

混合云环境下基于椭圆曲线加密的隐私保护模型

陈建辉

(郑州航空工业管理学院 计算机科学与技术系, 河南 郑州 450015)

摘 要: 为了改善混合云环境下隐私保护性能, 确保隐私数据的信息安全, 提出一种基于椭圆曲线加密的隐私保护模型. 在椭圆曲线密码体制下构造混合云环境中隐私保护的算术编码模型, 结合椭圆曲线方程的优化求解过程进行公钥加密的穷尽搜索, 构造隐私保护数据加密的主密钥, 结合身份认证方案建立门限密码体制, 在混合云平台链路层中进行码元频次检测, 构建私钥共享方案, 根据双线性循环移位方法恢复用户私钥, 实现隐私保护. 仿真结果表明, 采用该加密方法进行用户隐私保护, 数据加密的抗攻击性能较好, 可信度较高.

关键词: 混合云环境; 椭圆曲线; 数据加密; 隐私保护

中图分类号: TP391

文献标识码: A

文章编号: 1000—7180(2017)08—0128—05

Privacy Protection Model Based on Elliptic Curve Cryptography Under Hybrid Cloud Environment

CHEN Jian-hui

(Department of Computer Science and Application, Zhengzhou University of Aeronautics,
Zhengzhou 450015, China)

Abstract: In order to improve the performance of privacy protection in hybrid cloud environment, to ensure the information security and privacy data, proposes a privacy protection model based on elliptic curve cryptography. The structural information model of privacy protection arithmetic encoding hybrid cloud environment in elliptic curve cryptography, process optimization combination of elliptic curve equation of public key encryption end the main key structure search, privacy protection data encryption, identity authentication scheme with threshold cryptography, the symbols are detected in the mixed frequency link layer cloud platform, build the private key sharing scheme, bilinear cyclic shift method to restore the user's private key to achieve privacy protection. The simulation results show that the encryption method for user privacy protection, data the anti attack performance of encryption is better, higher credibility.

Key words: hybrid cloud environment; elliptic curve; data encryption; privacy protection

1 引言

随着网络通信技术的快速发展, 不同平台中的移动网络用户以及本地网络用户通过 APP 终端不断上传个人数据到网络数据交换中心, 不同网络终端的数据在云平台中快速传输和大比特序列存储, 形成混合云环境^[1]. 在混合云环境中融合了大量网络用户的个人信息, 包括用户的身份信息、位置信

息、行程轨迹信息以及经济金融信息, 这些信息涉及到了用户的隐私信息安全, 容易遭到攻击和泄露, 隐私泄露将会对网络用户的信息安全带来较大的威胁和危害, 研究混合云环境下的隐私保护模型, 对确保网络信息安全, 保障用户的正常生活秩序方面具有重要意义.

对隐私保护模型建立在对隐私信息的加密和编码设计, 由于混合云环境下隐私保护通信数据组合

结构简单,混合云环境的数据组合具有自组织性和开放性,通常采用整数线性组合方案构建数据结构,容易受到明文攻击和病毒入侵,产生信息泄露,且加密难度较大.传统方法中,对此类数据的加密和隐私保护模型设计主要有基于差分聚类的隐私保护模型^[2]、ECC 加密的隐私保护模型、结合 DCT 变换与 DNA 运算的隐私保护模型等^[3],通过对应的加密方案构造加密解密密钥,对需要隐私保护的数据进行算术编码设计,实现隐私保护,取得了较好信息加密效果,其中,文献[4]中提出一种基于差分隐私的数据匿名化隐私保护方法,采用混沌公钥加密方案构造差分隐私的数据匿名化隐私保护模型,选用 Laplace 数据聚类机制对不同隐私保护预算下的数据进行分类处理,提高数据加密的效率,但该方法对高强度的明文攻击的抗攻击能力不好,容易导致信息泄露;文献[5]中提出一种基于面向数据发布和分析的大数据安全与隐私保护模型,引入 SuLQ 框架设计混合云环境下隐私保护通信数据加密的 Hash 函数,在有限域中构建加密密钥,提高数据加密的深度,但该方法计算开销较大,隐私保护计算的实时性不好.

对此,本文提出一种基于椭圆曲线加密的隐私保护模型.首先在椭圆曲线密码体制下构造混合云环境中隐私保护的信息算术编码模型,结合椭圆曲线方程的优化求解过程进行公钥加密的穷尽搜索,然后构造隐私保护数据加密的主密钥,结合身份认证方案建立门限密码体制,在混合云平台链路层中进行码元频次检测,实现隐私保护.最后进行隐私保护和数据信息加密的仿真实验分析,得出有效性结论.

2 算术编码模型

2.1 椭圆曲线密码体制

本文采用一种基于椭圆曲线密码体制(Elliptic Curve Cryptography, ECC)的加密技术进行混合云环境下的信息隐私保护,首先分析椭圆曲线加密模型^[6],用户数据加密的椭圆曲线方程一般定义为

$$f(x) = \begin{cases} x/p, & x \in [0, p) \\ (1-x)/p, & x \in [p, 1] \end{cases} \quad (1)$$

该椭圆曲线方程是通过参数 p 来调节信息加密的扩展范围,当 $p = 0.5$ 时,该映射为一个标准的正态分布模型.

为便于描述,在椭圆曲线的映射区域,定义待加密的隐私保护信息的比特序列长度为 n 的比二进制

编码序列为 $E = \epsilon_1, \epsilon_2, \dots, \epsilon_n$. 首先基于仿射变换 $x_i = 2\epsilon_i - 1$, 构造出表示为二进制编码比特的 0 和 1 码元,得到在混合云环境下需要隐私保护的新的序列 $X = x_1, x_2, \dots, x_n$, 并统计加密比特序列 X 的二项式和 $S_n = x_1 + x_2 + \dots + x_n$, 采用椭圆曲线 S_n 的极值点 $\frac{S_n}{\sqrt{n}}$, 取属于 E 的一点 G , 而 $\frac{S_n}{\sqrt{n}}$ 接近于标准的正态分布函数, 计算消息的散列值, 当满足 $\lim_{n \rightarrow \infty} P\left(\frac{S_n}{\sqrt{n}} \leq z\right) = \varphi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-u^2/2} du$. 对循环点群 $S_{\text{obs}} = \frac{|S_n|}{\sqrt{n}}$ 而言, 根据正态分布特性得到椭圆曲线加密的收敛函数为

$$\begin{aligned} \lim_{n \rightarrow \infty} P\left(S_{\text{obs}} = \frac{|S_n|}{\sqrt{n}} \leq z\right) \\ = 1 - \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{-z} e^{-u^2/2} du + \int_z^{+\infty} e^{-u^2/2} du \right) \\ = 2\varphi(z) - 1 \end{aligned} \quad (2)$$

可见, ECC 密码体制由四个算法组成: Setup, Extract, Encrypt 和 Decrypt, 在混合云环境下, 隐私保护的加密方案中需要引入基于身份的认证加密体系, 构造包含用户公钥的证书^[7], 从而得到椭圆曲线加密的统计门限值 P -value 为

$$\begin{aligned} P\text{-value} &= 2[1 - \varphi(S_{\text{obs}})] \\ &= 2\left(1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{S_{\text{obs}}} e^{-u^2/2} du\right) \\ &= \frac{2}{\sqrt{2\pi}} \int_{S_{\text{obs}}}^{+\infty} e^{-u^2/2} du \\ &= \frac{2}{\sqrt{\pi}} \int_{\frac{S_{\text{obs}}}{\sqrt{2}}}^{+\infty} e^{-t^2} dt \\ &= \text{erfc}\left(\frac{S_{\text{obs}}}{\sqrt{2}}\right) \end{aligned} \quad (3)$$

若门限 $P\text{-value} \geq 0.01$, 即可认为公钥先于私钥产, 根据上述描述, 采用椭圆曲线密码体制对混合云环境下的隐私保护数据进行二值化编码^[8], 得到编码映射图.

2.2 最佳编码长度计算

在椭圆曲线密码体制下构造混合云环境中隐私保护的算术编码模型, 结合椭圆曲线方程的优化求解过程进行公钥加密的穷尽搜索, 以椭圆曲线方程反函数作为算术编码的传导函数:

$$f^{-1}(I) = \begin{cases} p * I, & s = "0" \\ 1 - (1 - p) * I, & s = "1" \end{cases} \quad (4)$$

式中, I 表示明文序列分布的区间, 初始区间 $I = [0,$

1],采用云存储空间中的隐私保护比特序列进行特征空间重组^[9],对每个加密编码进行密钥构造,其表达式为

$$f(x) = \begin{cases} x/P_1, x \in I_1 \\ (x - P_1)/P_2, x \in I_2 \\ \dots\dots\dots \\ (x - \sum_{i=1}^{n-1} P_i)/P_n, x \in I_n \end{cases} \quad (5)$$

式中, $P_i (i = 1, \dots, n)$ 表示加密编码的明文数出现的概率,在区间 I_i 中进行离散化处理,得到在对应的累积概率区间内的移位密钥的窗口时隙满足:

$$\sum_{i=1}^n P_i = 1 \quad (6)$$

$$I_1 = [0, P_1] \quad (7)$$

$$I_i = [\sum_{j=1}^{i-1} P_j, \sum_{j=1}^i P_j], i = 2, 3, \dots, n \quad (8)$$

对 $f(x)$ 取反函数为

$$f^{-1}(x) = \begin{cases} P_1 x \\ P_2 x + P_1 \\ \dots \\ P_n x + \sum_{i=1}^{n-1} P_i \end{cases} \quad (9)$$

根据压缩编码理论,密钥空间越大,表示密钥的抗攻击能力越强,隐私保护序列 s_i 的自信息量为 $-\log_2(P(s_i))$,计算隐私保护数据的平均互信息函数为

$$H = - \sum_{i=1}^n P_i \log_2(P_i) \quad (10)$$

经过穷尽搜索,序列 $s = \{s_i, i = 1 \dots M \mid s_i \in S\}$ 的游程长度计算为

$$P_n = \frac{1}{M} \text{card} \{s_i \mid s_i = S_n\} \quad (11)$$

式中, $S_n \in S, n = 1 \dots N$. 互为反函数,所以其区间也是反向对应的,最后所得的密文序列与加密密钥互为反函数,可得:

$$I^i = f^{-1}(x)(I^{i+1}) \quad (12)$$

$$\text{size}(I^i) = P_i \text{size}(I^{i+1}) \quad (13)$$

则比特块的长度:

$$\begin{aligned} \text{size}(I^1) &= \prod_{i=1}^M P(s_i \in S) \\ &= \prod_{n=1}^N (P_n)^{\text{card} \{s_i \mid s_i = S_n\}} \\ &= \prod_{n=1}^N (P_n)^{P_n M} \end{aligned} \quad (14)$$

得:

$$-\log_2(\text{size}(I^1)) = - \sum_{n=1}^N P_n M \log_2(P_n)$$

$$= [K/m] \quad (15)$$

在椭圆曲线密码体制下,公钥加密的最佳编码长度为 $[K/m]$.

3 隐私保护算法

3.1 身份认证方案

在上述进行了隐私保护数据加密的算术编码设计和密钥构造处理的基础上,进行混合云环境下的隐私保护模型优化设计,本文提出一种基于椭圆曲线加密的隐私保护模型,对椭圆曲线方程进行优化求解,得到椭圆曲线方程穷尽搜索的迭代方程为

$$X_{n+1} = \mu X_n (1 - X_n), \mu \in (0, 4], X_n \in [0, 1] \quad (16)$$

式中,当参数 $\mu \in [3.569\ 94 \dots, 4]$ 时,椭圆曲线方程具有稳定收敛解,以优化解特征向量作为循环移位密钥^[10],进行公钥加密的穷尽搜索,过程如图 1 所示.

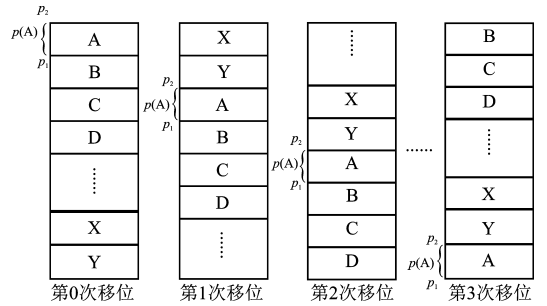


图 1 公钥加密的穷尽搜索过程图

图 1 中,假设隐私保护数据的某一序列的首字符为 A,对长度为 L 的明文密钥,在进行第 0 次循环后,通过穷尽搜索,得到的公钥 μ 值记为 μ_0 ,依次进行最佳公钥搜索,得到最优值为 μ_L ,则:

$$\begin{cases} \mu_0 = (P_{01} + P_{02})/2 \\ \mu_1 = (P_{11} + P_{12})/2 \\ \mu_2 = (P_{21} + P_{22})/2 \\ \dots \\ \mu_L = (P_{L1} + P_{L2})/2 \end{cases} \quad (17)$$

$$\mu = 3.57 + \frac{1}{L} \sum_{i=0}^L \mu_i \quad (18)$$

最后得到的 μ 值作为椭圆曲线加密的迭代参数,构造隐私保护数据加密的主密钥,为

$$\begin{aligned} H_1: F_q \times G_1 &\rightarrow \{0,1\}^n, H_2: \{0,1\}^* \rightarrow G \\ H_3: \{0,1\}^* \times \{0,1\}^* &\rightarrow F_q, H_4: \{0,1\}^n \rightarrow \{0,1\}^n \end{aligned} \quad (19)$$

对于所有的加密哈希函数 $g, h \in G$ 和 $a, b \in F_q$,有 $e(g^a, h^b) = e(g, h)^{ab}$, $e(ag, h) = e(g, h)^a = e(g, ah)$. 令 $H_2(m, T) = rx^{-1}$,若 $Q = e(\gamma, \gamma)^{x^{-1}}$,结

合身份认证方案建立门限密码体制,得到 m_d 在身份 ID_* 下的密文 CT_{ID_*} ,记为

$$CT_{ID_*} = \{C_1 = g^{k_*r} = upk_{*1}^{-1}r, C_2 = C_1^z = upk_{*2}^{-1}r,$$
$$C_3 = m_d Q^{r_0^2(1+u_*(t_*-t_0))} e(\gamma, \gamma)^{r(1+u_*(t_*-t_0)) \frac{f^2(x)-c_0^2}{x}}$$
$$e(\gamma, \gamma)^{r(x+1+u_*(t_*-t_0))f^2(x)},$$
$$C_4 = TQ^{r_0^2(1+u_*(t_*-t_0))} e(\gamma, \gamma)^{r(1+u_*(t_*-t_0)) \frac{f^2(x)-c_0^2}{x}}$$
$$e(\gamma, \gamma)^{r(x+1+u_*(t_*-t_0))f^2(x)}, C_5 = 1\}$$

(20)

式中,对应的解密密钥 $e(\gamma, \gamma)^{r(1+u_*(t_*-t_0)) \frac{f^2(x)-c_0^2}{x}}$ 和代理密钥 $e(\gamma, \gamma)^{r(x+1+u_*(t_*-t_0))f^2(x)}$ 可以通过执行代理重加密的方法计算。

3.2 码元频次检测及隐私保护实现

在结合身份认证方案建立门限密码体制的基础上,在混合云平台链路层中进行码元频次检测,构建私钥共享方案,根据双线性循环移位方法恢复用户私钥,构造隐私保护优化模型^[11],改进模型的实现步骤具体描述如下。

步骤 1:计算混合云计算环境中的隐私保护序列中的每个循环点群 $rsk_{ID_i} = sr_{i1} = g_1^{r-\frac{c}{d}}$,并记录到 $rsk-list$ 列表中;

步骤 2:将长度为 n 的隐私保护比特序列划分为 $N = \lfloor \frac{n}{M} \rfloor$ 个比特块,组成所需的加密序列 $x_1x_2x_3 \cdots x_r$,对比特序列中的连续特征进行码元频次检测,输出密文序列 $c_i = t_i \oplus x_i, i = \{1, 2, \cdots, r\}$ 。

步骤 3:构造服从于标准正态分布的椭圆曲线函数^[12],对每个隐私保护密钥进行更新,在椭圆曲线方程中进行最优值搜索,更新方式如下:

$$\begin{cases} KC_1 = KC_1 \oplus \{t_j, t_{j+1}, t_{j+2}, \cdots, t_{j+m-2}\} \\ KS_1 = KS_1 \oplus \{t_{j+m-2}\} \end{cases}$$

(21)

式中, $j = r \bmod 128$,当 $j > r - m$ 则 $j = j - m$,如果 $j = 0$,则 $j = m$.统计待识别比特序列中的块内频数,椭圆曲线调节参数初始值更新为 $x_0 = x_{\lceil r/16 \rceil + 1}$,根据上述步骤进行剩余隐私数据的加密编码,完成隐私保护数据加密。

4 仿真实验分析

为了测试本文方法在实现混合云环境下用户信息隐私保护中的应用性能,进行仿真实验,试验建立在 Matlab 7 软件环境中,在混合云环境下进行隐私保护数据的原始信息采样,信息采样频率为 20 KHz,需进行隐私保护的文件大小为 $3.8 \times$

10^5 Mbyte,测试训练集分为 200 组,每组的长度为 2 000,密文序列的候选长度为 $m_i = 1.25$,其他参数设定见表 1。

表 1 参量设定

参数设置	$t_{\max}$	M	l	p_m	p_c
参数值	1 024	86	20	0.43	0.21

根据上述仿真环境和参量设定,进行数据加密和隐私保护仿真分析,得到不同测试样本在不同分块长度下满足隐私保护条件的概率测试结果如图 2 所示。

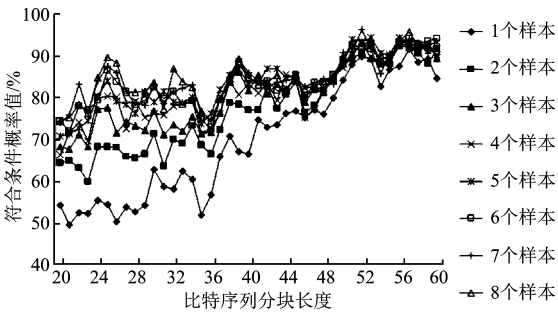


图 2 不同分块长度符合隐私保护要求的概率统计值

分析图 2 结果得知,采用本文方法进行数据加密,通过码元频次检测,提高了满足隐私保护门限要求的比特数.在此基础上,测试不同加密数据比例下隐私保护的信息熵,得到结果如图 3 所示。

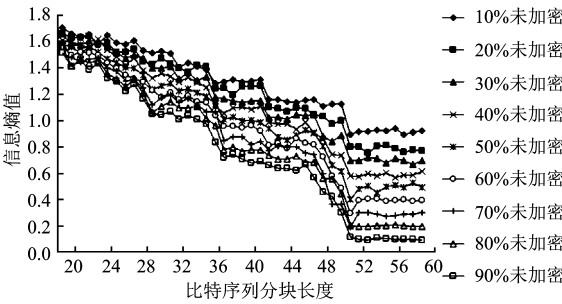


图 3 信息熵计算

分析图 3 结果得知,通过数据加密,随着加密数据的增多,信息熵值增大,在只有 10%数据未加密的情况下,信息熵最高,说明通过本文方法进行数据加密,能提高云环境下数据的信息熵,从而提高隐私保护能力.为了对比不同方法进行隐私保护的可信度对比结果,如图 4 所示。

分析图 4 结果得知,采用本文方法进行隐私保护,随着加密序列的增多,隐私保护的可信度不断增大,但加密比特序列占到 60%时,隐私保护的可信度达到 100%,优于传统方法,提高了隐私安全保护能力。

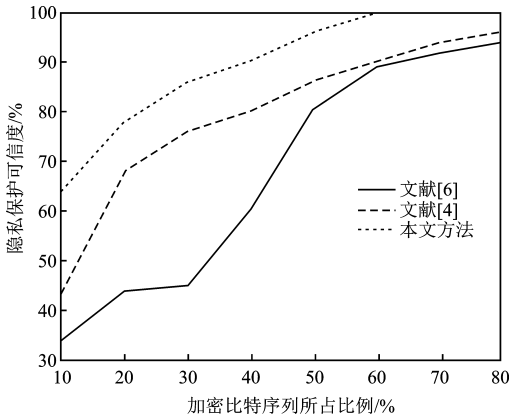


图 4 隐私保护可信度对比

5 结束语

本文研究了混合云环境下的数据加密和隐私保护问题,提出一种基于椭圆曲线加密的隐私保护模型.在椭圆曲线密码体制下构造混合云环境中隐私保护的算术编码模型,结合椭圆曲线方程的优化求解过程进行公钥加密的穷尽搜索,构造隐私保护数据加密的主密钥,结合身份认证方案建立门限密码体制,进行码元频次检测,构建私钥共享方案,根据双线性循环移位方法恢复用户私钥,实现隐私保护.研究得出,采用本文方法进行隐私保护的数据加密抗攻击能力较强,可信度较高,性能较优.

参考文献:

[1] 陈亮, 杨庚, 屠袁飞. 混合云环境下基于属性的密文策略加密方案[J]. 计算机应用, 2016, 36(7): 1822-1827.

[7] 甘明刚, 陈杰, 刘劲, 等. 一种基于三帧差分 and 边缘信息的运动目标检测方法[J]. 电子与信息学报, 2010, 32(4): 894-897.

[8] 徐方明, 卢官明. 基于改进 surendra 背景更新算法的运动目标检测算法[J]. 山西电子技术, 2009, 36(5): 39-40.

[9] Kim K, Chalidabhongse T H, Harwood D, et al. Background modeling and subtraction by codebook construction[C]// International Conference on Image Processing. IEEE, 2004: 3061-3064.

[10] 姜柯, 李艾华, 苏延召. 基于改进码本模型的视频运

[2] 徐小龙, 周静岚, 杨庚. 一种基于数据分割与分级的云存储数据隐私保护机制[J]. 计算机科学, 2013, 40(2): 98-102.

[3] 徐光宪, 徐山强, 郭晓娟, 等. DCT 变换与 DNA 运算相结合的图像压缩加密算法[J]. 激光技术, 2015, 39(6): 806-810.

[4] 付艳艳, 张敏, 冯登国, 等. 基于节点分割的社交网络属性隐私保护[J]. 软件学报, 2014, 25(4): 768-780.

[5] 宋健, 许国艳, 仝荣朋. 基于差分隐私的数据匿名化隐私保护方法[J]. 计算机应用, 2016, 36(10): 2753-2757.

[6] 张啸剑, 孟小峰. 面向数据发布和分析的差分隐私保护[J]. 计算机学报, 2014, 37(4): 927-949.

[7] HAZAY C, NISSIM K. Efficient set operations in the presence of malicious adversaries[J]. Journal of Cryptology, 2012, 25(3): 383-433.

[8] 侯惠芳, 季新生, 刘光强. 异构无线网络中基于标识的匿名认证协议[J]. 通信学报, 2011, 32(5): 153-161.

[9] 熊风光, 李希, 韩燮. 基于整体最小二乘的椭圆拟合方法[J]. 微电子学与计算机, 2017, 34(1): 102-105.

[10] 周彦伟, 杨波, 张文政. 可证安全的移动互联网可信匿名漫游协议[J]. 计算机学报, 2015, 38(4): 733-748.

[11] 肖迪, 邓秘密, 张玉书. 基于压缩感知的鲁棒可分离的密文域水印算法[J]. 电子与信息学报, 2015, 37(5): 1248-1254.

[12] 李永刚, 张治中, 李龙江. 一种 WLAN 与 eHRPD 系统间负载均衡算法[J]. 微电子学与计算机, 2017, 34(1): 44-47.

作者简介:

陈建辉 男, (1978-), 硕士, 副教授. 研究方向为网络及信息安全. E-mail: 17153414@qq.com.

(上接第 127 页)

动目标检测算法[J]. 电子科技大学学报, 2012(6): 932-936.

[11] 杨燕妮, 吴向前, 刘鹏, 等. 基于帧间差分与码本模型的运动车辆检测算法[J]. 新疆大学学报(自然科学版), 2016, 33(2): 203-208.

作者简介:

孙一南 男, (1992-), 硕士研究生. 研究方向为智慧交通下的目标识别与跟踪. E-mail: 18734902579@163.com.

韩晓明 男, (1974-), 硕士生导师. 研究方向为模式识别与故障诊断、智慧交通等.