论 文

一种基于演化密码和 HMM 改进的 Koblitz 安全曲线产生新方法

王潮^{①*}, 张焕国^{②③}, 刘礼黎^④

① 上海大学通信与信息工程学院, 上海 200072

② 武汉大学计算机学院, 武汉 430072

③ 空天信息安全与可信计算教育部重点实验室, 武汉 430072

④ 华为公司, 上海 201206

* 通信作者. E-mail: wangchao@staff.shu.edu.cn

收稿日期: 2010-08-25; 接受日期: 2011-06-28

国家自然科学基金 (批准号: 60970006, 60970115, 61272096, 91018008)、空天信息安全与可信计算教育部重点实验室开放基金 (批准号: AISTC2009_04) 和上海市重点学科和科委重点实验室项目 (批准号: S30108, 08DZ2231100) 资助

摘要 考虑未来云计算攻击和量子计算机攻击, 需要储备安全强度更高的 ECC 安全曲线. 利用隐 Markov 模型 (HMM) 预测迹向量解决基点计算难题, 完善基于演化密码思想提出的 Koblitz 安全曲线产生新算法, 完成了 $F(2^{2000})$ 以内 Koblitz 安全曲线的搜索实验, 产生的安全曲线基域的覆盖范围、曲线的规模和产生效率均超过美国 NIST 的公开报道参数. 可提供的安全曲线的基域和基点最高超过 1900 bit, 远超过美国 NIST 公布的 571 bit. 在 NIST 公布的 $F(2^{163}) \sim F(2^{571})$ 范围之间还有新的安全曲线发现. 对产生的安全曲线进行了详细的安全分析, 表明与 NIST 推荐的安全曲线具有相同的安全准则.

关键词 安全曲线选择 演化密码 蚁群算法 Koblitz 曲线 HMM

1 引言

椭圆曲线密码 (elliptic curve cryptography, ECC) 算法的标准化, 大有取代 RSA 的趋势, ECC 已经逐渐被包含到 IEEE, ANSI, ISO 和 NIST 等发布的标准中. 但是由于安全曲线产生方法等原因, 目前各国际标准组织推荐的安全曲线参数较少, 不超过 30 条. 美国 NIST 推荐的 15 条曲线也成为目前工程应用中常选用的椭圆曲线.

1999 年, 张焕国和覃中平教授借鉴生物进化的思想, 将密码学与演化计算结合起来, 提出了演化密码的概念和利用演化密码实现密码设计和分析自动化的方法. 并在演化密码体制, 演化密码芯片, 密码部件 (S 盒、p 置、轮函数) 的设计自动化, Bent 函数的分析与演化设计, 密码的演化分析, 密码协议的演化设计等方面获得实际成功 [1~5].

我们采用演化密码思想加速 ECC 安全曲线选择 [6], 提出一种不同于传统密码研究方法, 初步完成了 $F(2^{800})$ 以内 Koblitz 安全曲线的选择实验 [7], 产生的安全曲线基域的覆盖范围、曲线的规模和产生效率均超过目前国内和美国国家标准与技术研究所 (NIST) 公布的相应参数.

引用格式: 王潮, 张焕国, 刘礼黎. 一种基于演化密码和 HMM 改进的 Koblitz 安全曲线产生新方法. 中国科学: 信息科学, 2013, 43: 322–334

尽管目前对 ECC 的攻击尚未超过 CertiCom 公司提出的第一能级, 但是根据 NIST 的云计算定义, 需要考虑到目前云计算^[8]可能给黑客带来便利的攻击能力; 而 2011 年 5 月加拿大 D-Wave 公司推出的首台 128Qbit 商用量子计算机问世, 则需要考虑量子计算对 ECC 攻击带来新的隐患, 因此需要储备更高安全强度的 ECC 安全曲线.

但是进一步分析发现, ECC 安全曲线基点的选择是困难的, 不利于更高安全强度的 ECC 安全曲线产生. 本文考虑利用隐 Markov 模型 (hidden Markov model, HMM) 预测迹向量解决基点计算难题, 完善现有的基于演化密码思想的 ECC 安全曲线选择方法, 以期获得更高安全强度的安全曲线.

2 安全曲线选择进展与 NIST 推荐的安全曲线

2.1 美国 NIST 推荐的安全曲线

2000 年 2 月 NIST 颁布 FIPS 186-2^[9], 推荐了 15 条安全椭圆曲线. 其中 Koblitz 曲线涉及的域为 $F(2^{163})$, $F(2^{233})$, $F(2^{283})$, $F(2^{409})$, $F(2^{571})$, NIST 同时给出基点参数. NIST 推荐的这 15 条曲线也成为目前工程应用中常选用的椭圆曲线^[10,11].

2.2 安全曲线选择进展

目前安全曲线选择的方法主要有两种: 复乘法 (complex multiplication, CM) 和随机选择法. 从整体上看, 近些年来在安全曲线产生方面没有更好的方法提出, 随机选择法依旧是最为安全但比较费时的方法, 利用传统纯数学理论分析改进产生算法受到较大制约.

对于 CM 方法, 它是根据给定的阶来选取符合此阶的 ECC 曲线, 实现的速度相对较快, 但是这种方法的缺陷是产生的曲线具有某种特定的结构特征, 存在潜在的安全威胁. 因此, 一般都利用随机选择法来选择曲线. 随机选取的椭圆曲线的安全性大多数依赖于曲线的阶的大小, 求阶方法有以下几种: Schoof's algorithm^[12]、SEA (Schoof-Elkies-Atkin) 算法^[13]、Satoh 算法^[14,15]、Fouquet 算法^[16]、SST 算法^[17]、AGM 算法^[18]和 MSST 算法^[19]等. 另外, 以 Koblitz 曲线为代表的一类特征为小素数的特殊椭圆曲线, 其曲线阶的计算存在更为简便的 Weil 定理算法. 这类曲线的安全性主要依赖于曲线阶的大小以及安全基域的搜索.

3 基于演化密码的 Koblitz 安全曲线产生算法需进一步解决的问题分析

我们提出一种基于演化密码的安全曲线选择算法^[6,7], 应用蚁群算法来搜索 Koblitz 曲线的安全基域取得了成功, 表明了演化密码也是可以应用在公钥密码领域.

3.1 Koblitz 曲线选择的蚁群系统设计^[6,7]

蚁群算法是一种新兴的演化计算方法, Dorigo 等^[20]提出蚁群算法后, 于 1996 年在 IEEE 发表了奠基性论文, 建议将蚁群算法用于组合优化问题求解, 蚁群系统的正反馈特征可以解释蚁群寻优的快速性, 分布式计算则避免蚁群在寻优空间中的过早收敛, 而贪婪式搜索模式则有助于在搜索过程的早期阶段就找到可行的问题解答. 目前蚁群算法已经成功应用于城市旅行商问题 (traveling salesman problem, TSP)^[21]和二次分配问题^[22].

基于蚁群的优点, 并参照蚁群算法在 TSP 中的应用, 我们设计出一套适合 Koblitz 曲线安全基域搜索的蚁群模型, 这个模型类似于最大最小蚂蚁系统 (maximum minimum ant system, MMAS) [23].

STEP1: 参数的初始化, 取 $\alpha = 1$, $\beta = 5$, 这里更注重启发信息在探索中的作用. 信息素蒸发因子 $\rho = 0.02$, 实验表明 ρ 取较小的数能够避免蚁群算法过快收敛.

位于点 i 的蚂蚁 k , 根据伪随机比例规则选择点 j 作为下一个访问的点. 根据文献 [23],

$$j = \begin{cases} \arg \max \{\tau_i [\eta_i]^\beta\}, & \text{如果 } q \leq q_0, \\ J, & \text{否则,} \end{cases} \quad (1)$$

其中 q 是均匀分布在区间 $[0, 1]$ 中的一个随机变量, $q_0 (0 \leq q_0 \leq 1)$ 是一个参数.

在 MMAS 中, 只有每次迭代中最优的蚂蚁才被允许在每次迭代后释放信息素. 根据文献 [23], 信息素更新规则如下:

$$\tau_i \leftarrow (1 - \rho)\tau_i + \Delta\tau_i^{\text{best}}, \quad (2)$$

其中 $\rho (0 < \rho \leq 1)$ 是信息素的蒸发率.

STEP2: 每次迭代开始, 将 5 只蚂蚁随机放在不同点. 蚂蚁每到达一个点都要用验证目标函数 - 判定该点对应的阶是否为大素数, 然后按照模型构建的规则选择下一个要到达的点. 每只蚂蚁每走一步都要进行一次局部更新; 每次迭代结束后, 选择一个最优蚂蚁进行一次全局更新.

3.2 实验结果 [6,7]

(1) 产生的安全曲线基域的范围最大超过 700 bit (PC 机耗时 3.5 h), 并另外提供了不同的基点, 远超过美国 NIST 公布的最高 571 bit 和国内公开报道的最高 163 bit; 已经产生 163 bit 以上的安全曲线基域若干. 如对于下面形式的 Koblitz 曲线 $E_1: y^2 + xy = x^3 + x^2 + 1$, 701 bit 基域产生结果如下: k 表示基域, $k = 2^{701}$. n 为基域为 k 的阶,

$n=5260135901548373507240989882880128665550339802823173859498280903068732 \ 1542970808221$
 $136665362775884512269800074472057387507859154454647132730530 \ 677414059685643347943137538780$
 $32816084302756649401756057061240038011$.

(2) 同时也覆盖了美国 NIST 公布的 $F(2^{163}) \sim F(2^{571})$ 5 条 Koblitz 安全曲线, 对比如图 1 和 2 所示.

3.3 需进一步解决的问题分析

图 1 和 2 中给出的计算时间, 仅仅是基域的产生时间. 然而基点的计算时间随着阶的增大而暴涨, 如何加速基点的搜索, 是需要进一步解决的问题.

求解基点时我们需要求解 $z^2 + z + c = 0$ 二次方程, 我们定义“迹”的概念, 用于快速判定方程是否有解. 假设 a 是域上元素, 即 $a \in \text{GF}(2^m)$, 则 α 的迹定义为 $\text{Tr}(\alpha) = \alpha + \alpha^2 + \alpha^{2^2} + \cdots + \alpha^{2^{m-1}}$. 在二进制域中, $\text{Tr}(c)$ 的值不是 0 就是 1. 当 $\text{Tr}(c)$ 值是 0 的时候, 表明方程有解; 当 $\text{Tr}(c)$ 是 1 的时候, 表明方程无解.

通常情况下, 我们采用以下方法计算迹:

INPUT : $a \in \text{GF}(2^m)$
 OUTPUT : $T = \text{Tr}(\alpha)$

m	571
a	0
$p(t)$	$t571+t10+t5+t2+1$
h	4
n	1932268761508629172347675945465 9936721494636648532174993286176 2572575957114478021226813397852 2706711834706712800825351461273 6749740666173119296824216170925 03555733685276673
Gx	0x 69E 202A 15A4 739E 0FEE ECDA 1B17 1AB0 7AC0 8EC3 6369 7701 8FCF 52D2 BF79 7DFD 9E29 19EE A954 1D83 2EE8 E37D DC92 7003 79F5 CBDB F146 8B63 582F ECAB DADA 8AD9 4A21 EF89 4D82 F454
Gy	0x 28C 3E1D 87E3 0527 4B9A C627 9532 E328 BE6A 607C AB08 9FC3 4300 8E22 E46E 9ADD 451D FCBF 8798 E765 A443 40D9 195B 0DBE 98B0 8B95 D737 57AF 5B0E BDD1 CF13 DAD6 DC8A 3D5E A7FF 0A71
Average time	16 min

图 1 演化密码方法 $m=571$ 的安全曲线参数

Figure 1 Our secure curve

	Curve K-571
a	0
$p(t)$	$t571 + t10 + t5 + t2 + 1$
n	193226876150862917234767594546 599367214946366485321749932861 762572575957114478021226813397 852270671183470671280082535146 127367497406661731192968242161 7092503555733685276673
Gx	0x26eb7a859923fbc82189631f8103f e4ac9ca2970012d5d4602480480184 1ca44370958493b205e647da304db4 ceb08cbbd1ba39494776fb988b4717 4dca88c7e2945283a01c8972
Gy	0x349dc807f4fbf374f4aeade3bca953 14dd58cec9f307a54ffc61efc006d8a2 c9d4979c0ac44aea74fbbbbb9f772ae dcb620b01a7ba7af1b320430c85919 84f601cd4c143ef1c7a3

图 2 美国 NIST 公布的 $m=571$ 安全曲线参数

Figure 2 Secure curve recommended by NIST

1. $T \leftarrow \alpha$

2. For I from 1 to $m-1$

$T \leftarrow T^2 + \alpha$

3. Return T

更有效的方法是引入一个 m 维向量 $\text{Trace_Vector} = \{t_{m-1}, \dots, t_1, t_0\}$, 称为“迹向量”, $\text{Tr}(\alpha) = \alpha \& \text{Trace_Vector}$. 只要知道 m 对应的 Trace_Vector , 不必每次都运行一遍求迹算法.

随着 m 的增大, 迹向量的计算时间暴涨. 实验表明, 在配置为 CPU AMD sempron 2800+、内存 256 MB 的台式机上, 求取 $m = 409$ 的迹向量需耗时 1 h, $m = 571$ 的迹向量需耗时 13.8 h, 迹向量计算方法需要进一步改进.

4 基于 HMM 预测的基点快速产生算法改进

HMM^[24] 最早于 1957 年被提出来, 是用于统计学习的并具有很强预测能力的概率模型. HMM 由双随机过程组成, 其中之一是 Markov 链; 另外一个随机过程描述状态和观察者之间的统计对应关系. 观察者不能直接看到 Markov 链中的状态, 但能够通过当前状态预测下一状态的可能结果.

图 3 HMM 模型

Figure 3 HMM

4.1 HMM 模型设计

将已知的既约多项式作为观察值, 而迹向量作为状态值. 由此, 我们设计出 HMM 的两个随机过程, 具体模型如图 3 所示.

4.2 HMM 参数定义

N : 模型中 Markov 链的状态数目, 这里取 $N = 3$, 同时记 t 时刻的状态为 q_t . 这 3 个状态值为:

θ_1 : 迹向量结果为 1;

θ_2 : 迹向量结果为既约多项式前一段;

θ_3 : 迹向量结果为既约多项式前一段再在高位补充几位.

M : 每个状态对应的可能出现的观察值数目, 这里取 $M = 2$, 同时记 t 时刻的观察值为 o_t . 这两个观察值为:

V_1 : 既约多项式的非零系数个数为 3;

V_2 : 既约多项式的非零系数个数为 5.

π : 初始状态矢量, 在本模型中指的是实验开始时, 选取的是某种迹向量 θ_1 (随机 m 值对应的迹向量) 概率.

$$\pi = (\pi_1, \pi_2, \pi_3), \quad \text{其中 } \pi_i = P(q_1 = \theta_i), \quad 1 \leq i \leq N.$$

A : 状态转移概率矩阵, $A = (a_{ij})_{3 \times 3}$, 其中

$$a_{ij} = P(q_{t+1} = \theta_j | q_t = \theta_i), \quad 1 \leq i, j \leq N.$$

在本模型中指的是如果当前选取的迹向量是 θ_i , 转移到下一个迹向量 θ_j 的概率.

B : 观察值概率矩阵, $B = (b_{jk})_{N \times M}$, 其中

$$b_{jk} = P(o_t = V_k | q_t = \theta_j), \quad 1 \leq j \leq N, \quad 1 \leq k \leq M.$$

本模型中 b_{jk} 指的是在出现 θ_j 迹向量情况下, 既约多项式系数个数为 V_k (3 或 5) 的概率. 藉此, HMM 模型可记为 $\lambda = (N, M, A, B, \pi)$, 可简写为 $\lambda = (A, B, \pi)$.

4.3 HMM 预测迹向量计算过程

(1) HMM 初始化.

从以上我们得知: 处于状态 θ_1 的数据有 8 个, 处于状态 θ_2 的数据有 4 个, 处于状态 θ_3 的数据有 13 个. 整个状态序列是一个随机过程, 因而我们可以得到大致的状态转移矩阵以及初始状态概率:

$$A = \begin{pmatrix} 0.32 & 0.16 & 0.52 \\ 0.32 & 0.16 & 0.52 \\ 0.32 & 0.16 & 0.52 \end{pmatrix}, \quad (3)$$

$$\pi = \{0.32 \quad 0.16 \quad 0.52\}. \quad (4)$$

另外, θ_1 对应多项式系数为 3 的数据有 6 个, 对应多项式系数为 5 的数据有 2 个; θ_2 对应多项式系数为 3 的数据有 4 个, 对应多项式系数为 5 的数据有 0 个; θ_3 对应多项式系数为 3 的数据有 0 个, 对应多项式系数为 5 的数据有 13 个. 由此我们可以得到观察值概率矩阵

$$B = \begin{pmatrix} 0.75 & 0.25 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (5)$$

(2) HMM 参数训练.

观察序列是一个随机过程链, 将 $\{13, 17, 19, \dots, 409, 571\}$ 中 25 个数据进行随机组合, 随机分成 5 组, 每组 5 个元素. 假定以下一种随机分组, 以及对应的观察序列:

$$(83, 109, 359, 277, 41) \implies (V_2 \ V_2 \ V_1 \ V_2 \ V_1),$$

$$(113, 103, 13, 331, 409) \implies (V_2 \ V_1 \ V_2 \ V_2 \ V_1),$$

$$(347, 101, 233, 19, 97) \implies (V_1 \ V_2 \ V_2 \ V_2 \ V_2),$$

$$(23, 131, 107, 571, 311) \implies (V_2 \ V_2 \ V_1 \ V_2 \ V_1),$$

$$(349, 239, 283, 163, 17) \implies (V_1 \ V_1 \ V_2 \ V_2 \ V_1).$$

采用 $(83, 109, 359, 277, 163)$ 对 HMM 进行训练. 得到第一次训练后的结果:

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0.28 & 0.38 & 0.33 \end{pmatrix}, \quad (6)$$

$$B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (7)$$

结果分析: 以上模型得到的迹向量只是最大概率的相似, 与精确计算的结果可能存在一定误差, 这种误差的最直接表现是对二次方程 $z^2 + z + c = 0$ 是否有解的判定产生错误, 即把不是循环群上的点判定为落在循环群上, 从而进行无效的 $kP = O$ 检验, 浪费了计算资源. 为了便于分析, 我们可以用两者之间的汉明距离来表示这种误差. 实验表明, 误差的汉明距离一般不超过 3, 也就是说, 随机选取的点有 $1/8$ 的概率能够正确判定.

PC 机: CPU 为 AMD sempron 2800+; 内存为 256 MB; 硬盘 80 GB.
软件平台: Microsoft Visual C++ 2005.

- $a=0, b=1$ 时有 9 个: 233, 571, 1913 等;
- $a=1, b=1$ 时有 10 个: 163, 283, 701 等;

6 安全性分析

1、椭圆曲线的阶能够被一个大素数 (大于 2^{160}) 整除. ECC 曲线的阶 $\#E(GF(p))$ 的包含大于 2^{160} 的大素数因子, 则平方根攻击 (Baby Step/Giant Step, Pollard- ρ [25]) 方法对于曲线攻击是无效的, 同时保证 Pohlig-hellman 攻击 [26] 的 $O(\sqrt{n})$ 计算量不能实现.

329

表 1 演化计算实验得到得 $m=1913$ 的安全域各个参数、基点以及运算时间

Table 1 The parameters, base points and computation time for the evolutionary computing experiments to find a secure base field with $m = 1913$

m	1913
a	0
$P(t)$	$t1913 + t462 + 1$
h	4
n	18549051793122879180755832753102533201307123170836936863828311783317415290370319116231599 10318771582347127907195077732409786694854454271860823301532849986915566161378680316362858 33935343042490907923065103318890199370041031605701878307179269364175607357713678657834158 33417081448956274817152939109510502080446621307340506344124445380758520589281270514743788 09227569635054287482828860058338975911755934640281128910856977249191069538500634179279407 84289995805895418122858005792857164394481730055626536047325501369261533212606286712416728 239385836218680905417981825652174545911343
Gx	0x171AE2E4006571ED0FC82A2EF714ABA236FE448D044B2BC749AF4B8C04EE882F8DFDEE4001 F34E9BCD4A1F38E1822D10A822F8BA360D5FE27FC6B4877E544512885C00705D3199A48E0C9F3F 558E90F09171A4A157AFEE940EE7BE528D36CC7C6E69266BEAF137096197C411FBFF86142A2AF4 A4E1172BCD167DD09FE3F3B0759CB8B9C3B3BEE8EE03ADC792D9F05FCFA75DF356843B6B296 2317FC94AF5E1ACFDDDD215234BD7D6FA5E4B87A740BA0791639B286BE0B8735438B13D732ACFA 8E41818A1C824D8170323937B93BDE6705C021C38B5ABF2BDACD6668BFAE0937526A550F3A99E3 80C724428CE0F825AFA
Gy	0x9C3C42689601F773755657361F05D7B4EDAC6F75DBEBB9CC5228E82D31CB1F673405745CB34F 169BDF04A1FD2F7A2CA71D7957FFDCC3F6CF07CDCD7D04E7CAB98AE66A76FFAA764EF6761 07A26040FF0389ECD699BBD4044E553DF38BAA7A0C28C8C711C6F2FBCB72DE91CD9145AB424 9C57A8EBA3309DA372519558719E5C0814E649D59A0DAC36A73A7928C85B60C8854C99AF793A8 B3D808A2EC6C34141110568E0D4AEDB1F11B82197297FA4D96FCAD54D69783DC9599D9ED88D E0FCAAFB45A2BE9E9284E1F20D0E354292934EDDE6EEA1585E82EF70C198E8A92ED7F5DA76 72DA9CECF835645A23CBCA25585B
Time	3.3 h

我们只选择是大素数的阶).

2、必须避免受到 MOV 攻击. 为抵御抗 MOV 攻击^[27], 须有 $p^i \neq 1 \bmod n$, 其中 $i = 1, 2, \dots, \log_2^{p/8}$, n 为基点的阶. 实验表明, MOV 攻击只对很少一部分的曲线有效, 本文还是将这个因素考虑在内. 由于 MOV 检测只涉及到基域 p 及基点的阶 n , 因此该部分我们安排在求阶计算之后, 而不是在基点选取之后. 具体流程如图 4 所示.

我们将搜索得到的安全基域 $p = 2^m (m = 163, 223, \dots, 571, 701, \dots, 1913)$ 以及相对应基点的阶 n 代入以上模型进行检验, 结果表明以上基域均不会遭受 MOV 攻击, 因此建立在上述参数的椭圆曲线是安全的.

3、若曲线的阶等于基域, 这样的曲线不能使用. 在 Koblitz 安全曲线选择实验中我们并未发现有基点阶等于基域的情况出现. 事实上, 即便这种情况出现了, 由于 Koblitz 曲线的基域都是以 2 为特征

图 4 MOV 攻击检验流程图

Figure 4 MOV detection model

的, 当基点的阶等于基域的时候, 这样的阶必定不是大素数. 因此在 Koblitz 安全曲线选择中可以不必担忧这类攻击.

4、避免将来可能存在的对特殊非奇异曲线攻击. 由于我们这里研究的是 Koblitz 类型曲线, 所以不存在特殊非奇异曲线攻击的问题. 另外, 我们考虑了 Weil 下降攻击 (GHS 攻击) [28]. 与 MOV 攻击一样, 该攻击能够成功的概率很小. Menezes 等 [18] 的研究表明, 当基域 2^m 中的参数 m 为素数且大于 160 的时候, GHS 攻击将会变得很困难. 我们所公布的安全曲线基域参数 $m = 163, 223, \dots, 1913$ 都是素数.

5、小余因子的问题. 另外, FIPS 186-3 也明确指出 ANSI X9.62 (1998) 在安全曲线选择时没有考虑余因子大小问题, 这样选择的曲线会有一定的安全隐患. FIPS 186-3 给出了不同基域范围下, 余因子所能够达到的最大值, 同时指出: 余因子应当越小越好. 本文 Koblitz 安全曲线选择中余因子只取 2 和 4 两种情况, 因此也满足这方面的安全要求.

综上所述, 本文选择的安全准则与美国 NIST 推荐安全曲线的安全准则基本一致, 因此本文所选出来的曲线完全涵盖了美国 NIST 公布的 5 条 Koblitz 安全曲线. 在此基础上, 本文还发现了其他安全曲线存在, 同时将安全基域范围扩大到更高的数量级.

7 结论和展望

考虑云计算攻击和未来量子计算机攻击, 本文基于演化密码结合 HMM 的思想, 解决了基点产生的困难问题, 提出一种新的安全基域选择算法. 并在 PC 机上初步完成了 $F(2^{2000})$ 以内 Koblitz 安全曲线 (基域和基点) 产生实验, 而且产生的安全曲线基域和基点的范围最大超过 1900 bit (PC 机耗时 5 个小时), 远远超过美国 NIST 公布的最高 571 bit. 本文在 NIST 公布的 $F(2^{163}) \sim F(2^{571})$ 范围之间还发现了新的安全曲线, 理论分析表明, 这些新发现的曲线也同样拥有较高的安全特性.

本文选择的这样的安全准则与美国 NIST 推荐安全曲线的安全准则一致, 因此本文所选出来的曲线可以抵抗常见的 Baby Step/Giant Step 法、Pollard- ρ 方法、Pohlig-Hellman 等攻击, 同时可以保证选择出曲线不是超奇异的, 因此可以抵抗 MOV 攻击.

本文理论分析和实验进一步证明了演化密码思想于对称密码和公钥密码都是安全有效的.

参考文献

- 1 Zhang H G, Feng X T, Tan Z P, et al. Evolutionary cryptography and the evolutionary design for DES. *J Commun*, 2002, 5: 57–64 [张焕国, 冯秀涛, 覃中平, 等. 演化密码与 DES 密码的演化设计. 通信学报, 2002, 5: 57–64]
- 2 Meng Q S, Zhang H G, Wang Z Y, et al. Designing bent functions using evolving method. *Acta Electron Sin*, 2004, 11: 1901–1903 [孟庆树, 张焕国, 王张宜, 等. Ben 函数的演化设计. 电子学报, 2004, 11: 1901–1903]
- 3 Wang Z Y, Li B, Zhang H G. Research on security of Hash functions. *Comput Eng Appl*, 2005, 12: 18–19 [王张宜, 李波, 张焕国. Hash 函数的安全性研究. 计算机工程与应用, 2005, 12: 18–19]
- 4 Wang Z Y, Li L, Zhang H G. Automatic design approach of security protocols. *Comput Eng Appl*, 2005, 12: 16–17 [王张宜, 李莉, 张焕国. 网络安全协议的自动化设计策略. 计算机工程与应用, 2005, 12: 16–17]
- 5 Chen L J, Zhao Y, Zhang H G. Cryptanalysis for stream cipher based on evolutionary computation. *Comput Appl*, 2008, 8: 1912–1915 [陈联俊, 赵云, 张焕国. 一种基于演化计算的序列密码分析方法. 计算机应用, 2008, 8: 1912–1915]
- 6 Zhang H G, Wang C, Shi X Y, et al. Fast Generating Algorithm for ECC Secure Curve Based on Evolutionary Computation. China Patent, ZL200910200504. 2010 [张焕国, 王潮, 时向勇, 等. 基于演化计算的安全椭圆曲线快速选择算法. 中国专利, ZL200910200504. 2010]
- 7 Wang C, Zhang H G, Liu L L. The experiment of Koblitz elliptic curves generating based on evolutionary cryptography theory and verifying the parameters recommended by NIST. *China Commun*, 2011, 8: 41–49
- 8 Dustin O. Securing elasticity in the cloud. *Commun ACM*, 2010, 53: 46–51
- 9 NIST. Digital Signature Standard. Federal Information Processing Standards Publication, 2000
- 10 Brown M, Hankerson D, Lopez J, et al. Software implementation of the NIST elliptic curves over prime fields. In: *Topics in Cryptology – CT-RSA 2001, Proceedings Lecture Notes in Computer Science*, 2001. 250–265
- 11 Guneyasu T, Paar C. Ultra High Performance ECC over NIST Primes on Commercial FPGAs. In: *Lecture Notes in Computer Science*. Berlin/Heidelberg: Springer, 2008. 62–78
- 12 Schoof R. Elliptic curves over finite fields and the computation of square roots mod. *Math Comput*, 1985, 44: 483–494
- 13 Elkies N D. Elliptic and modular curves over finite fields and related computational issues. *Comput Perspect Number Theory*, 1998, 7: 21–76
- 14 Satoh T. The canonical lift of an ordinary elliptic curve over a finite field and its point counting. *J Raman Math Soc*, 2000, 15: 483
- 15 Satoh T. On p-adic point counting algorithms for elliptic curves over finite fields. In: *Algorithmic Number Theory, 5th International Symposium, ANTS-V, Sydney, 2002*. LNCS 2369. 43–66
- 16 Fouquet M, Gaudry P, Harley R. An extension of Satoh's algorithm and its implementation. *Raman Math Soc*, 2000, 15: 281
- 17 Satoh T, Skjernaa B, Taguchi Y. Fast computation of canonical lifts of elliptic curves and its application to point counting. *Finite Field Their Appl*, 2003, 9: 98–101
- 18 Menezes A, Qu M. Analysis of the Weil descent attack of Gaudry, Hess and Smart. In: *Topics Cryptology—CT-RSA 2001*. Berlin/Heidelberg: Springer-Verlag, 2001. LNCS 2020. 308–318
- 19 Gaudry P. A comparison and a combination of SST and AGM algorithms for counting points of elliptic curves in characteristic 2. In: *Asiacrypt 2002*. LNCS 2501. 311
- 20 Dorigo M, Maniezzo V, Colorni A. Introduction to natural algorithms. *Rivista Inform*, 1994, 24: 179–197
- 21 Colorni A, Dorigo M, Maniezzo V. An investigation of some properties of an ant algorithm. In: *Proceedings of the Parallel Problem Solving from Nature Conference (PPSN 92)*. Elsevier Publishing, 1992. 509–520
- 22 Dorigo M. Optimization, learning and natural algorithms. Dissertation for the Doctoral Degree. Milan: DEI, Politecnico di Milano, 1992
- 23 Wu Q D, Wang L. Intelligent Ant Group Algorithm and It's Application. Shanghai: Shanghai Press of Science Technology and Education, 2004 [吴启迪, 汪镭. 智能蚁群算法及应用. 上海: 上海科技教育出版社, 2004]
- 24 Eddy S R. Hidden Markov models. *Curr Opin Struct Biol*, 1996, 6: 361–365
- 25 Oorschot P, Wiener M J. Parallel collision search with cryptanalytic applications. *J Cryptol*, 1999, 12: 1–28
- 26 Pohlig S C, Hellman M E. An improved algorithm for computing logarithms over GF(P) and the cryptographic significance. *IEEE Trans Inf Theory*, 1978, 24: 106–110

- 27 Menezes A T, Okamoto T, Vanstone S A. Reducing elliptic curve logarithms in a finite field. *IEEE Trans Inf Theory*, 1993, 39: 1639–1646
- 28 Diem C. The GHS-attacks in odd characteristic. *J Ramanujan Math Soc*, 2003, 18: 1–32

Evolutionary cryptography theory based generating method for a secure Koblitz elliptic curve and its improvement by a hidden Markov models

WANG Chao¹, ZHANG HuanGuo^{2,3} & LIU LiLi⁴

1 *Key Lab of Specialty Fiber Optics and Optical Access Network, Ministry of Education, Shanghai University, Shanghai 200072, China;*

2 *Computer School of Wuhan University, Wuhan 430072, China;*

3 *Key Laboratory of Aerospace Information Security and Trusted Computing Ministry of Education, Wuhan 430072, China;*

4 *Huawei Technologies CO. LTD., Shanghai 201206, China*

*E-mail: wangchao@staff.shu.edu.cn

Abstract Considering potential attacks from cloud-computing and quantum-computing, it is becoming necessary to provide higher security elliptic curves. The hidden Markov models are introduced for designing the trace-vector computation algorithm to accelerate the search for elliptic curve (EC) base-points. We present a new algorithm for secure Koblitz EC generation based on evolutionary cryptography theory. The algorithm is tested by selecting a secure Koblitz EC over the field $F(2^{2000})$, with experiments showing that both the base field and base point of the secure curve generated exceed the parameter range for Koblitz curves recommended by NIST. The base fields generated go beyond 1900 bits, which is higher than the 571 bits recommended by NIST. We also find new secure curves in the range $F(2^{163})$ – $F(2^{571})$ recommended by NIST. We perform a detailed security analysis of those secure curves, showing that those we propose satisfy the same security criteria as NIST.

Keywords secure EC selection, evolutionary cryptography theory, ant colony algorithm, Koblitz EC, HMM

WANG Chao was born in 1971. Currently, he is a Professor of Shanghai University. His research interests include WSN, information security, ECC, cryptography analysis and design by quantum computing. He is the secretary of IEEE Shanghai Section, the vice Chair of IEEE Shanghai CAS Chapter and the vice Chair and secretary of IEEE Shanghai Computer Chapter and a CCF Senior Member.

ZHANG HuanGuo was born in 1945. Currently, he is a Professor of Wuhan University. His research interests include information security, cryptography, and post quantum cryptography. He is the managing director of Chinese Association for Cryptologic Research (CACR) and a member of China Science Editorial Committee.

LIU LiLi was born in 1985. He received his master degree from Shanghai University in 2011. His research interests include ECC and evolutionary cryptosystem. Now he is a Software Engineer at Huawei Technologies Co., LTD.