorithm, LPA) 给原数据添加 suitably-chosen 噪音. 噪音依据 Laplace 分布产生. Laplace 分布的 PDF 公式如下:

$$Pr=(Lap(b)=Z)=\frac{1}{2b}\mathrm{e}^{-\varepsilon|Z|/b},$$

其中 $Lap(b)$ 是服从均值为 0, 方差为 $2b^2$ 的 Laplace 分布的随机变量.

LPA 算法计算和输出 $\widetilde{m}=m+Lap(b)$. 若 $b=\Delta_1(m)/\varepsilon$, 则 $\mathrm{LPA}(m,\varepsilon)$ 满足 ε-差分隐私, 其中 $\Delta_1(m)$ 是请求电量数据的灵敏度, 即改变任意一次请求的电量数据对总数据 m 所造成的 L_1 距离[18]的最大误差, 其表达式如下:

$$|m-m'|\leqslant\Delta_1(m),$$

其中 $m'\in nbrs(m)$.

2.4 二项分布

在概率论和统计学中, 二项分布是 n 个独立的是/非试验中成功次数的离散概率分布, 记为 $X\sim(n,p)$, 其中 X 表示实验结果, n 为独立重复实验的次数, p 为每次试验成功的概率.

如果事件发生的概率是 p, 则不发生的概率为 $q=1-p$, N 次独立重复试验中发生 k 次的概率是:

$$P=(X=k)=C_n^k p^k q^{n-k}, k=0,1,\cdots,n. \tag{1}$$

最多发生 k 次的概率是:

$$F=(X=k)=P(X\leqslant k)=\sum_{j=0}^{k} C_n^j p^j q^{n-j}, k=0,1,\cdots,n. \tag{2}$$

3　系统模型

系统模型, 如图 1 所示, 其主要包括 3 种类型的参与者: 智能电表、用户和电力公司. 下面依次介绍它们在系统模型中的功能和作用.

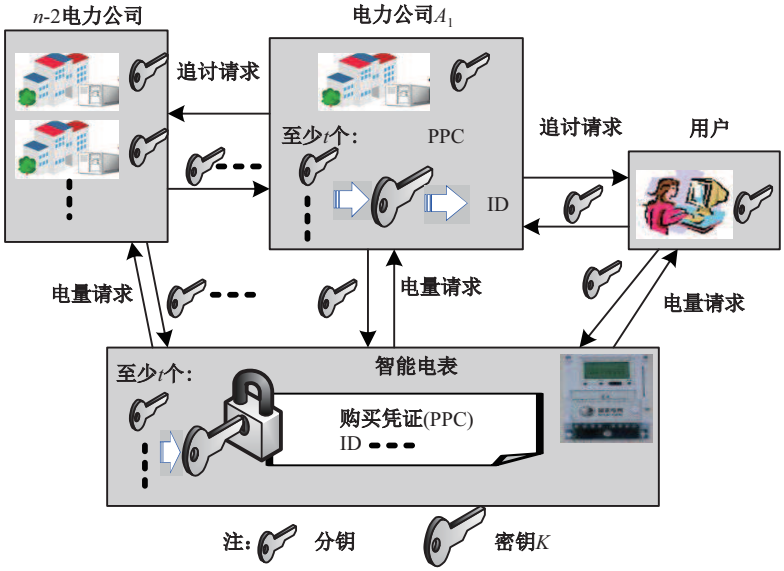

图 1　系统模型

Fig. 1　System mode

智能电表: 智能电表生成密钥 K, 并将密钥 K 划分为 n 个分钥分发给 n 个参与者(即 $n-1$ 个电力公司和一个用户), 最后删除密钥 K 及其相关的关键信息, 如分钥. 智能电表向电力公司请求电量之前, 智能电表需进行两方面工作: ①基于Shamir(t,n) 密钥共享方案生成购买凭证 (Power Purchase Credential, PPC): 智能电表至少需向 t 个参与者发送电量请求索要分钥, 并根据得到的分钥进行密钥 K 的恢复, 然后利用密钥 K 加密自己的身份 ID 生成 PPC; ②基于 Laplace 干扰算法干扰请求的电量数据: 智能电表利用 4 个高斯随机变量生成 Laplace 噪音干扰请求的电量数据, 然后将干扰后的电量数据发送给电力公司, 并根据 D-H 协议将 Laplace 噪音随机数秘密地传送给电力公司实现真实电量的计算.

用户: 用户是智能电表的使用者. 假设用户身份 ID 与智能电表身份 ID 一致. 用户秘密地保存一个分钥, 根据智能电表或电力公司的请求提交分钥.

电力公司: 假设有 $n-1$ 个电力公司. 电力公司应能够对未按时缴费的用户进行电费追讨. 智能电表(用户)凭借 PPC 向电力公司购得电量. 电力公司没有密钥 K 无法解密 PPC 获

得智能电表身份 ID. 假设电力公司 A_1 要追讨电费, 它需先获得密钥 K: A_1 持有一个分钥, 它至少需向 $t-1$ 个参与者发送追讨请求索要分钥, 根据得到的分钥恢复密钥 K 并解密 PPC 获得智能电表身份 ID, 然后根据 ID 追讨电费.

攻击模型

(1) 内部攻击: 允许用户和电力公司都可以是恶意的. 恶意的用户可能妨碍电力公司追讨电费, 其类型有以下两种. a. Liar, 发送错误分钥给电力公司的用户; b. Rejecter, 拒绝发送分钥给电力公司的用户. 恶意的电力公司可能是 Collaborator, 它可能与其它恶意的电力公司相勾结意图获得密钥 K. 我们假设绝大多数电力公司是可信的, 这种假设也符合实际情况.

(2) 外部攻击: 外部攻击者可以分为以下两种类型. a. Eavesdropper/Interceptor, 在传输过程中窃听或截取智能电表实时请求的电量信息意图窥探用户隐私的攻击者; b. Intruder, 意图攻击参与者(如智能电表)获得关键性信息(如密钥 K)的攻击者.

隐私目的: 实现电力公司分时电价计费的同时保护用户隐私不被电力公司窥探; 智能电表实时请求的电量数据在传输过程中保持一定的隐私水平防止被窃听或截取而泄露用户隐私.

4 方案

方案分为 4 个阶段: 注册阶段、电量请求阶段、电费追讨阶段和密钥 K 更新阶段. 下面依次详细描述各个阶段所做的工作, 其中所用字符含义如表 1 所示.

4.1 注册阶段

(1) 电力公司为每个智能电表分配一个身份 ID. 每个参与者生成自己的公钥/私钥对(公钥/私钥对基于公钥加密算法, 在此不作详细阐述), 并将自己的公钥公布给其他参与者.

(2) 智能电表随机生成一个 $t-1$ 次多项式, $f(x)=(K+a_1x+a_2x^2+\cdots+a_tx^{t-1})\bmod Q$, 其中 $a_1,a_2,\cdots,K\epsilon F_Q,Q(Q>K)$ 为大质数, F_Q 为一个有限域. 智能电表记录密钥 K 生成时间 t_0. 然后, 智能电表从多项式 $f(x)$ 曲线上选择 n 个点, 即 $(x_i,y_i)(x_i\neq 0,1\leqslant i\leqslant n,y_i=f(x_i))$ 作为分钥. 然后, 智能电表用各个参与者的公钥加密分钥以及密钥 K 生成时间 t_0, 签名[28]之后分发给各个参与者. 最后, 智能电表删除密钥K及与其相关的一些关键性信息(如多项式 $f(x)$, 分钥 (x_i,y_i) 等)并秘密保存密钥 K 的生成时间 t_0.

$$M\rightarrow A_i/U:$$
$$E_{Pub_{A_i/U}}((x_i,y_i)|t_0)|Sig_{Pri_M}(E_{Pub_{A_i/U}}((x_i,y_i)|t_0)).$$

(3) 电力公司和用户收到消息后, 它们首先用存有的智能电表公钥验证智能电表的签名. 成功后, 它们各自用自己的私钥解密消息得到各自的分钥和密钥 K 生成时间并秘密保存.

表 1 字符含义

Tab. 1 The definition of characters

字符	含义
M	智能电表
A_i	第 i 个电力公司
U	用户
K	Shamir(t,n) 门限密钥共享方案的密钥
Pub_R/Pri_R	实体 R 的公钥/私钥
A_i/U	表示电力公司或用户
$E_{Pub_R}(J)$	表示用实体 R 的公钥加密信息 J
$D_W(J)$	表示用密钥 W 解密信息 J
Sig_{Pri_R}	表示实体 R 的签名
$J\|Q$	表示信息 J 和信息 Q 的连接, 无实质意义

4.2　电量请求阶段

电量请求阶段分为两个过程: PPC 生成和 Laplace 噪音干扰.

4.2.1　PPC 生成

智能电表每次请求电量之前, 需要先获得 PPC. 下面是 PPC 生成的详细步骤.

(1) 智能电表 M 随机地从 n 个参与者(即 $n-1$ 个电力公司和一个用户)中选择 t 个参与者, 并将电量请求 (power request, PRE) 广播给它们进行分钥的索要. PRE 主要包含密钥 K 生成时间 t_0.

$$M \to A_i/U:$$
$$E_{Pub_{A_i/U}}(\mathrm{PRE})|Sig_{Pri_M}(E_{Pub_{A_i/U}}(\mathrm{PRE})).$$

(2) 当 t 个参与者收到广播后, 它们首先验证签名. 成功后, 它们分别用智能电表公钥加密各自的分钥并附上签名后发送给智能电表.

$$A_i/U \to M:$$
$$E_{Pub_M}((x_i,y_i)|t_0)|Sig_{Pub_{A_i/U}}(E_{Pub_M}((x_i,y_i)|t_0)).$$

(3) 当智能电表收到参与者的消息后, 它首先验证签名. 成功后, 智能电表用自己的私钥解密消息得到 t 个分钥并根据多项式插值方法重构多项式 $f(x)$, 密钥 K 即为 $x=0$ 时多项式 $f(x)$ 的值(本文利用拉格朗日插值进行了实现). 然后智能电表用恢复的密钥 K 加密自己的身份 ID 并附上 t_0 得到 PPC, 即 $E_K(ID)|t_0$.

4.2.2　Laplace 噪音干扰

接着智能电表要对请求的电量数据进行 Laplace 噪音干扰. LPA 需要计算 $\widetilde{m}=m+Lap(b)$: 其中 $Lap(b)$ 是服从均值为 0, 尺度参数为 b 的 Laplace 分布的随机变量. 定义 $N(\mu,\sigma)$ 是服从均值为 μ, 方差为 σ^2 的高斯随机变量. 本文利用以下性能(证明见完全版 [22])生成 Laplace 噪音随机数.

引理 1　设 $Y_j \sim N(0,b)(j=1,2,3,4)$ 是高斯随机变量, 则 $Z=Y_1^2+Y_2^2-Y_3^2-Y_4^2$ 是服从 $Lap(2b^2)$ 的随机变量. 此性能将 D-H 协议次数由 $O(4)$ 减少到 $O(1)$, 提高了效率. 下面是 Laplace 噪音干扰的详细步骤.

(1) 假设智能电表每隔特定时间(如 15 s 或 15 min)请求一次电量, 每 4 次请求为一个循环, 本文以智能电表每 15 s 请求一次电量, 即 1 min 为一个循环为例作说明. 假设每次请求的电量为 $m_j, j=1,2,3,4$. 智能电表每次随机地生成一个服从 $N(0,\sqrt{b/2})$ 分布的随机数 $y_j, j=1,2,3,4$, 然后计算 $\widetilde{m}_j=m_j+y_j^2$. 假设智能电表向电力公司 A_1 请求电量. 前 3 次请求电量时, 智能电表用电力公司 A_1 的公钥加密 $\widetilde{m}_j(j=1,2,3)$ 和 PPC($E_K(ID)|t_0$), 附上签名后发送给电力公司 A_1.

$$M \to A_1:$$
$$E_{Pub_{A_1}}(E_K(ID)|t_0|\widetilde{m}_j)|Sig_{Pri_M}(E_{Pub_{A_1}}(E_K(ID)|t_0|\widetilde{m}_j)).$$

①当智能电表第 4 次 ($j=4$) 请求电量时, 智能电表选择一个素数 p 及其整数模 n 乘法群原根 g, 并生成一个秘密整数 a, 计算 $A=g^a \bmod p$. 智能电表用电力公司 A_1 的公钥加密 p、g、A 和 PPC, 并附上自己的签名发送给电力公司 A_1.

$$M \to A_1:$$

$$E_{Pub_{A_1}}(p|g|A|E_K(ID)|t_0)|Sig_{Pri_M}(p|g|A|E_K(ID)|t_0).$$

②电力公司 A_1 收到消息后先验证签名. 成功后, 电力公司 A_1 用自己的私钥解密消息获得 p、g、A. 然后电力公司 A_1 随机地生成一个秘密整数 b, 计算 $B = g^a \bmod p$ 及密钥 $s = A^c \bmod p$. 随后电力公司 A_1 用智能电表的公钥加密 B 和 PPC, 并附上自己的签名发送给智能电表.

$$A_1 \to M:$$

$$E_{Pub_M}(B|E_K(ID)|t_0)|Sig_{Pri_{A_1}}(E_{Pub_M}(B|E_K(ID)|t_0)).$$

③智能电表收到消息后先验证签名和 PPC 的正确性. 成功后, 智能电表计算密钥 $s = B^a \bmod p$ 及 $Lap(b)$ 的噪音随机数 $z = y_1^2 + y_2^2 - y_3^2 - y_4^2$. 智能电表用密钥 s 加密噪音随机数 z, 即 $E_s(z)$. 然后, 智能电表用电力公司 A_1 的公钥加密 PPC、电量 $\widetilde{m}_4$ 和 $E_S(z)$, 并附上签名发送给电力公司 A_1.

$$M \to A_1:$$

$$E_{Pub_{A1}}(E_K(ID)|t_0|\widetilde{m}_4|E_S(z))|Sig_{Pri_M}(E_{Pub_{A_1}}(E_K(ID)|t_0|\widetilde{m}_4|E_S(z))).$$

(2) 电力公司 A_1 收到请求的电量消息后先验证签名. 成功后, 电力公司 A_1 用自己的私钥解密消息获得 PPC 和干扰后电量 $\widetilde{m}_j (j = 1, 2, 3, 4)$. 然后电力公司 A_1 计算 $\widetilde{m} = \sum_{j=1}^{4} \widetilde{m}_j$, 并用密钥 s 解密 $E_s(z)$ 获得 Laplace 噪音随机数 z. 接着电力公司 A_1 作 $(\widetilde{m} - z)$ 计算获得智能电表一分钟内请求的真实电量 m. 电力公司 A_1 虽然知道智能电表请求的实时(即每分钟)电量 m, 但电力公司没有密钥 K 无法解密 PPC 获得智能电表身份 ID, 所以电力公司 A_1 无法窥探用户隐私.

(3) 在一定的时期(如一个月或两个月), 智能电表将总的电量及 PPC 加密并附上签名发送给电力公司 A_1.

$$M \to A_1:$$

$$E_{Pub_{A_1}}(m_{total}|E_K(ID)|t_0)|Sig_{Pri_M}(E_{Pub_{A_1}}(m_{total}|E_K(ID)|t_0)).$$

(4) 电力公司 A_1 收到消息后先验证签名. 成功后, 电力公司 A_1 用自己的私钥解密消息获得 m_{total}、$E_K(ID)|t_0$. 电力公司将标有同样 PPC 的实时电量 m 进行加和得到 m'_{total}, 并比较 m'_{total} 与 m_{total} 是否相等. 若两者相等, 电力公司 A_1 根据该智能电表实时请求的电量进行分时电价计费; 若两者相差较大, 电力公司则增加额外收费作为惩罚. 另外, 电力公司可以将干扰后的电量数据向外公布, 以供其他部门(如发电部门)参考使用.

4.3 电费追讨阶段

如果用户未在规定的时间(如一个月或两个月)上交电费, 电力公司要向此用户追讨电费. 电力公司需先获得密钥 K. 假设电力公司 A_1 要追讨电费. 因电力公司 A_1 仅有一个分钥, 它至少需要向 $t-1$ 个参与者索要分钥来恢复密钥 K. 下面是此阶段的详细步骤.

(1) 电力公司 A_1 随机地选择 $t-1$ 个参与者, 并广播追讨请求 (dunning request, DRE) 给它们. DRE 主要包含要追讨用户的 PPC.

$$A_1 \to A_i/U : E_{Pub_{A_i/U}}(\mathrm{DRE})|Sig_{Pri_{A_1}}(E_{Pub_{A_i/U}}(\mathrm{DRE})).$$

(2) 参与者们收到消息后, 它们先验证签名. 成功后, 参与者们将附有 t_0 的分钥并加上它们的签名发送给电力公司 A_1.

$$A_i/U \to A_1:$$
$$E_{Pub_{A_1}}((x_i,y_i)|t_0)|Sig_{Pri_{A_i/U}}(E_{Pub_{A_1}}((x_i,y_i)|t_0)).$$

(3) 电力公司 A_1 收到消息后, 它先验证签名. 成功后, 电力公司 A_1 用自己的私钥解密消息得到 t 个分钥, 并恢复密钥 K. 电力公司 A_1 用密钥 K 解密 PPC 获得智能电表的身份 ID. 然后, 电力公司 A_1 根据 ID 进行电费的追讨.

然而, 经过电费的追讨电力公司 A_1 得知了智能电表的身份, 它可能将智能电表身份 ID 和该智能电表实时请求的电量信息联系起来窥探用户隐私. 解决这个问题的方法只需要更新密钥 K(见 5.4 节).

4.4 密钥 K 更新阶段

用密钥 K 加密智能电表身份 ID 生成 PPC 的核心部分, 即 $E_k(ID)$, 更新密钥 K 相当于更新 PPC 则电力公司 A_1 无法将身份 ID 或旧的 PPC 与新的 PPC 相关联. 密钥 K 的更新只需重新进行注册阶段的 (2) 和 (3) 步骤即可.

5 安全性分析

本节根据第 3 节的攻击模型, 从内部攻击和外部攻击两个方面进行安全性分析.

5.1 内部攻击

在电费追讨阶段, 电力公司至少需要向 $t-1$ 个参与者索要分钥来解密 PPC(即 $E_K(ID)|t_0$) 获得智能电表身份 ID. 恶意的用户可能拒交分钥 (Rejecter) 或上交错误的分钥 (Liar) 妨碍电力公司追讨电费. 但智能电表可以向其它 $n-2$ 个电力公司索要分钥恢复密钥 K 实现电费的追讨.

恶意的电力公司之间可能相互勾结 (Collaborators) 意图根据已掌握的分钥恢复密钥 K 解密 PPC 获得智能电表身份 ID, 然后关联已掌握的实时电量信息窥探用户隐私. 恢复密钥 K 至少需要 t 个分钥, 所以只要恶意电力公司的数目少于 t, 它们的意图就无法实现. 在实际中绝大多数的电力公司是可信的, 这一攻击成功率很小.

5.2 外部攻击

Eavesdropper/Interceptor 意图通过窃听或截取实时电量信息窥探用户隐私. 智能电表实时请求的电量信息在传输之前进行了 Laplace 噪音干扰, 传输中的电量信息已具有 ε-差分隐私水平, 即使攻击者窃取了电量信息, 它们也无法窥探用户隐私.

Intruder 意图攻击实体参与者获取关于密钥 K 的关键信息以恢复密钥 K. 智能电表生成密钥 K 和分钥, 并将分钥分发给各个参与者之后删除密钥 K 及其有关的关键信息, 如多项式 $f(x)$, 分钥 (x_i,y_i) 等, 并秘密地保存密钥 K 生成时间 t_0. 即使 Intruder 攻击智能电表获得密钥 K 生成时间 t_0, 它也无法得到密钥 K. 另外, 智能电表不存储 PPC, 所以 Intruder 也无法通过攻击智能电表获取 PPC 而冒充智能电表购电. 用户和电力公司的分钥 (x_i,y_i) 均被秘密地保存, Intruder 很难通过攻击它们获取分钥.

6 实验

6.1 最优门限值 t 的确定

本节权衡购买凭证生成效率及密钥 K 安全性确定最优门限值 t. 购买凭证生成效率取决于两个关键因素: 一是基于 Shamir 门限密钥共享方案的密钥 K 恢复时间, 其与门限值 t 有关; 二

是基于对称加密算法的加密时间, 其与门限值 t 无关. 因此, 最优门限值 t 的确定只需考虑密钥 K 恢复时间及密钥 K 安全性.

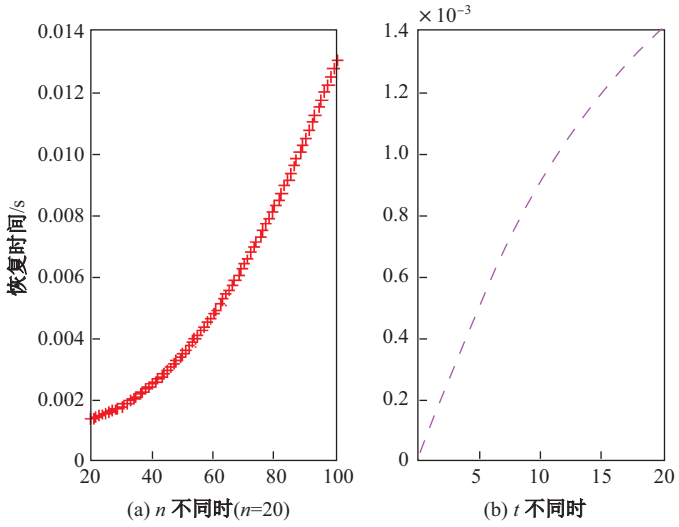

图 2 密钥K恢复的时间

Fig. 2 The recovery time of K

图 2(a) 是 $n(t=20)$ 取不同值时, 密钥 K 恢复的时间情况(其中 n 表示所参与的电力公司和用户的总数目, 即一个用户和 $n-1$ 个电力公司, t 表示门限值, 为了说明问题我们将 n 最高设为 100). 从图 2(a) 可以看出, 密钥 K 恢复的时间随着 n 的增大而变长. 所以, 方案设定智能电表(生成 PPC 或更新密钥 K)或电力公司(追讨电费)选择 t 个分钥而不是 $n(n>t)$ 个分钥来恢复密钥 K, 以降低恢复密钥 K 的时间花销提高效率.

图2(b)是 t 取不同值时, 密钥 K 恢复的时间情况. 从图 2(b) 可以看出, 密钥 K 恢复时间随着门限 t 的增大而变长. 单从时间考虑, 门限 t 越小越好, 但若考虑密钥 K 的安全性, 门限值 t 并非越小越好, 如图 3 所示. 文献 [23] 利用数互补判断矩阵排序算法分析了 Shamir 密钥共享方案中 $t\backslash n$ 值的选择, 但没有量化分析 Shamir 密钥共享方案的安全性. 我们引入二项分布量化分析密钥 K 的安全性. 实验如下:

在 3.4 节已提到, 二项分布标记为 $B(n,p)$. 这里假设 n 表示所参与的电力公司和用户的总数目, p 为一个分钥泄露的概率, 是以分钥持有者(用户)的信誉度和智能电表的坚强度综合考量的量化分析, 其中用户的信誉度主要依据银行信誉度, 智能电表的坚强度主要依据制造商的信誉度及顾客的反馈. 依据 Shamir(t,n) 门限密钥共享方案的门限特性, 少于 t 个分钥泄露不会威胁密钥 K 的安全, 因此, 协议安全度的关系式如下框所示:

图 3 是 $n=20$ 时, 密钥 K 的安全性情况. 从图 3 可以看出: 当 n 和 p 一定时, 安全度随着 t 的增大先为 0 保持不变然后迅速上升达到最大值之后保持不变; 当 n 一定并且 p 很小时, 选择很小的 t 值就能达到很高的安全度. 如图 3 所示, 当 $p=0.5$ 时, 门限 $t=17$ 时安全度已达最高(约为 100%, 虽然实际情况不能达到此安全度, 但能说明情况), 所以 $t=17$ 是 $n=20$、$p=0.5$ 条件下的最优门限值. 由图 3 还可以得出, 当 n 一定时, 在 t 未达到所对应 p 的最优门限值之前, 安全度随着 p 的减小而提高. 另外从图 2(b) 可以得出, 当 $t=17$ 时密钥 K 恢复所需时间约为 1.3×10^{-3} s. 此外, 在能确保分钥泄露概率 p 很小的情况下, 选择较小的门限值 t 可达到较高安全性同时也能提高效率. 例如, 当 $p=0.01$ 时, $t=3$ 时的安全度已达最高(见图 3), 此时密钥 K

恢复时间仅约为 0.25×10^{-3} s(见图 2(b)).

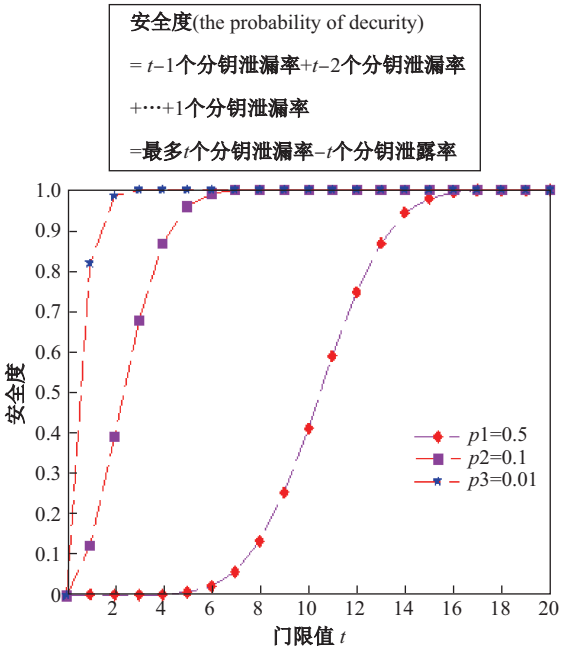

图 3　密钥K的安全性

Fig. 3　The security of K

6.2　加解密时间测试

实验采用 ThinkPad Core 2 CPU, E425 @1.90GHz, C 语言编程实现 1 024 位 RSA 公钥加密算法以及 64 位 DES 对称加密算法. 实验以每次加解密最长的信息为例, 测试不同电量请求时间下的加解密时间, 实验结果如图 4 所示. 从图 4 可以看出: DES 加解密时间相对 RSA 加密时间很短, 可以忽略不计; RSA 加解密时间随着电量请求的时间增加而线性增加; RSA 加解密时间相对于有效时间的比例基本是定值且很小, 仅约为 0.46%. 例如(见图 4), 智能电表在 12 h 内请求电量, 加解密耗时约 200 s, 仅占有效时间 12 h 的 0.46%.

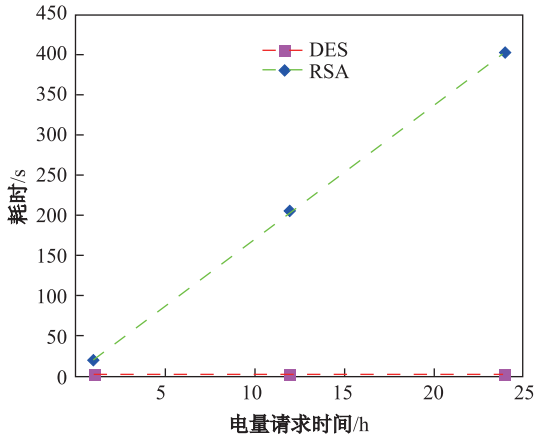

图 4　加解密时间

Fig. 4　Encrption and decrptin time

6.3 Laplace噪音干扰效果验证

为了验证Laplace噪音干扰对电量数据有 ε-差分隐私保护, 假设一组智能电表在 20 min 内实时请求的电量数据, 该数据在 0~1 kwh 范围内变动且随时间线性递增. 不失一般性, 假设差分隐私水平 $\varepsilon = 1$, 请求电量数据的灵敏度 $\Delta_1(m) = 1$ kwh. 为了满足 ε-差分隐私, Laplace 噪音随机数服从 分布. 图 5 是干扰前后数据的对比. 从图 5 可以看出, 干扰前后数据相差很大, 攻击者不可能根据干扰后的数据推测出智能电表用户的用电规律(即随时间线性增加), 验证了 Laplace 噪音干扰的 ε-差分隐私保护效果. 另外, 从图 5 可以看出干扰后的数据实用性不高, 电力公司可以在发布数据之前, 对干扰后的数据作 Fourier Perturbation Algorithm (FPA_k)[18]变换提高干扰后数据的实用性.

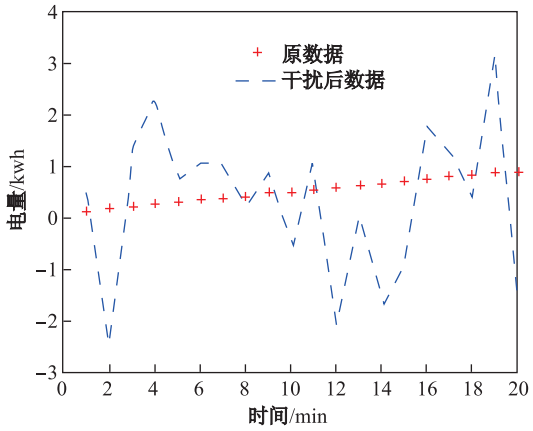

图 5 干扰前后数据的对比

Fig. 5 The comprison between the original data and the disturbed data

6.4 可行性比较

在文献 [1] 中, 智能电表生成环签名并凭借环签名[27]实时请求电量. 利用哈希函数 MD5 对文献 [1] 方案中环签名进行了实现, 并在不同电量请求时间下对环签名的生成耗时与本文方案中购买凭证 PPC 的生成耗时作了比较, 结果如图 6 所示. 从图 6 可以看出: 环签名生成所需的时

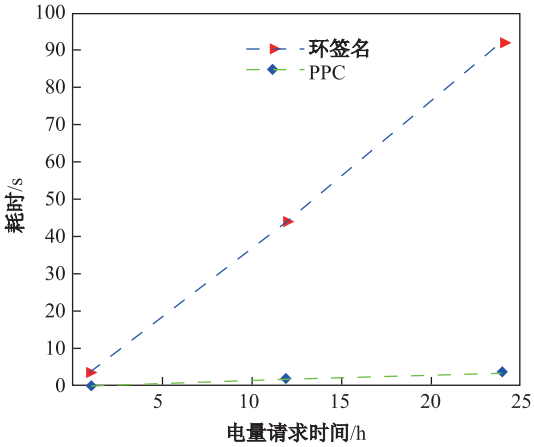

图 6 环签名和PPC生成耗时的对比

Fig. 6 The comparison of generation time between ring signature and PPC

间随着电量请求时间的增加而呈线性增长, 其相对于有效时间的比值基本为定值约为 0.1%; 而 PPC 生成所需的时间相对于有效时间的比值也基本为定值仅约为 0.04%. 可见 PPC 生成效率远大于环签名生成效率, 本文所提出的电量请求方案有很好的可行性.

7 总　　结

本文基于Shamir(t,n)门限密钥共享方案匿名智能电表身份 ID 生成购买凭证 PPC, 智能电表凭借 PPC 实时向电力公司请求电量. 电力公司依据智能电表请求的电量信息及 PPC 进行分时电价的计费, 其不知晓智能电表身份 ID 无法窥探用户隐私. 在一定条件下电力公司可恢复密钥 K 解密 PPC 获得智能电表身份 ID 实现电费的可追溯性. 利用 Laplace 噪音干扰智能电表实时请求的电量数据, 进一步降低电量数据在传输中被窃取而造成的用户隐私泄露风险. 实验从以下 4 个方面验证了提出方案的有效性和可行性: 安全性的定量分析及最优门限值 t 的确定、时间效率的测试分析、Laplace 噪音干扰的ε-差分隐私保护效果的验证分析以及方案可行性的比较. 接下来的工作是提高 PPC 生成效率以及进一步提高分钥的安全性, 例如, 如何使电力公司在不知晓自己分钥机密信息的情况下实现电费的可追溯性.

[参　考　文　献]

[1] YUC M, CHEN C Y, KUO S Y, et al. Privacy-preserving power request in smart grid networks[J]. IEEE Systems Journal, 2014, 8(2): 441-449.

[2] LI F J, LUO B, LIU P. Secure information aggregation for smart grids using homomorphic encryption[C]//Proc of the SmartGridComm. Gaithersburg. MD: IEEE, 2010: 327-332.

[3] GENTRY C. A fully homomorphic encryption scheme[D]. Palo Alto: Stanford University, 2009.

[4] VARODAYAN D, KHISTI A. Smart meter privacy using a rechargeable battery: Minimizing the rate of information leakage[C]// Proc of the Acoustics, Speech, and Signal Processing. Prague: IEEE, 2011: 1932-1935.

[5] KALOGRIDIS G, EFTHYMIOU C, DENIC S, et al. Privacy for smart meters: towards undetectable appliance load signatures[C]// Proc of the SmartGridComm. Gaithersburg, MD: IEEE, 2010, 4(6): 232-237.

[6] CHEUNG J C L, CHIM T W, YIU S M, et al. Credential-based privacy-preserving power request scheme for smart grid network[C]// Proc of the Global Telecommunications Conference. Houston: IEEE, 2011, 5(9): 1-5.

[7] EFTHYMIOU C, KALOGRIDIS G. Smart grid privacy via anonymization of smart metering data[C]// Proc of the SmartGridComm. Gaithersburg, MD: IEEE, 2010, 4(6): 238-243.

[8] MARKHAM M M, SHENOY P, FU F, et al. Private memoirs of a smart meter[C]// Proc of the Embedded Systems for Energy-Efficient Buildings. Zurich: ACM, 2010.

[9] GOLDWASSER S, MICALI S, RACKOFF C. The knowledge complexity of interactive proof-systems[J]. SIAM Journal of Computing, 1989.

[10] CHIM T W, YIU S M, LUCAS C K, et al. PASS: Privacy-preserving authentication scheme for smart grid network[C]// Proc of the SmartGridComm. Brussels: IEEE, 2011: 196-201.

[11] KIM Y S, HEO J. Device authentication protocol for smart grid systems using homomorphic hash [J]. Communications and Networks, 2012, 14(6): 606-613.

[12] LEE W B, CHEN T H, SUN W R, et al. An s/key-like one-time password authentication scheme using smart cards for smart meter[C]// Proc of the Advanced Information Networking and Applications Workshops. Victoria: IEEE, 2014: 281-286.

[13] LEE S, BONG J, SHIN S, et al. A security mechanism of smart grid ami network through smart device mutual authentication[C]// Proc of the Computer Communications Workshops. Phuket: IEEE, 2014: 592-595.

[14] FOUDA M M, FADLULLAH Z M, KATA N, et al. A lightweight message authentication scheme for smart grid communications[J]. IEEE Transactions on Smart Grid, 2011, 2(4): 675-685.

[15] FOOUDA M M, FADLULLAH Z M, KATA N, et al. Towards a light-weight message authentication mechanism tailored for smart grid communications[C]// Proc of the Information Networking. Shanghai: IEEE, 2011: 1018-1023.

[16] KAKALI C, ASOK D, DAYA G. Mutual authentication protocol using hyperelliptic curve cryptosystem in constrained devices[J]. International Journal of Network Security, 2013, 15(1): 9-15.

[17] RIHM A, HEBA A, SALWA H E. New real time multicast authentication protocol[J]. International Journal of Network Security, 2011, 12(1): 13-20.

[18] RASTAGI V, NATH S. Differentially private aggregation of distributed time-series with transformation and encryption[C]// Proc of the Management of data. Indiana: ACM, 2010: 6-11

[19] SARATHY R, MURALIDHAR K. Evaluating laplace noise addition to satisfy differential privacy for numeric data[J]. Transactions on Data Privacy, 2011, 4(1): 1-17.

[20] SHAMIR A. How to share a secret[J]. Communications of the ACM, 1979, 22(11): 612-613.

[21] TIAN X X, SHA C F, WANG X L, et al. Privacy preserving query processing on secret share based data storage[C]// Proc of the Database Systems for Advanced Applications. Hong Kong: Springer, 2011: 108-122.

[22] RASTOGI V, NATH S. Differentially private aggregation of distributed time-series with transformation and encryption[R]. Tech Rep MSR-TR-2009-186, Microsoft Research, 2009.

[23] LI Q D, ZHOU Y H. Research and application based on A. Shamir's (t, n) threshold secret sharing scheme[C]// Proc of the Computer Science & Education. Melbourne: IEEE, 2012(6): 14-17.

[24] DWORK C, MCSHERRY F, NISSIM K, et al. Calibrating noise to sensitivity in private data analysis[C]// Proc of the 3rd Theory of Cryptography Conference. New York: Springer, 2006: 265-284.

[25] DWORK C. Differential privacy: A survey of results[C]// Proc of the Theory and Applications of Models of Computation. China: Springer, 2008: 1-19.

[26] 田秀霞, 高明, 王晓玲, 等. 数据库服务–安全与隐私保护 [J]. 软件学报, 2010, 21(5): 991-1006.

[27] CHAUMM D. Blind signatures for untraceable payments[C]// Proc of the Advances in Cryptology. USA: Springer, 1982: 199-203.

[28] 张明武, 杨波, 祝胜林. 可信模块隐私保护的自证明签密方案 [J]. 北京邮电大学学报, 2009, 32(1): 60-64.

[29] LIUY L, JIN Z G. Security enhancement of WAPI access authentication protocol(WAI)[J]. Journal of Harbin Institute of Teehnolo (New Series), 2012, 19(6): 42-46.

(责任编辑: 张 晶)

(上接第 86 页)

[12] WANG X, LI W, CUI Y, et al. Click-through rate estimation for rare events in online advertising[G]//HUA X S, MEI T, HANJALIC A. Online Multimedia Advertising: Techniques and Technologies . Hershey: IGI Global, 2010. doi: 10.4018/978-1-60960-189-8.ch001.

[13] AGARWAL D, BRODER A Z, CHAKRABARTI D, et al. Estimating rates of rare events at multiple resolutions[C]// ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - Kdd. ACM, 2007: 16-25.

[14] AGARWAL D, CHEN B C, ELANGO P. Spatio-temporal models for estimating click-through rate[C]// International Conference on World Wide Web. ACM, 2009: 21-30.

[15] SCHONLAU M. Boosted regression (boosting): An introductory tutorial and a stata plugin[J]. Stata Journal, 2005, 5(3): 330-354.

[16] BURGES C J C. From ranknet to lambdarank to lambdamart: An overview[R]. Microsoft Research Technical Report, 2010.

[17] FANG Y, LIU J. A novel prior-based real-time click through rate prediction model[J]. International Journal of Machine Learning & Cybernetics, 2014, 5(6): 887-895.

[18] FAIN D C, PEDERSEN J O. Sponsored search: A brief history[J]. Bulletin of the American Society for Information Science & Technology, 2010, 32(2): 12-13.

[19] RICHARDSON M, DOMINOWSKA E, RAGNO R. Predicting clicks:estimating the click-through rate for new ads[C]// International Conference on World Wide Web. ACM, 2007: 521-530.

[20] JOACHIMS T, GRANKA L, PAN B, et al. Accurately interpreting clickthrough data as implicit feedback[C]// Proceedings of the 28th Annual International ACM SIGIR, 2005: 154-161.

(责任编辑: 李万会)