МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ

УЧРЕЖДЕНИЕ ОБРАЗОВАНИЯ
«БРЕСТСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Кафедра «Интеллектуальные информационные технологии»

ПРИМЕНЕНИЕ ЭЛЛИПТИЧЕСКИХ КРИВЫХ В КРИПТОГРАФИИ

Методические указания к выполнению лабораторных работ
по дисциплине
«Криптографические методы защиты информации»

для студентов специальности
1-40 03 01 «Искусственный интеллект»

Брест 2012

УДК 347 77/681.3

В методических указаниях приведены необходимые теоретические сведения по эллиптическим кривым, содержится информация о протоколах обмена ключевой информацией с использованием эллиптических кривых, также алгоритмы и указания для выполнения операции экспоненцирования над точкой эллиптической кривой.

Методические указания предназначены для использования студентами специальности 1-40 03 01 «Искусственный интеллект» в ходе выполнения лабораторных работ по дисциплине «Криптографические методы защиты информации».

Составитель: Хацкевич М.В., старший преподаватель

Рецензент: Пролиско Е.Е. доцент кафедры моделирования Учреждения образования

Учреждение образования

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Алгоритмы традиционной криптографии строятся путем комбинирования большого числа относительно несложных преобразований способом, обеспечивающим хорошие характеристики итогового алгоритма. Практические основы были заложены в работах Хорста Фейстеля, и с тех пор принципы построения одноключевых шифров изменились не очень сильно. Помимо некоторого расширения набора базовых преобразований и большего разнообразия в архитектурах изменения носят в основном количественный характер и отражают развитие используемой вычислительной базы. Важно, что типовой размер ключей и блоков данных, применение которых считается безопасным, вырос примерно в 2 раза. Иная картина наблюдается в современной криптографии. Стойкость ее алгоритмов базируется на недоказанной пока вычислительной невозможности эффективного решения некоторых математических задач, то есть на гипотезе, которая может оказаться ошибочной. Например, стойкость криптосистемы RSA базируется на сложности задачи факторизации больших чисел, а стойкость современных схем ЭЦП, большинство из которых являются вариациями обобщенной схемы Эль-Гамаля, - на сложности задачи логарифмирования в конечных полях. В настоящее время в современной криптографии существуют следующие проблемы:

- Ограниченность числа рабочих схем. В отличие от алгоритмов классической криптографии, которые могут быть созданы в неограниченном количестве путем комбинирования различных элементарных преобразований, каждая «современная» схема базируется на определенной «нерешаемой» задаче. Как следствие, количество рабочих схем криптографии с открытым ключом весьма невелико.
- Постоянное увеличение размера блоков данных и ключей обусловлено прогрессом математики и вычислительной техники. Например, «безопасный» размер чисел в RSA вырос практически на порядок; похожая картина наблюдается и для других схем, тогда как в традиционной криптографии этот размер увеличился всего вдвое.
- Потенциальная ненадежность базиса. В настоящее время теорией вычислительной сложности исследуется вопрос о возможности решения задач данного типа за полиномиальное время (гипотеза P = NP). В рамках теории уже доказана связь большинства используемых вычислительно сложных задач с другими аналогичными задачами. Это означает, что, если будет «взломана» хотя бы одна современная криптосистема, многие другие также не устоят.
- Отсутствие дальней перспективы. Уже известны квантовые вычисления, с помощью которых оказалось возможным решать многие задачи гораздо быстрее, чем на традиционных компьютерах. Правда, в настоящее время они существуют лишь в теории, из достижений можно отметить только успешную факторизацию числа 15. Специалисты полагают, что «серьезные» квантовые компьютеры появятся примерно через несколько десятков лет, поэтому будущее современной криптографии туманно. Для современной криптографии актуальна проблема повышения стойкости и уменьшения размера блоков данных путем модификации уже существующих криптосистем.

В большинстве современных продуктов и стандартов криптографии применяются методы с открытым ключом, основанные на проблеме факторизации больших чисел (RSA) и дискретного логарифмирования (Эль- Гамаль). Однако для их надежной защищенности число битов ключа в последние годы резко возросло, что обусловило рост нагрузки

на вычислительные системы. Подход на основе эллиптических кривых Е (которые являются источником конечных абелевых групп) обеспечивает эквивалентную защиту, в сравнении с ранее разработанными протоколами, при меньшем числе разрядов.

Криптосистемы на основе эллиптических кривых предложены в 1986 году В. Миллером и Н. Коблицем. Сложность атаки на ключ криптосистемы на основе эллиптической кривой экспоненциально связана с длиной ключа. Стойкость же алгоритмов, основанных на проблеме факторизации больших чисел (RSA) и задачи дискретного логарифмирования (Эль-Гамаля), субэкспоненциальная. При одинаковой стойкости криптосистемы на основе эллиптических кривых имеют размер модуля на порядок меньше. На основе эллиптических кривых был утвержден ряд стандартов: например, FIPS 186-2-2000, ГОСТ Р34.10-2001.

Эллиптические кривые над конечными полями доставляют неисчерпаемый источник конечных абелевых групп, которые (даже если они велики) удобны для вычислений и обладают богатой структурой. Во многих отношениях эллиптические кривые – естественный аналог мультипликативных групп полей, но более удобный, так как существует большая свобода в выборе эллиптической кривой, чем в выборе конечного поля.

1 МАТЕМАТИЧЕСКИЕ ОСНОВЫ ЭЛЛИПТИЧЕСКИХ КРИВЫХ

1.1 ОПРЕДЕЛЕНИЕ ЭЛЛИПТИЧЕСКОЙ КРИВОЙ

Определение. Эллиптической кривой Е над полем К называется множество точек $(x, y) \subset K \times K$, удовлетворяющих уравнению

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, a_i \in K \tag{1.1}$$

вместе с точкой О – точка на бесконечности.

Вместо (1.1) используется и функция двух переменных

$$F(x,y) = y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6 - 0 \tag{1.2}$$

Введение операции сложения над парами точек Е позволяет построить абелевую группу точек, если все точки Е неособые (имеют однозначные производные). Такую кривую называют гладкой, или несингулярной.

Определение. Кривая Е называется сингулярной (особой), если существует хотя бы одна точка *(x,y)*, в которой частные производные (1.2) одновременно обращаются в 0, т. е.

$$\partial F/\partial x = \partial F/\partial y = 0 \tag{1.3}$$

В противном случае кривая Е называется несингулярной (неособой). Такие кривые представляют интерес для криптографии.

Вместо общей записи уравнения (1.1) часто рассматривают уравнения трех типов кривых:

$$E: y^2 = x^3 + ax + b, p \neq 2,3 \tag{1.4}$$

$$Es: y^2 + y = x^3 + ax + b, p = 2 \tag{1.5}$$

$$E_N: y^2 + xy = x^3 + ax^2 + b, p = 2, b \neq 0 \tag{1.6}$$

(1.4) описывает все кривые над полями характеристики, не равной 2 и 3, (1.5) и (1.6) - кривые над полями характеристики 2.

Несингулярная кривая (1.4), таким образом, не имеет кратных корней кубического трехчлена $f(x)$ (т. е. кубическое уравнение имеет три различных вещественных корня, либо один вещественный и два комплексно-сопряженных корня).

Для несингулярной кривой (1.4) должно выполняться условие:

$$\Delta = (4a^3 + 27b^2) \neq 0 \qquad (1.7)$$

Величину Δ называют дискриминантом кубического трехчлена.

2 ЗАКОНЫ СЛОЖЕНИЯ ТОЧЕК ЭЛЛИПТИЧЕСКОЙ КРИВОЙ И ПОСТРОЕНИЕ АБЕЛЕВОЙ ГРУППЫ ТОЧЕК

Перед обсуждением конкретных примеров эллиптических кривых над различными полями отметим чрезвычайно важное свойство точек эллиптической кривой: они образуют абелеву группу относительно операции сложения точек, о которой будет подробнее сказано ниже. Чтобы объяснить наглядно, как это получается, временно будем полагать, что K = R, т.е. что эллиптическая кривая – обычная плоская кривая (с добавлением еще одной точки O «в бесконечности»).

Определение.

Пусть E – эллиптическая кривая над вещественными числами, и пусть P и Q – две точки на E. Определим точки -P и P+Q по следующим правилам:

1.Точка O – тождественный элемент по сложению («нулевой элемент») группы точек. В следующих пунктах предполагается, что ни P, ни Q не являются точками в бесконечности.

2. Точки P = (x, y) и -P имеют одинаковые x- координаты, а их y- координаты различаются только знаком, т.е. – (x, y) = (x,–y). Из (1) сразу следует, что (x,–y) – также точка на E.

$$-P = (x_1, -y_1) \qquad (2.1)$$

3. Если P и Q имеют различные x-координаты, то прямая L = PQ имеет с E еще в точности одну точку пересечения R (за исключением двух случаев: когда она оказывается касательной в P, и тогда полагаем R = P, или касательной в Q, и тогда полагаем R =Q). Определяем теперь P + Q как точку -R, т.е. как отражение от оси x третьей точки пересечения. Геометрическое построение, дающее P + Q, приводится ниже в примере 1.

4. Если Q = -P (т. е. x- координата Q та же, что и у P, а y- координата отличается лишь знаком), то полагаем P + Q = O («точке в бесконечности»; это является следствием правила 1).

5. Остается возможность P = Q. Тогда считаем, что L – касательная к кривой в точке P. Пусть R – единственная другая точка пересечения L с E. Полагаем P + Q = -R (в качестве R берем P, если касательная прямая в P имеет «двойное касание», т. е. если P есть точка перегиба кривой).

Пример 1. На рисунке 2.1 слева изображена эллиптическая кривая $y^2 = x^3 - x$ в плоскости xOy и приведен типичный случай сложения точек P и Q. Чтобы найти P + Q, проводим прямую PQ и в качестве P + Q берем точку, симметричную относительно оси x третьей точке, определяемой пересечением прямой PQ и кривой. Если бы P совпадала с Q, т.е. если бы нам нужно было найти 2P,то использовали бы касательную к кривой в P : тогда точка 2P симметрична третьей точке, в которой эта касательная пересекает кривую.

На рисунке 2.1 справа аналогичным образом проиллюстрировано сложение точек R и Q на кривой $y^2 = x^3 + x + 1$.

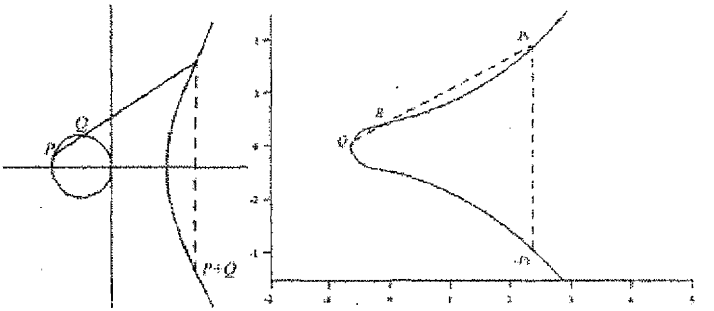
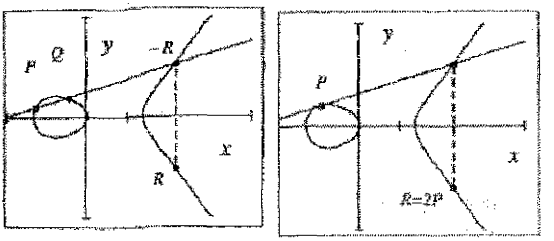


Рисунок 2.1 – Примеры геометрического построения суммы точек эллиптической кривой

Обозначим $(x_1, y_1), (x_2, y_2)$ и (x_3, y_3) – координаты точек P, Q и P+Q соответственно. Необходимо выразить (x_3, y_3) через $(x_1, y_1), (x_2, y_2)$.

Имеют место два случая:



а *б*

Рисунок 2.2 – Примеры геометрического построения суммы точек эллиптической кривой

1. $P \neq \pm Q$. Уравнение прямой линии, проходящей через точки P и Q (рисунок 2.1, а), имеет вид

$$y = \lambda x + \beta; \lambda = \frac{y_2 - y_1}{x_2 - x_1}; \beta = y_1 - \lambda x_1; \qquad (2.2)$$

Уравнение (1.2) в канонической форме (1.4) можно переписать

$$F(x,y) = y^2 - x^3 - ax - b = 0. \qquad (2.3)$$

Точки пересечения кривой E и прямой (2.2) имеют по оси x координаты x_1, x_2, x_3 точек P, Q и -R соответственно. Поскольку они являются общими для функций (2.2) и (2.3), последнее уравнение можно записать в виде

$(\lambda x = \beta)^2 - x^3 - ax - b = 0$, или $-(x - x_1)(x - x_2)(x - x_3) = 0$.

Приравнивая в этих кубических уравнениях коэффициенты при переменных x_2, получим:

$$\lambda^2 = x_1 + x_2 + x_3. \qquad (2.4)$$

Параметр λ прямой (2.2) можно также выразить в виде

$$\lambda = \frac{-y_3 - y_1}{x_3 - x_1}.$$

Из (2.4) и последнего соотношения окончательно имеем координаты точки R=P+Q= (x_3, y_3);

$$\begin{cases} x_3 = \lambda^2 - x_1 - x_2, \\ y_3 = -y_1 - \lambda(x_3 - x_1), \lambda = \dfrac{y_2 - y_1}{x_2 - x_1} \end{cases} \tag{2.5}$$

2. P = Q, R = 2P. В этом случае $x_1 = x_2$ и параметр λ не определен. Дифференциал функции (1.4)

$2ydy = 3x^2dx + adx;$

тогда при $x = x_1$ производная равна параметру ν касательной $y = \nu x + \beta$ к кривой в точке P

$$\nu = \left.\frac{dy}{dx}\right|_{x=x_1} = \frac{3x_1^2 + a}{2y_1}.$$

Теперь можно записать координаты точки R = 2P = (x_3, y_3):

$$\begin{cases} x_3 = \nu - 2x_1, P = Q; \\ y_3 = -y_1 - \nu(x_3 - x_1), \nu = \dfrac{3x_1^2 + a}{2y_1}. \end{cases} \tag{2.6}$$

Если n – целое число, то, как и в любой абелевой группе, nP обозначает сумму n точек P при n > 0 и сумму |n| точек -P, если n ≤ 0 .

Формулы сложения (2.5) и удвоения (2.6) справедливы для кривых E над всеми полями, в том числе и конечными, кроме полей характеристик 2 и 3. В последнем случае, как видно из (2.6), редукция по модулю 2 или 3 ведет к некорректности формул удвоения и следует использовать другие канонические уравнения кривых. Заметим, что координаты сложения и удвоения точек определяются с помощью всех операций в поле, т. е. сложения (вычитания), умножения и деления.

Для построения абелевой группы точек E определим O группы как

$P + (-P) = O; \forall P \in E;$

Если провести прямую через точки P и - P, то третья точка пересечения прямой и E уходит в бесконечную точку вдоль оси y . Поэтому O группы точек E называют «точкой на бесконечности».

Смысл перехода к обратной к точке пересечения прямой и кривой E при определении суммы R = P + Q становится понятным, если выразить, например, точку P как P = R - Q. В этом случае прямая проходит через точки R, - Q и -P, а обратной к этой третьей точке является точка P. Для точек E выполняется ассоциативность сложения

P+(Q+S) = (P+Q)+S , и коммутативность P+Q=Q+P.

Таким образом, множество точек E замкнуто относительно операции сложения, удовлетворяет свойствам ассоциативности, коммутативности, имеет обратный элемент и O (нуль группы), т. е. удовлетворяет всем условиям аддитивной абелевой группы.

Пример 2. Пусть P = (-3, 9) и Q = (-2, 8) – точки на эллиптической кривой $y_2 = x_3 - 36x$. Найти P + Q и 2P.

Решение. Подстановка $x_1 = -3, y_1 = 9, x_2 = -2, y_2 = 8$ в первое из уравнений (2.5) дает $x_3 = 6$; тогда второе из уравнений (2.5) дает $y_3 = 0$. Непосредственной подстановкой координат точки P + Q = (6, 0) в уравнение кривой можно убедиться в том, что она также лежит на ней.

Далее, подставляя $x_1 = -3, y_1 = 9, a = -36$ в первое из уравнений (2.6), получаем для x- координаты точки 2P значение 25/4, а второе из уравнений (2.6) дает для y- координаты значение -35/8. Точка 2P = (25/4, -35/8) также принадлежит рассматриваемой кривой.

3 ПОРЯДОК ГРУППЫ ЭЛЛИПТИЧЕСКОЙ КРИВОЙ И ПОРЯДОК ТОЧКИ

Аналогией с экспоненцированием в мультипликативной группе в аддитивной группе точек является k-кратное сложение элементов (в нашем случае - точек P), которое обозначается как:

$$P + P + P + ... + P = kP.$$

Точку kP называют скалярным произведением, а процесс ее вычисления – экспоненцированием точки P (возведением в степень).

Определение. Порядком N_E эллиптической кривой называется число всех ее точек (x, y) вместе с точкой на бесконечности (точкой O).

Определение. Порядком точки P эллиптической кривой называется наименьшее натуральное число $m \neq 0$, для которого mP= O.

Порядки N_E и m могут быть бесконечными и конечными. В группе бесконечного порядка (например, в группе точек E над полями R или Q) могут быть точки конечного порядка. В частности, в группе точек E над R всегда есть точки 2-го и 3-го порядков.

Координаты x_i точек 2-го порядка - это корни кубического трехчлена правой части (1.4). Для кубического трехчлена с тремя действительными корнями a_1, a_2, a_3 (рисунок 2.2, а) имеем три точки второго порядка $(a_1,0),(a_2,0),(a_3,0)$. В этих точках $(a_i,0) = -(a_i,0) \Rightarrow 2(a_i,0) = O$.

Вместе с точкой O эти точки образуют подгруппу 4-го порядка группы точек E бесконечного порядка. Кубический трехчлен с одним действительным корнем (рисунок 2.2, б) порождает подгруппу 2-го порядка точек 2-го порядка.

Точки 3-го порядка в группе точек E над R можно найти из соотношений

2P = -P => 3P=O => $x_3 = x_1$. Согласно (2.6)

$$\left(\frac{3x_1^2 + a}{2y_1}\right)^2 - 2x_1 = x_1 \Rightarrow x_1 \geq 0.$$

Отсюда с учетом (1.4)

$$3x_1^4 + 6ax_1^2 + 12bx_1 - a^2 = 0, x_1 \geq 0. \qquad (3.1)$$

В частности, при a = 0

$x_1(x_1^3 + 4b) = 0 \Rightarrow x_1 = 0.$

В общем случае x-координата точки третьего порядка совпадает с x-координатой точки перегиба эллиптической кривой E.

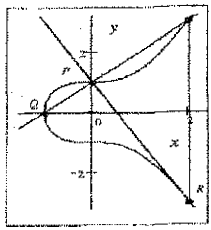


Рисунок 3.1 – График кривой $y^2 = x^3 + 1$

Точки конечного порядка кривой образуют так называемые подгруппы кручения. На кривой $y^2 = x^3 + 1$ на рисунке 3.1 имеются точки 2, 3 и 6-го порядков, образующие циклические подгруппы кручения тех же порядков.

Точка Q - точка 2-го порядка, являющаяся генератором подгруппы кручения $G_2 = \{O, Q\}$ 2-го порядка. Точка P - точка 3-го порядка, генерирующая подгруппу $G_3 = \{O, P, 2P\}$ того же порядка. Точка R - точка 6-го порядка, образующая подгруппу $G_6 = \{O, R, 2R = 2P = -P, 3R = Q, 4R = P, 5R = -R\}$ порядка 6.

Результат 2R = 2P следует из удвоения суммы P + Q= R. Очевидно так же, что группа G_6 может быть выражена через прямую сумму циклических подгрупп 2-го и 3-го порядков

$$G_6 = G_2 \otimes G_3 = \{0,Q\} \otimes \{0,P,2P\}.$$

Таким образом, любая точка группы G_6 выражается через сумму точек из подгрупп G_2 и G_3.

Наряду с циклическими подгруппами в группах точек E встречаются нециклические подгруппы кручения.

Например, уравнение $y^2 = (x-a_1)(x-a_2)(x-a_3)$ над полем R с тремя вещественными корнями порождает 3 точки 2-го порядка и соответствующие нециклические подгруппы кручения (нециклическая подгруппа 4-го порядка, из 3 циклических подгрупп (точек 2-го порядка) кручения; нет генерирующей точки для этой подгруппы).

Пример 3. Найти порядок точки P = (2, 3) на $y^2 = x^3 + 1$.

Решение. Применяя (2.6), находим, что 2P = (0, 1), и вновь, с помощью (2.6), что 4P =2(2P) = (0, -1). Поэтому 4P = -2P и, следовательно, 6P = O. Тем самым порядок P может быть равен 2, 3 или 6. Но 2P = (0,1) ≠ O, а если бы P имела порядок 3, то было бы 4P = P, что неверно. Итак, P имеет порядок 6.

4 ПРОЕКТИВНЫЕ КООРДИНАТЫ

При рассмотрении E полезным оказывается переход от аффинных координат (x,y) к проективным (X, Y, Z), связывающий точки кривой E в этих координатах отношением эквивалентности. Привлечение новой переменной Z позволяет задать координаты нуля группы E (точки на бесконечности). В операциях над конечными полями проективные координаты позволяют избежать трудоёмких вычислений обратного элемента поля при сложении точек.

Определение. Проективной плоскостью над полем K называется множество классов эквивалентности троек (X,Y,Z), в которых хотя бы один элемент ненулевой. Эквивалентными считаются тройки, если элементы одной из них получаются из другой умножением на скаляр:(X',Y',Z') ~ (X,Y,Z), если для некоторого элемента $\lambda \in K : (\lambda X', \lambda Y', \lambda Z') = (X, Y, Z)$.

Такие классы эквивалентности называются проективными точками. Например, две точки (7, 1, 1) и (8, 3, 3) эквивалентны в проективной плоскости над F_{13} (λ =3). Проективные точки с ненулевым элементом Z принадлежат классу эквивалентности, содержащему единственную точку, вида $(x, y, 1)$: она просто вычисляется x=X/Z, y=Y/Z.

В аффинных координатах уравнение кривой (1.4) запишем как

$$F(x,y) = \left(\frac{Y}{Z}\right)^2 - \left(\frac{X}{Z}\right)^3 - a\left(\frac{X}{Z}\right) - b = 0.$$

Умножим данное уравнение на Z_3

$$F(X,Y,Z) = Z^3 F(x,y) = Y^2 Z - X^3 - aXZ^2 - bZ^3 = 0. \qquad (4.1)$$

Исключая из рассмотрения начало координат (0, 0, 0), для любой тройки (X,Y, Z) класс эквивалентности задается проективной точкой (λ X, λ Y, λ Z), где λ - скаляр, X, Y, Z - фиксированы. В трехмерном пространстве этот класс представляет собой прямую линию, проходящую через начало координат. При $Z \neq 0$ любая такая прямая пересекает плоскость Z = 1, в которой, как видно из (4.1), возвращаемся к записи кривой в аффинных координатах. Таким образом, проективная плоскость может быть определена, как множество всех точек (x, y) обычной (аффинной) плоскости с дополнением точек, для которых Z=0.

Кроме точек с $Z \neq 0$ уравнению (4.1) удовлетворяет еще одна точка.

Подставим Z=0 в уравнение, получим $X_3 = 0$, это означает, что $X = 0$. Но имеется только один класс эквивалентности, где оба элемента X и Z нулевые – класс, содержащий (0,1,0). В проективной плоскости она задает координаты бесконечно удаленной точки или нулевого элемента группы точек E. Точка O является третьей точкой пересечения точек P и -P на бесконечности.

5 ДИСКРИМИНАНТ И J-ИНВАРИАНТ ЭЛЛИПТИЧЕСКОЙ КРИВОЙ

При изучении многих свойств и преобразований эллиптических кривых часто полезными оказываются дискриминант Δ и j-инвариант кривой. В частности, условие $\Delta \neq 0$ является необходимым и достаточным условием несингулярности кривой над любым полем.

Для кривой (1.4) дискриминант кубического уравнения

$$x^3 + ax + b = (x - e_1)(x - e_2)(x - e_3) = 0$$

определен формулой (1.7).

Найдем корни кубического трехчлена. Будем искать решение в форме

$$x = \sqrt[3]{r} + \sqrt[3]{s},$$

при этом

$$x^3 = r + s + 3\sqrt[3]{rs}(\sqrt[3]{r} + \sqrt[3]{s}) = r + s + 3x\sqrt[3]{rs}.$$

Сравнивая это уравнение с исходным, получаем $3\sqrt[3]{rs} = -a, r+s = -b$ и, следовательно,

$$rs = -\frac{a^3}{27}; r+s = -b \Rightarrow r^2 + br - \frac{a^3}{27} = 0.$$

Решение этого квадратного уравнения дает пару значений

$$r,s = -\frac{b}{2} \pm \frac{1}{6}\sqrt{\frac{-\Delta}{3}}; \Delta = -(4a^3 + 27b^2).$$

Дискриминант Δ кубического трехчлена совпадает с дискриминантом квадратного уравнения, определяющего значения r, s. Корни кубического уравнения теперь можно найти из

$$e = x = \sqrt[3]{r} + \sqrt[3]{s} = \sqrt[3]{r} - \frac{a}{3\sqrt[3]{r}}. \tag{5.1}$$

В общем случае дискриминант полинома *n*-й степени

$$f(x) = x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 = (x-e_1)(x-e_2)(x-e_3)...(x-e_n)$$

определяется как произведение квадратов разностей всех корней

$$\Delta = \prod_{i>k}(e_i - e_k)^2. \tag{5.2}$$

Для кривой (1.1) общего вида определим вспомогательные коэффициенты c_{2i}, d_4, дискриминант Δ и j-инвариант равенствами:

$$c_2 = a_1^2 + 4a_2;$$
$$c_4 = 2a_4 + a_1 a_3;$$
$$c_6 = a_3^2 + 4a_6;$$
$$c_8 = a_1^2 a_6 + 4a_2 a_6 - a_1 a_3 a_4 + a_2 a_3^2 - a_4^2;$$
$$d_4 = c_2^2 - 24c_4;$$
$$\Delta = -c_2^2 c_8 - 8c_4^3 - 27c_6^2 + 9c_2 c_4 c_6;$$
$$j = \frac{d_4^3}{\Delta}, \Delta \neq 0.$$

В частности, для кривой (1.4) получим

$$\Delta = -16 \bullet (4a^3 + 27b^2); j(E) = \frac{-4^3 12^3 a^3}{\Delta} = \frac{12^3 4a^3}{4a^3 + 27b^2}, \Delta \neq 0. \tag{5.3}$$

Отсюда видно, что кривые с коэффициентом a = 0 - кривые с нулевым j-инвариантом, а кривые с коэффициентом b = 0 - кривые с j-инвариантом, равным $12^3 = 1728$.

Для кривых (1.5), (1.6) характеристики 2 соответственно имеем

$$E_S : \Delta = 1, j = 0; \tag{5.4}$$
$$E_N : \Delta = b, j = 1/\Delta = b^{-1}. \tag{5.5}$$

Изоморфные кривые и так называемые кривые кручения имеют один и тот же j-инвариант. Кривые с нулевым j-инвариантом над некоторыми полями не рекомендуются для криптографических применений.

Можно заметить, что всегда можно найти кривую с заданным $j(E) = j_0$.

Пусть $a = 3k(\mathrm{mod}\, p), b = 2k(\mathrm{mod}\, p)$, тогда $j_0 = 1728\frac{k}{k+1}\mathrm{mod}\, p$ и $k = \frac{j_0}{1728 - j_0}\mathrm{mod}\, p$.

6 ЭЛЛИПТИЧЕСКИЕ КРИВЫЕ НАД ПРОСТЫМИ ПОЛЯМИ ГАЛУА

Рекомендуемыми в появившихся за последние годы стандартах являются два типа полей - простые поля Галуа F_p и расширенные поля F_2^m характеристики 2.

Рассмотрим свойства кривых E над простыми полями F_p, p > 3. В ряде случае операции над простым полем оказываются менее сложными (и более быстрыми), чем над расширенным полем. В частности, ГОСТ Российской Федерации Р34.10- 2001 рекомендует именно такое поле.

Пусть кривая (1.4) с целыми коэффициентами a и b определена над полем рациональных чисел и p > 3 - простое число. Тогда сравнение

$$y^2 = x^3 + ax + b(\mathrm{mod}\, p) \qquad (6.1)$$

называется редукцией кривой по модулю p. При этом и коэффициенты кривой, и координаты (x, y) точек являются целыми сравнимыми по модулю p числами.

Редукция называется хорошей, если p не делит дискриминант или

$$\Delta = -(4a^3 + 27b^2)\mathrm{mod}\, p \neq 0. \qquad (6.2)$$

В этом случае все корни кубического уравнения различны и кривая является несингулярной (без особых точек). Будем рассматривать только такие кривые.

Редукция (6.1) равнозначна переходу от поля Q к конечному полю, в частности, простому полю Галуа, т. е.

$$y^2 = x^3 + ax + b; a, b \in F_p. \qquad (6.3)$$

Абелеву группу точек (x, y), удовлетворяющих уравнению (6.3), вместе с точкой на бесконечности O обозначим E_p. В отличие от коэффициентов a, b кривой координаты (x, y) точек могут рассматриваться как элементы любого расширения $F_p^m (m = 1,2,3,...)$ поля F_p вплоть до алгебраического замыкания $\overline{F}_p$.

Законы сложения точек (2.5), (2.6) справедливы для группы E_p ($p \neq 2, 3$) после введения редукции по модулю p, а операция деления равнозначна умножению на обратный элемент поля F_p. Из конечности числа элементов поля, очевидно, следует и конечность числа точек кривой, т. е. ее порядка.

Так как кривая всегда содержит точку на бесконечности O, а для каждого решения x уравнения (6.3) имеются по два значения y, для числа N_E точек кривой можно получить грубую оценку

$$1 < N_E < 2p + 1, или \left|p + 1 - N_E\right| < p.$$

Более точную оценку порядка N_E эллиптической кривой E над конечным полем F_q получил в 1934 г. немецкий математик *Г. Хассе*.

Теорема (*Хассе*): Для эллиптической кривой E над конечным полем F_q, справедлива следующая оценка порядка кривой N_E

$$q+1-2\sqrt{q} \le N_E \le q+1+2\sqrt{q}, q=p^m, m=1,2,3,\ldots. \quad (6.4)$$

В частности, для простого поля она может быть записана как

$$|p+1-N_E| < 2\sqrt{p}. \quad (6.5)$$

Пусть $\chi(z)$ - квадратичный характер элемента z поля F_q, определяемый как

$$\chi(z) = \begin{cases} 1, z = a^2, a \in F_q, \\ -1, z \neq a^2, \\ 0, z = 0, \end{cases} \quad (6.6)$$

Иными словами, если z имеет корень квадратный в поле F_q, $\chi(z) = 1$, в противном случае $\chi(z) = -1$. В первом случае говорят, что z является квадратичным вычетом, во втором - квадратичным невычетом. Тогда с учетом (6.3) порядок кривой можно рассчитать перебором всех элементов поля F_q как сумму

$$\sum_{x \in F_q} \{\chi(x^3+ax+b)+1\} = q+1+\sum_{x \in F_q} \chi(x^3+ax+b) = q+1-t, \quad (6.7)$$

где

$$t = -\sum_{x \in F_q} \chi(x^3+ax+b).$$

Первая единица в выражении (6.7) учитывает точку на бесконечности O, а под знаком суммы каждое решение *x* уравнения (6.3) даст по две точки E.

Исключением являются точки второго порядка с координатой y=0, которые в соответствии с (6.6), (6.7) учитываются по одному разу. Значение t в (6.7), не превышающее по абсолютной величине $2\sqrt{q}$, может быть положительным или отрицательным в зависимости от преобладания квадратичных вычетов или невычетов $f(x) = x^3 + ax + b$.

7 МЕТОДЫ ЭКСПОНЕНЦИРОВАНИЯ ТОЧКИ ЭЛЛИПТИЧЕСКОЙ КРИВОЙ

Наиболее распространенной операцией во всех криптографических алгоритмах является *k*-кратное сложение точки *P*, обозначаемое как *kP*. Эту операцию обычно называют скалярным умножением, или, в терминологии мультипликативной группы, экспоненцированием точки кривой.

Дадим краткое описание методов повышения производительности при вычислении точки *kP*. Подход к расчету точки *kP* может отличаться в зависимости от того, является ли точка *P* фиксированной (заранее известной) или произвольной точкой. В первом случае всегда можно пользоваться предвычислениями точек, например, $2^i P$, которые хранятся в памяти. Двоичное представление числа *k* позволяет селектировать те из них, которые в результате суммирования образуют точку *kP*. Во втором, более общем случае, все вычисления приходится проводить в реальном времени.

Пусть порядок $Ord\ P = r_1[\log_2 r] = L$ и число k представлено в двоичной системе $k = \sum_{i=0}^{L-1} k_i 2^i$.

7.1 АЛГОРИТМ «УДВОЕНИЕ - СЛОЖЕНИЕ»

Самый простой метод, при котором вычисления осуществляются по формуле

$$kP = \sum_{i=0}^{L-1} k_i 2^i P = \sum_{i=0}^{L-1} k_i P_i, \quad P_i = 2^i P.$$

Вычисления осуществляются с помощью следующего алгоритма.

Вход: $k = (k_{L-1}, k_{L-2}, ..., k_0)$, $P \in E$. Выход: Kp.

1. $R \leftarrow 0$.
2. **for** $i = 0$ **to** $L-1$ **do**
   2.1. **if** $k_i = 1$ **then** $R \leftarrow R + P$.
   2.2. $P \leftarrow 2P$.
3. **return** R.

Реализация метода требует *(L - 1)* операций *D* удвоения точки и *W(k) - 1* сложений *A*, где *W(k)* - вес Хэмминга двоичного вектора *k* (число единиц этого вектора). Так как в среднем число единиц случайного вектора *k* равно *1/2*, общее число групповых операций оценивается величиной *0.5LA + LD*.

7.2 АЛГОРИТМ «УДВОЕНИЕ - СЛОЖЕНИЕ - ВЫЧИТАНИЕ»

Предыдущий алгоритм можно усовершенствовать, если ввести дополнительную операцию - вычитания точки. Например, число $k = 31$ в двоичной системе имеет вес *W(k)* = 5, но его можно представить как $2^5 - 1$ с весом 2. Снижение веса Хэмминга (и, соответственно, числа групповых операций) реализуется переходом от двоичного представления числа k к троичному *NAF(k)* с коэффициентами $k_i \in \{0,1,-1\}$ (*NAF* - non-adjacent form). Одно из свойств представления *NAF(k)* — отсутствие в нем смежных пар ненулевых элементов, благодаря чему возрастает удельный вес нулевых элементов k_i.

Для расчета *NAF(k)* используется следующий алгоритм.

Вход: положительное целое число *k*. Выход: *NAF(k)*.

1. $i \leftarrow 0$.
2. **while** $k \geq 1$ **do**
   2.1. **if** k **is odd then:** $k_i \leftarrow 2 - (k \bmod 4), k \leftarrow k - k_i$; *// odd – нечетное*
   2.2. **else** $k_i \leftarrow 0$;
   2.3. $k \leftarrow k/2, i \leftarrow i + 1$.
3. return $(k_0, k_1, ..., k_{L-1})$.

После расчета *NAF(k)* вычисляется точка kP методом слева - направо с помощью следующего алгоритма.

Вход: *NAF(k)*, $P \in E$. Выход: kP.

1. $R \leftarrow 0$.

2. **for** $i = 0$ **to** $L-1$ **do**
   2.1. **if** $k_i = 1$**then** $R \leftarrow R + P$.
   2.2. **if** $k_i = -1$**then** $R \leftarrow R - P$.
   2.3. $P \leftarrow 2P$.
3. return R.

NAF - представление числа k может оказаться на один бит длиннее двоичного. В то же время для случайного *k* вероятность ненулевых элементов 1 и -1 снижается от 1/2 до 1/3, т. е. в среднем для *L*-разрядного числа их количество оценивается величиной *L/3*. Тогда общее среднее число групповых операций сложения *A* и удвоения *D* можно оценить как сумму *(L/3)A + LD*.

7.3 МЕТОД ОКОН С АЛГОРИТМОМ «УДВОЕНИЕ – СЛОЖЕНИЕ – ВЫЧИТАНИЕ»

Если в криптосистеме есть резервы памяти, их можно задействовать для дальнейшего увеличения скорости вычислений. Вместо точки ***P*** можно экспоненцировать и в дальнейшем складывать смежные блоки или окна шириной ω в *NAF*-представлении точки kP.

Назовем ω-окном числа $NAF_\omega(k) = \sum_{i=0}^{L-1} k_i 2^i$ нечетный коэффициент $k_i, |k_i| < 2^{\omega-1}$, содержащий хотя бы один ненулевой элемент. Заметим, что $NAF_2(k) = NAF(k)$.

Например, при ω=4 имеем 8 различных k_i.

-7_{10}=(-1,0,0,1),
-5_{10}=(0,-1,0,-1),
-3_{10}=(0,-1,0,1),
-1_{10}=(0,0,0,-1),
1_{10}=(0,0,0,1),
3_{10}=(0,1,0,-1),
5_{10}=(0,1,0,1),
7_{10}=(1,0,0,-1).

Этих окон достаточно для формирования числа произвольной длины L. Четные коэффициенты в *NAF*-представлении числа k избыточны, так как они образуются удвоением нечетных. На первом этапе предвычислений рассчитываются и записываются в память 8 точек: ±P, ±3P, ±5P и ±7P.

В общем случае в памяти хранится $2^{\omega-1}$ точек. Число $NAF_\omega(k)$ может быть определено с помощью модифицированного алгоритма вычисления *NAF(k)*.

Модификация состоит в следующем: на шаге 2.1 вместо «$k_i \leftarrow 2 - (k \bmod 4)$» следует записать «$k_i \leftarrow k \bmod 2^\omega$», где $k \bmod 2^\omega$ означает целое число $u = k \bmod 2^\omega$, определенное в интервале $-2^{\omega-1} \le u \le 2^{\omega-1}$. Далее вычисляется точка kP по следующему алгоритму.

Вход: $\omega, NAF_\omega(k) = \sum_{i=0}^{L-1} k_i 2^i, P \in E$. Выход: kP.

1. $P_i = iP, i = 1,3,5,\ldots,2^{\omega-1}-1$.
2. $R \leftarrow 0$.
3. **for** $i = 0$ **to** $L-1$ **do**
   3.1. **if** $k_i \neq 0$ **then:**

if $k_i > 0$ **then** $R \leftarrow R + P_{ki}$

else $R \leftarrow R - P_{ki}$

3.2. $P \leftarrow 2P$.

4. **return** R.

Например, $\omega = 4$, $k = 249$, при этом $NAF(k) = (1,0,0,0,0,-1,0,0,1)$ и $NAF_4(k) = (1, 0, 0, 0, 0, 0, 0, 0, 0, -7_{10})$. Использование троичного $NAF(k)$ требует, очевидно, двух сложений точек, тогда как во втором случае за счет предварительного расчета точки -7P достаточно одного сложения. Число удвоений одинаково в обоих случаях. Ясно также, что выигрыш за счет окна появляется лишь при сравнительно больших длинах L числа k. Рост ширины ω окна ведет к увеличению сложности вычислений на первом шаге (и объема памяти) и снижению временной сложности на третьем шаге.

8 КРИПТОГРАФИЧЕСКИЕ ПРОТОКОЛЫ С ИСПОЛЬЗОВАНИЕМ ЭЛЛИПТИЧЕСКИХ КРИВЫХ

8.1 АНАЛОГ ОБМЕНА КЛЮЧАМИ ПО СХЕМЕ ДИФФИ-ХЕЛЛМАНА

По аналогии между экспоненцированием элементов в мультипликативной группе поля и k-кратным сложением точки эллиптической кривой можно построить протокол Диффи-Хеллмана.

Обмен ключами с использованием эллиптических кривых выполняется по следующей схеме. Сначала выбирается простое число $p \approx 2^{180}$ и параметры a и b для эллиптической кривой в уравнении $y^2 = x^3 + ax + b(\mathrm{mod}\, p)$. Это задает группу точек на эллиптической кривой. Затем в этой группе выбирается генерирующая точка $G = (x_1, y_1)$. При выборе G важно, чтобы наименьшее значение n, при котором $nG=O$ оказалось очень большим простым числом. Параметры p,a,b и G криптосистемы являются параметрами, известными всем участникам.

Обмен ключами между пользователями A и B проводится:

1. Сторона A выбирает целое число n_A, меньшее n. Это число будет личным ключом участника A. Затем участник A генерирует открытый ключ $P_A = n_A G$.

Открытый ключ представляет собой некоторую точку из группы точек на эллиптической кривой.

2. Точно так же участник B выбирает личный ключ n_B и вычисляет открытый ключ P_B.

3. Участник A генерирует секретный ключ $K = n_A P_B$, а участник B генерирует секретный ключ $K = n_B P_A$.

Два последних выражения дают один и тот же результат, поскольку $n_A P_B = n_A(n_B G) = n_B(n_A G) = n_B P_A$.

Чтобы взломать эту схему, противник должен будет вычислить k по данным G и kG, что предполагается трудной задачей.

Общий секретный ключ представляет собой пару чисел. Если этот ключ предполагается использовать в качестве сеансового ключа для симметричного шифрования, то из этой пары чисел необходимо генерировать одно подходящее значение, например, использовать просто координату x или некоторую простую функцию от x...

Пример 4. Возьмем $p = 211$, $E_{11}(0,-4)$ (что соответствует кривой $y^2 = x^3 - 4$) и $B = (2, 2)$. Можно подсчитать, что $241B = O$. Личным ключом пользователя А является $a = 121$, поэтому открытым ключом A будет $aB = 121(2, 2) = (115, 48)$. Личным ключом пользователя Б является $b = 203$, поэтому его открытым ключом будет $203(2, 2) = (130, 203)$. Общим секретным ключом является $121(130, 203) = 203(115, 48) = (161, 169)$.

Протокол Диффи-Хеллмана, однако, не защищен от противника С, который имеет доступ к каналу связи и может подменять пересылаемые точки P_A и P_B своими точками $P_C = n_C G$. Он, таким образом, может либо выступать от имени одного из пользователей, установив секретную связь с другим, либо, контролируя канал, быть транслятором их переписки, свободно расшифровывая и читая все сообщения. Такого активного криптоаналитика С называют «man in the middle». Для защиты от перехвата и подлога чрезвычайно важной становится задача аутентификации пользователей. Рассмотрим протокол, в котором устраняется недостаток протокола Диффи-Хеллмана.

8.2 ПРОТОКОЛ РАСПРЕДЕЛЕНИЯ КЛЮЧЕЙ MQV

Протокол распределения ключей на эллиптической кривой (ECKEP-Elliptic Curve Key Establishment Protocol), предложенный Менезисом, Кью и Ванстоуном (A.Menezes, M. Qu, S. Vanstone) и называемый *MQV*-протоколом, предполагает использование как долговременных ключей d_A и d_B, так и разовых ключей k_A, k_B пользователей. Для формирования общего секретного ключа KAB пользователи выполняют следующие действия:

1. Пользователь А:
   - а) генерирует случайное целое число $0 < k_A < n$;
   - б) вычисляет точку $R_A = k_A G = (x_A, y_A)$;
   - в) вычисляет точку $Y_A = k_A Q_B = (x_1, y_1)$;
   - г) вычисляет целое число $S_A = (k_A + d_A x_A x_1) \bmod n$;
   - д) отправляет точку R_A пользователю В.
2. Пользователь В:
   - а) генерирует случайное целое число $0 < k_B < n$;
   - б) вычисляет точку $R_B = K_B G = (x_B, y_B)$;
   - в) вычисляет точку $Y_B = k_B Q_A = (x_2, y_2)$;
   - г) вычисляет целое число $S_B = (k_B + d_B x_B x_2) \bmod n$;
   - д) отправляет точку R_B пользователю А.
3. Пользователь А:
   - а) вычисляет точку $Y_B = d_A R_B = (x_2, y_2)$;
   - б) вычисляет точку $K = S_A(R_B + x_B x_2 Q_B)$.
4. Пользователь В:
   - а) вычисляет точку $Y_A = d_B R_A = (x_1, y_1)$;
   - б) вычисляет точку $K = S_B(R_A + x_A x_1 Q_A)$.

В результате вычислений пользователи получают одну и ту же точку K.

Действительно, для пользователя А с учетом равенства $Q_B = d_B G$ и согласно соотношений 2 (б), (г) и 3 (б) имеем

$$K = S_A(R_B + x_B x_2 Q_B) = S_A(k_B + d_B x_B x_2)G.$$

Аналогично вычисления пользователя В согласно 4 (б) дают

$$K = S_B(R_A + x_A x_1 Q_A) = S_B(k_A + d_A x_A x_1)G.$$

В связи с коммутативностью операции сложения в группе точек Е результаты совпадают. Координаты секретной точки К могут быть известным способом преобразованы в разовый ключ симметричного шифрования.

Протокол *MQV* можно модифицировать без расчета параметров S_A, S_B согласно 1 (г) и 2 (г), а вычисляя точку K тождественными операциями с точками кривой Е. Для этого в пункте 3 пользователь А вычисляет точку $V = (R_B + x_B x_2 Q_B)$, а затем точки $k_A V$ и $d_A x_A x_1 V$. Сумма двух последних точек дает точку К. Так же действует и пользователь В. Этот алгоритм, более трудоемок, так как операции с точками сложней операций в поле F_q.

При выполнении этого протокола у активного криптоаналитика С («man in the middle») возникают проблемы. Если бы пользователи просто складывали свои секретные ключи, т. е. $S_{A,B} = (k_{A,B} + d_{A,B}) \bmod n$, то ситуация для С будет такой же благоприятной, как и в протоколе Диффи- Хеллмана. Идея защиты от навязывания противником С своего разового ключа k_C состоит в том, что соотношение

$$S_A = (k_A + d_A x_A x_1) \bmod n$$

согласно пункту 1(в) нелинейно связывает ключи k_A, d_A и d_B, так как $x_1 = f(k_A, k_B)$. Тем самым противник С лишается возможности свободной подтасовки ключа k_C. Чтобы рассчитать свое значение S_A, в котором ему известно лишь x_A, противнику С придется определить долговременные ключи d_A и d_B т. е. дважды вычислить дискретный логарифм в группе Е.

8.3 ЭЛЕКТРОННАЯ ЦИФРОВАЯ ПОДПИСЬ (ЭЦП)

Цифровая подпись для сообщения является числом, зависящим от самого сообщения и от секретного, известного только подписывающему субъекту ключа. При этом подпись должна быть легко проверяемой без знания секретного ключа. При возникновении спорной ситуации, связанной с отказом подписавшего от факта подписи им некоторого сообщения либо с попыткой подделки подписи, третья независимая сторона (арбитр) должна иметь возможность разрешить спор.

Применение ЭЦП позволяет решить следующие задачи:

1) осуществить аутентификацию источника сообщения;

2) установить целостность сообщения;

3) обеспечить невозможность отказа от факта подписи конкретного сообщения.

В настоящее время используются различные схемы ЭЦП. Их можно разделить на три класса:

1) схемы на основе симметричных систем шифрования;

2) схемы на основе систем шифрования с открытыми ключами;

3) схемы со специально разработанными алгоритмами вычисления и проверки;

4) подписи.

Замечание. Распространенной практикой является формирование ЭЦП не для самого сообщения, а для его хеш-образа при соответствующем выборе хеш-функции.

8.4 ГОСТ Р34.10-2001

ГОСТ Р34.10-2001 принят в Российской Федерации в октябре 2001 года . Он регламентирует процедуры формирования и проверки электронной цифровой подписи (ЭЦП) на основе арифметики эллиптических кривых, определенных над простым полем Галуа (предыдущий ГОСТ Р34.10-1994, принятый в 1994 году, основан на проблеме дискретного логарифмирования в конечном поле – схема Эль - Гамаля).

Общесистемные параметры в стандарте определены с ограничительными условиями:

F_p - простое поле Галуа порядка $p > 2^{255}$;

E - эллиптическая кривая $y^2 = x^3 + ax + b$ над полем F_p с числом точек $m = nq$ и дискриминантом $(4a^3 + 27b^2) \neq 0$; эллиптическая кривая задается через величину

$$J(E) = 1728\frac{4a^3}{4a^3 + 27b^2} \bmod p$$

(затем рассчитывается $k = \frac{J(E)}{1728 - J(E)} \bmod p, a = 3k(\bmod p), b = 2k(\bmod p)$);

P - генератор криптосистемы простого порядка ; $h(M)$ - хэш-функция, отображающая двоичные сообщения M произвольной длины в двоичный вектор длины 256 бит. Хэш-функция определена в ГОСТ Р34.11-1994.

Кроме этих ограничений, общесистемные параметры криптосистемы проверяются на стойкость с помощью трех тестов:

1. MOV-атака. Порядок q криптосистемы не должен делить порядок мультипликативной группы расширенного поля Галуа $(p^k - 1)$ с расширением $k = 2,3,...,B, B \geq 31$.

2. Тест на аномальность кривой. Должно выполняться условие $m \neq p$. В противном случае кривая аномальна и не приемлема для криптографии.

3. Тест на инвариант (и суперсингулярность), j-инвариант кривой $J(E) \neq 0 \bmod 1728$. Это, в частности, требует выполнения условий $a \neq 0$ и $b \neq 0$ для уравнения кривой.

Параметры пользователя

Пользователь А:

- генерирует и хранит в секрете долговременный секретный ключ как целое число $0 < d_A < q$;
- вычисляет открытый ключ как точку кривой $Q_A = d_A P$. Открытый ключ доступен для всех пользователей системы.

Формирование ЭЦП

Пользователь А:

1. Вычисляет хэш значение сообщения M: $e = h(M) \bmod q$. Если $e = 0$, то принять $e = 1$.
2. Генерирует случайное целое число $0 < k_A < q$.
3. Вычисляет точку $C = k_A P = (x_C, x_C)$.
4. Вычисляет параметр $r = x_C \bmod q$. При $r = 0$ возврат в пункт 2.
5. Вычисляет параметр $s = (k_A e + d_A r) \bmod q$. При $s = 0$ возврат в пункт 2.
6. Определяет цифровую подпись $\zeta = (r \| s)$ - в виде конкатенации двух двоичных векторов r и s.

Проверка ЭЦП

Пользователь В проверяет цифровую подпись пользователя А с помощью его открытого ключа Q_A, общесистемных параметров, алгоритма хэширования $h(M)$ и подписанного сообщения (M, ζ). Проверка заключается в вычислении на основе известных данных параметра r' и сравнении его с принятым значением r.

Умножив равенство в пункте 5 на инверсию e^{-1} и учитывая, что $Q_A = d_A P$, для точек криптосистемы получим равенство:

$k_A P = e^{-1} SP - e^{-1} r d_A P = uP + vQ_A$;

$u = e^{-1} S \bmod q$;

$v = -e^{-1} r \bmod q$.

Протокол проверки ЭЦП включает следующие вычисления.

Пользователь В:

1. По полученной подписи ζ вычисляет целые числа r и s. Если $0 < r < q$, $0 < S < q$, то перейти к следующему шагу. В противном случае подпись неверна.
2. Вычисляет хэш-значение полученного сообщения $h = h(M)$ как целое число.
3. Вычисляет $e = h \bmod q$. Если $e = 0$, то принять $e = 1$.
4. Вычисляет обратный элемент $e^{-1} \bmod q$ мультипликативной группы поля F_q.
5. Вычисляет параметры $u = e^{-1} S \bmod q, v = -e^{-1} r \bmod q$.
6. Вычисляет точку $C' = uP + vQ_A = (x_C', y_C')$.
7. Вычисляет параметр $r' = x_C' \bmod q$.
8. Сравнивает вычисленное r' и принятое r значения. При равенстве $r = r'$ цифровая подпись принимается, в противном случае она неверна.

При проверке требуется вычисление одной инверсии e^{-1} в поле. Кроме алгоритмов формирования и проверки ЭЦП, в ГОСТ Р34.10-2001 приведены общие положения, математические определения, а также контрольный пример процессов формирования и проверки подписи для заданных параметров схемы ЭЦП.

8.5 ECDSA

Алгоритм *ECDSA* был впервые предложен С. Ванстоуном в 1992 году, после чего он прошел длительный этап исследований на стойкость и различного рода усовершенство-

ваний и доработок. Лишь в конце XX века он был утвержден в ряде стандартов цифровой подписи: международном стандарте *ISO/IEC CD 15946-2* в1999 году, американском стандарте для финансовых служб *X9.62 ANSI (American National Standard Institute)* в 1999 года, американском национальном стандарте *FIPS 186-2 NIST* в 2000 года, *P1363 IEEE (Institute of Electrical and Electronics Engineers)* в 2000 году и других. Национальный стандарт *DSS* правительства США, принятый *NIST* в 2000 году, сменил предыдущий стандарт *FIPS* 186-1, действующий с 1994 года, и рекомендует *ECDSA* взамен *DSA*, построенного на арифметике простого поля Галуа. Вообще по новому стандарту ЭЦП может вырабатываться по одному из трех алгоритмов: *DSA* – на основе проблемы дискретного логарифмирования в конечном поле, *RSA DSA* – вариант схемы *RSA* и *ECDSA*. Сам алгоритм *ECDSA* в стандарте не приведен (ссылкой на описание алгоритма в стандарте *X9.62 ANSI*).

Описание ECDSA

Параметры пользователя

Пользователь А:

- генерирует и хранит в секрете долговременный секретный ключ как целое число $0 < d_A < n$;
- вычисляет открытый ключ как точку кривой $Q_A = d_A G$. Открытый ключ доступен для всех пользователей системы.

Формирование ЭЦП

Пользователь А:

1. Вычисляет хэш-значение сообщения M как целое число $e = h(M), e < n$.
2. Генерирует случайное целое число $0 < k_A < n$.
3. Вычисляет точку $R = k_A G = (x_1, y_1)$.
4. Вычисляет параметр $r = \pi(R) \bmod n$. При $r = 0$ возврат в пункт 2.
5. Вычисляет обратный элемент k_A^{-1} простого поля F_n.
6. Вычисляет параметр $s = k_A^{-1}(e + d_A r) \bmod n$. При $s = 0$ возврат в пункт 2.
7. Направляет пользователю В подписанное сообщение *(M, r, s)*, в котором $DS = (r, s)$ - цифровая подпись.

В пункте 4 преобразование точки R в целое число предполагает, что ее x - координата как элемент x_1 поля F_n переводится в целое число $\overline{x_1}$ с последующей редукцией по модулю $n (r = \pi(R) \bmod n = \overline{x_1} \bmod n)$.

Проверка ЭЦП

Пользователь В проверяет ЭЦП пользователя А, имея в распоряжении следующую информацию: открытый ключ пользователя А Q_A, общесистемные параметры, алгоритм хэширования *h(M)* и подписанное сообщение *(M, r, s)*. Суть проверки состоит в вычислении на основе известных данных параметра r' и сравнении его с принятым значением r.

Умножив равенство в пункте 6 на инверсию s^{-1} второго параметра подписи и учитывая, что $Q_A = d_A G$, для точек криптосистемы в результате экспоненцирования (скалярного произведения) получим равенство

$k_A G = s^{-1} eG + s^{-1} r d_A G = uG + vQ_A$;

$u = s^{-1} e \bmod n$;

$v = s^{-1} r \bmod n$.

Согласно пунктам 3 и 4 протокола формирования левая часть этого равенства определяет точку $R = (x_1, y_1)$ и, соответственно, параметр $r = \overline{x_1} \bmod n$.

Правая часть равенства включает известные получателю данные, которые он использует для вычисления параметра r' (он может оказаться отличным от параметра r при модификациях сообщения M и ошибках в канале связи).

Протокол проверки ЭЦП включает следующие вычисления

Пользователь В:

1. Вычисляет хэш-значение полученного сообщения $M: e = h(M), e < n$.
2. Вычисляет обратный элемент $s^{-1} \bmod n$ поля F_n.
3. Вычисляет параметры $u = s^{-1} e \bmod n, v = s^{-1} r \bmod n$.
4. Вычисляет точку $R' = uG + vQ_A = (x_1, y_1)$.
5. Вычисляет параметр $r' = \pi(R') = \overline{x_1} \bmod n$.
6. Сравнивает вычисленное r' и принятое значения r. При равенстве $r' = r$ цифровая подпись верна, в противном случае она отвергается.

В результате проверки пользователь В удостоверяется в подлинности отправителя А и целостности сообщения М.

В ряде проектов и стандартов определение параметра r подписи не регламентируется, а задается функцией $r = \pi(x_1, y_1) \bmod n$. Это, в частности, позволяет избежать неоднозначности определения $r = \overline{x_1} \bmod n$ в связи с наличием обратной точки $(n - k_A)G = -k_A G$, имеющей ту же x - координату, что и точка $k_A G$(и, следовательно, совпадающий параметр r).

FIPS 186-2-2000 рекомендует к использованию 10 полей и 15 эллиптических кривых, 5 из которых определены над простыми полями F_p и 10 – над расширенными полями F_{2^m}. Значения модулей простых полей:

$P_{192} = 2^{192} - 2^{64} - 1$;

$P_{224} = 2^{224} - 2^{96} + 1$;

$P_{256} = 2^{256} - 2^{224} + 2^{192} + 2^{96} - 1$;

$P_{384} = 2^{384} - 2^{128} + 2^{96} + 2^{32} - 1$;

$P_{521} = 2^{521} - 1$.

Расширения двоичного поля равны *m = 163, 233,283,409 и 571.*

Из кривых над простыми полями F_P *NIST* рекомендует кривую

$$y^2 = x^3 - 3x + b \bmod p.$$

Значения коэффициента b для пяти рекомендуемых кривых, координаты одного из возможных генераторов - точки $G = (G_x, G_y)$ порядка n приведены в стандарте.

Из 10 кривых над расширенными полями F_2^m 5 несуперсингулярных кривых $y^2 + xy = x^3 + ax^2 + b$ с коэффициентами $a = 1$ и $b \in F_2^m, b \neq 0,1$.

Еще 5 несуперсингулярных кривых вида $y^2 + xy = x^3 + ax^2 + 1$ над полем F_2^m с коэффициентами $a = 1$ или 0. Эти кривые называют кривыми Коблица, в стандарте они обозначены как $(K-m)$. В стандарте *FIPS 186-2-2000*, кроме порядков кривых, даны возможные значения координат точек G порядка n.

Кривые Коблица - наиболее технологичные кривые над полем характеристики 2, они обеспечивают наивысшую производительность вычислений в поле F_2^m В то же время они относятся к классу аномальных кривых, что снижает их стойкость в $\sqrt{m}$ раз по сравнению с кривыми $(B-m)$ с произвольным значением коэффициента b и таким же порядком.

Достаточно большой диапазон размеров поля и порядков криптосистем позволяет реализовать системы с различной степенью безопасности, работающие совместно с симметричными блочными шифрами с разной длиной ключа. Однако с ростом размера модуля падает скорость криптопреобразований и, соответственно, растет время шифрования и обмена ключами, формирования и проверки подписи.

В целом, при одинаковых алгоритмах и программной реализации арифметика эллиптических кривых над простым полем F_p выполняется в 2-3 раза быстрее, чем в поле F_2^m.

ПРИЛОЖЕНИЕ А. КАНОНИЧЕСКИЕ УРАВНЕНИЯМ С ВЫРАЖЕНИЯМИ АРИФМЕТИЧЕСКИХ ОПЕРАЦИЙ ДЛЯ ПРИВЕДЕННЫХ СЛУЧАЕВ

1. Тип поля и вариант кривой:

Поле характеристики, отличной от 2 и 3.

Каноническое уравнение кривой:

$$y^2 = x^3 + ax + b;$$

Формула сложения:

$$x = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)^2 - x_1 - x_2;$$

$$y = -y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x_1 - x);$$

Формула удвоения:

$$x = \left(\frac{3x_1^2 + a}{2y_1}\right) - 2x_1;$$

$$y = -y_1 + \frac{3x_1^2 + a}{2y_1}(x_1 - x);$$

2. Тип поля и вариант кривой:

Поле характеристики 3

Каноническое уравнение кривой:

$$y^2 = x^3 + ax^2 + bx + c;$$

Формула сложения:

$$x = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)^2 - x_1 - x_2 - a;$$

$$y = -y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x_1 - x);$$

Формула удвоения:

$$x = \left(\frac{ax_1 - b}{y_1}\right)^2 - a + x_1;$$

$$y = -y_1 + \frac{ax_1 - b}{y_1}(x_1 - x);$$

3. Тип поля и вариант кривой:

Поле характеристики 2, суперсингулярная кривая

Каноническое уравнение кривой:

$$y^2 + ay = x^3 + bx + c;$$

Формула сложения:

$$x = \left(\frac{y_2 + y_1}{x_2 + x_1}\right)^2 + x_1 + x_2;$$

$$y = y_1 + a + \frac{y_2 + y_1}{x_2 + x_1}(x_1 + x);$$

Формула удвоения:

$$x = \left(\frac{x_1^4 + b^2}{a^2}\right);$$

$$y = y_1 + a + \frac{x_1 + b}{a}(x_1 + x);$$

4. Тип поля и вариант кривой:

Поле характеристики 2, несингулярная кривая

Каноническое уравнение кривой:

$$y^2 + axy = x^3 + bx^2 + c;$$

Формула сложения:

$$x = \left(\frac{y_2 + y_1}{x_2 + x_1}\right)^2 + \frac{y_2 + y_1}{x_2 + x_1} + x_1 + x_2 + b;$$

$$y = y_1 + x + \frac{y_2 + y_1}{x_2 + x_1}(x_1 + x);$$

Формула удвоения:

$$x = x_1^2 + \frac{y_1^2}{x_1^2} + x_1 + \frac{y_1}{x_1} + b;$$

$$y = x_1^2 + \frac{x_1^2 + y_1}{x_1}x + x;$$

ЛИТЕРАТУРА

1 Бессалов, А.В. Криптосистемы на эллиптических кривых: Учебное пособие / А.В. Бессалов, А.Б. Теликенко. – К.: Политехника, 2004.

2 Болотов, А.А. Элементарное введение в эллиптическую криптографию: Протоколы криптографии на эллиптических кривых / А.А. Болотов, С.Б. Гашков, А.Б. Фролов. – М., 2006.

3 Ростовцев, А.Г. Алгебраические основы криптографии. — СПб.: Мир и Семья, 2000.

4 Ростовцев, А.Г. Введение в криптографию с открытым ключом / А.Г. Ростовцев, Е.Б. Маховенко. – СПб.: Мир и Семья, 2001.

5 Столингс, В. Криптография и защита сетей: принципы и практика. – 2-е изд. – М.: Издательский дом "Вильямс", 2001.

6 Boneh, D. Identity-based encryption from the Weil pairing, Crypto'2001: Lecture Notes in Computer Science, Springer-Verlag, 2001.

7 Digital Signature Standard: FIPS 186-2-2000 / National Institute of Standard and Technology. – 2000.

8 James, A. Muir and Douglas R. Stinson. Minimality and other proprieties of the widthw nonadjacent form / University of Waterloo, Canada, 2004.

9 Jue-Sam Chou and Yalin Chen and Jin-Cheng Huang. A ID-Based Deniable Authentication Protocol on pairings. Cryptology ePrint Archive, Report 2006/335.

10 Paterson, G. ID-based signatures from pairings on elliptic curves. Cryptology ePrint archive: Report 2002/004.

11 K. Phani Kumar and G. Shailaja and Ashutosh Saxena. Identity Based Strong Designated Verifier Signature Scheme. Cryptology ePrint Archive, Report 2006/134.

12 Quan Yuan and Songping Li, A New Efficient ID-Based Authenticated Key Agreement Protocol, Cryptology ePrint Archive, Report 2005/309.

УЧЕБНОЕ ИЗДАНИЕ

Составитель: Хацкевич Мария Викторовна

ПРИМЕНЕНИЕ ЭЛЛИПТИЧЕСКИХ КРИВЫХ В КРИПТОГРАФИИ

Методические указания к выполнению лабораторных работ
по дисциплине
«Криптографические методы защиты информации»

для студентов специальности
1-40 03 01 «Искусственный интеллект»

Ответственный за выпуск: Хацкевич М.В.
Редактор: Боровикова Е.А.
Компьютерная верстка: Горун Л.Н.
Корректор: Никитчик Е.В.

Подписано к печати 30.07.2012 г. Бумага «Снегурочка». Формат 60x84 1/16.
Гарнитура Arial Narrow. Усл. печ. л. 1,63. Уч. изд. л. 1,75.
Заказ № 840. Тираж 40 экз. Отпечатано на ризографе Учреждения образования
«Брестский государственный технический университет»
224017, г. Брест, ул. Московская, 267.