УДК 004.056.55

С.С. ТІМОХІН, І.Д. ГОРБЕНКО, д-р техн. наук

КОДУВАННЯ ІНФОРМАЦІЇ ТОЧКАМИ НА ЕЛІПТИЧНІЙ КРИВІЙ

Інформація є однією з найцінніших ресурсів. Недарма кажуть, що той, хто володіє інформацією, володіє світом. Тисячі років тому люди намагалися приховати інформацію від інших і саме тоді вони почали використовувати перші шифри.

Сьогодні питання захисту постає ще гостріше. Недаремно наш час прозвали інформаційною ерою. Шифрування інформації є одним з основних способів приховування інформації. Усі шифри поділяються на симетричні і асиметричні, кожні з яких мають свої переваги та недоліки. Серед асиметричних криптосистем широкого використання здобули криптосистеми на еліптичних кривих. Вони використовуються для цифрового підпису електронних документів, шифрування тощо. Але для використання шифрування постає інше питання – як представити інформацію у вигляді точки на еліптичній кривій?

Мета статті – огляд існуючих алгоритмів представлення інформації у вигляді точок на еліптичній кривій, їх порівняння та оцінка за критеріями складності виконання та швидкодії, використання ресурсів.

Поточне шифрування на еліптичних кривих

Шифрування на еліптичних кривих є одним з найбільш стійких схем направленого шифрування. Алгоритм шифрування має наступний вигляд:

1. Генеруються загальносистемні параметри (ЗСП):

$$ÇÑÏ = (a,b,G,n,p,u,h) \quad (1)$$

- u – порядок еліптичної кривої,
- h – це коефіцієнт кореляції,
- n – порядок базової точки,
- p – велике просте число,
- G – базова точка,
- a, b – коефіцієнти еліптичної кривої.

При застосуванні еліптичної кривої над розширенням поля p виступає у якості основи поля Галуа, а операції наводяться за модулем примітивного многочлена f(x).

2. Обирається особистий ключ $1 < d < n-1$.

3. Обчислюється відкритий ключ за формулою

$$Q = d\,G \bmod p \quad (2)$$

Використовуючи ЗСП та відкритий ключ, можна зашифрувати повідомлення. Для цього обидва користувачі мають асиметричну пару, наприклад (d_a, Q_a) та (d_b, Q_b). Нехай А хоче передати повідомлення до В. Тоді А генерує сеансовий ключ:

$$R = k\,G, де 1 < k < n-1 \quad (3)$$

Після цього абонент А обчислює криптограму:

$$C = m + kQ_b \quad (4)$$

Абонент А передає абоненту В криптограму та сеансовий ключ (R, C). Для розшифрування абоненту В необхідно обчислити:

$$\tilde{N} - d_b R = m + kQ_b - d_b R = m + kd_b G - d_b kG = m \quad (5)$$

Тепер абонент В може прочитати повідомлення m.

Кодування інформації методом Кобліца

У 1985 році Н. Кобліц [1] запропонував імовірнісний алгоритм кодування. Цей алгоритм дозволяє представити інформацію з довільним алфавітом потужності $м$, $0 \le m < M$. Обирається достатньо велике число k, при чому $p > kM$, що характеризує вірогідність помилки, тобто не буде існувати точки, що може представити даний блок тексту:

$$P_{îøèá} = 1/2^k \tag{6}$$

Очевидно, що при збільшенні збільшується надійність, а розмір алфавіту зменшується і навпаки. Цілі числа представляються у вигляді, $0 \le mk + j < M$, де $1 \le j \le k$. Якщо хоча б для одного з $mk + j$ знайдено корінь квадратний за модулем, то вважається, що інформація успішно закодована. Для декодування необхідно взяти x координату точки та обчислити її значення за модулем k: якщо x координата ділиться на k, тоді інформація обчислюється за формулою (7), інакше (8):

$$i = x / k \tag{7}$$

$$i = (x - x \bmod k) / k \tag{8}$$

Перевагою методу є його простота, та у пам'яті необхідно тримати тільки точку з координатами та число k. Алгоритм не вимагає жодних перед обчислень та дотикових витрат. Достатньо простою є процедура вилучення інформації. Недоліками методу є те що він імовірнісний, але при достатньо великому виборі параметра k, вірогідність помилки прямує до нуля. Але це зменшує максимальний розмір алфавіту. При достатньо великому вірогідність не закодувати інформацію виростає багатократно [3]. Наприклад, при $M = 10^6, k = 20$ вірогідність не закодувати повідомлення складає $1-(1-1/2^k)^M = 0.6$, тобто вірогідність закодувати інформацію більше ніж закодувати. За допомогою даного алгоритму можна закодувати алфавіт, що приблизно в *k* разів менше, ніж кількість точок на еліптичній кривій.

Кодування інформації з використанням дуальних еліптичних кривих

Інший алгоритм використовує поняття дуальних ЕК [2]. Дві ЕК (9) називаються дуальними, якщо виконується (10), де не є квадратом у полі:

$$E_{a,b} : y^2 = x^3 + ax + b, E_{a`,b`} : y^2 = x^3 + a`x + b` \tag{9}$$

$$\begin{cases} a` = v^2 a \bmod p \\ b` = v^3 b \bmod p \end{cases} \tag{10}$$

Існує теорема про порядок дуальних ЕК:

$$\# E_{a,b} + \# E_{a`,b`} = 2p + 2 \tag{11}$$

Обирається довільний $v \in Fp$ і обчислюються параметри дуальної ЕК. Необхідно знайти, для яких значень $i \in [0, 2M-2]$ значення виразу $i^3 + ai + b$ не є квадратом, а для яких є, а також для iv виразу $(iv)^3 + a`iv + b`$. Далі обирається ЕК, що має більше квадратів на цьому відрізку. Інформація кодується послідовно, кожний символ алфавіту у наступний квадрат поля, для якого обчислюється у координата над еліптичною кривою.

Перевагами даного метода є те що він детермінований і інформація буде закодована у будь-якому разі. Причому для кодування можуть використовуватися усі точки еліптичної кривої і потужність алфавіту може дорівнювати порядку еліптичної кривої. Але проведення обчислень у даному випадку є експоненційно складною задачею.

Недоліками цього метода є велика кількість перед обчислень, що необхідно зробити для побудови таблиці кодування. При великому розмірі алфавіту це може зайняти велику кіль-

кість часу. І найважливішим недоліком є значне послаблення стійкості криптосистеми, якщо для кодування використовується дуальна еліптична крива[3].

Наприклад, розглянемо еліптичну криву P-192, рекомендовану [4]. Порядок даної кривої:

$$\#E_{ab} = 6277101735386680763835789423176059013767194773182842284081 \quad (12)$$

Тоді за теоремою про порядок дуальних еліптичних кривих маємо:

$$\begin{aligned}\#E_{a'b'} &= 6277101735386680763835789423239273818400622627597807638479 = \\ &= 23 \cdot 1086437506056025160590067774 3 \cdot 25120401793443689936479125511\end{aligned} \quad (13)$$

Як бачимо, що найбільшим множником є число 25120417934436899364791255 11. Таким чином, при використанні метода Поларда для розв'язання дискретного логарифму необхідно виконати $\sqrt{25120401793443689936479125511} = 158494169588170$ операцій над точками еліптичної кривої.

Таким чином, для кодування інформації методом на дуальних еліптичних кривих необхідно виконати обчислення, що складається з O(M) операцій визначення квадратного кореня у полі. Алгоритм вимагає передчасної побудови таблиці кодування та декодування та на прийомній стороні користувачеві доведеться виконувати аналогічну процедуру.

Для виконання операції кодування та декодування у пам`яті необхідно тримати усю таблицю кодування алфавіту.

Кодування інформації у системі лишкових класів

Був запропонований інший детермінований алгоритм кодування інформації на еліптичній кривій. Для цього обирається особлива еліптична крива, побудована у кільці *Zq*, де

$$q = \prod_{j=1}^{n} p_j \quad (14)$$

p_j – попарно прості числа. Алгоритм використовує представлення числа у вигляді системи лишкових класів, використовуючи китайську теорему про лишки.Розглянемо алгоритм кодування. Вхідними даними кодування є ЕК побудована у кільці *Zq*. Необхідно побудувати n масивів обчислюючи точки на даній еліптичній кривій для кожного з множників *q*. Тобто $\forall i, 0 < i < n, \forall x, 0 < x < p_i$ обчислюється символ Лежандра:

$$\left(\frac{a}{p}\right) = \left(\frac{x^3 + ax + b}{p_i}\right). \quad (15)$$

Якщо вираз (15) дорівнює одиниці, тоді у *i*-й масив записуються дві точки з координатами:

$$Ð\left(x, sqrtmod\left(x^3 + ax + b\right)\right), Ð(x,\ p_i - sqrtmod\ (x^3 + ax + b)) \quad (16)$$

Функція *sqrtmod*(*a*) обчислює корінь квадратний за модулем. Якщо ж вираз (15) дорівнює нулю, тоді у цей масив записується одна точка з координатами:

$$P(x, 0) \quad (17)$$

У випадку, якщо вираз (15) дорівнює -1, тоді у масив не записується нічого. Тоді максимальна кількість точок, що може бути закодована алфавітом на (18):

$$Í\ ÑÊ\ (l_i, \forall i, 0 < i < n) \quad (18)$$

При використанні даного алгоритму інформація представляється масивом точок довжини *n*, тобто однією точкою за кожного масиву. Для кожного символу *m* алфавіту *M* обираєть-

ся n точок з кожного з сформованих масивів з індексом $m\ modp_i$. Таким чином, інформація представляється масивом точок за кожним з множників розкладу розміру кільця.

Для того, щоб декодувати отриману інформацію необхідно сформувати тривимірний масив $k[n][p_i][p_i]$. Масив заповнюється наступним чином: $\forall i, 0 < i < n, \forall x, 0 < x < p_i$, якщо символ Лежандра (15) дорівнює одиниці, тоді:

$$k_{i,x,sqrtmod(x^3+ax+b)} = t, k_{i,x,p_i-sqrtmod(x^3+ax+b)} = t+1, t = t+2 \tag{19}$$

Якщо він дорівнює нулю, тоді маємо:

$$k_{i,x,0} = t, t = t+1 \tag{20}$$

Отримавши n точок, необхідно обрати з масиву n чисел підставивши по порядку $0<i<n$ у першу координату масиву та x-, y-координати у відповідно другу та третю координати масиву. Отримані числа є представленням закодованого числа у системі лишкових класів.

Для того щоб відновити число з системи лишкових класів, необхідно скористатися формулою

$$X = (\sum_{i=0}^{n-1} m_i (x_i m_i^{-1} mod p_i)) mod q \tag{21}$$

$$m_i = \frac{q}{p_i} \tag{22}$$

x_i – числа, вилучені з масиву.

Перевагами цього методу є те, що він є детермінованим і допомагає закодувати алфавіт, що складається з великої кількості слів. Але він використовує особливі еліптичні криві, що, як правило, не використовуються в криптографії, так як розклад числа на множники послаблює стійкість криптосистеми. Також можуть виникнути помилки у випадку, якщо $l_i < p_i$. Тобто індекс у масиві під час кодування може вийти за межі масиву і не буде точки, якою б можна було представити число.

Висновки

Отже, з проаналізованих алгоритмів краще використовувати імовірнісний алгоритм Кобліца, адже він допомагає достатньо швидко закодувати алфавіт, що складається з великої кількості слів. На жаль, використовуються не всі точки еліптичної кривої і існує вірогідність помилки, але при оптимальному виборі параметрів алгоритму можна створити швидку, ефективну і надійну систему.

Список літератури: 1. *Коблиц, Н.* Курс теории чисел и криптографии. – Москва : ТВП, 2001. – 254с. 2. *Лёвин, В.Ю.* Кодирование алфавитов точками эллиптических кривых : дипломна робота, 2007. 3. *Бабенко, М. Г.* Разработка программного комплекса шифрования данных, на основе использования точек эллиптической кривой : наук.-дослід. робота, 2011.

Харківський національный університет ім. В.Н. Каразіна

Надійшла до редколегії 15.02.2014