# ІНФОРМАТИКА ТА МЕТОДИКА ЇЇ ВИКЛАДАННЯ

УДК 004:003.26

**Пірус Є.М., Дікарєв С.С.**

[1] *старший викладач кафедри алгебри, ДВНЗ «ДДПУ»*

[2] *студент 5 курсу фізико-математичного факультету, ДВНЗ «ДДПУ»*

e-mail: pirus@ukr.net

## ПРОГРАМНА РЕАЛІЗАЦІЯ АЛГОРИТМІВ АСИМЕТРИЧНИХ МЕТОДІВ ШИФРУВАННЯ ІНФОРМАЦІЇ ЗАСОБАМИ СЕРЕДОВИЩА ПРОГРАМУВАННЯ LAZARUS

Дана робота присвячена вивченню проблеми практичної реалізації алгоритмів криптографічних систем з відкритим ключем та побудови програм реалізацій алгоритмів таких криптосистем засобами середовища програмування Lazarus. Об'єктом дослідження є криптографічні системи з відкритим ключем. Предметом дослідження виступають криптографічні протоколи та математичний апарат, які дозволяють реалізовувати криптографічні системи з відкритим ключем на основі поняття однонаправленої функції.

**Ключові слова:** *алгоритм, шифр, однонаправлена функція, протокол, шифрування інформації.*

## Вступ

Інформацію, яка має певну цінність, треба захищати від двох небезбек — природних (інформація може бути пошкоджена або знищена при поломках апаратури, внаслідок шумів у каналах зв'язку, і т.п.) і від зловмисників. Захистом від першої небезбеки — головним чином, від пошкодження інформації при передачі — займається теорія кодування. Тут головний метод захисту — ще до передачі таким чином перетворити (закодувати) текст, ввівши в нього додаткові символи, щоб навіть у разі пошкодження не дуже великої частини символів можна було з великою ймовірністю відновити початковий текст. Найпростіший спосіб — продублювати кожен символ кілька разів. Таким чином, із можливими втратами інформації борються за рахунок надлишковості тексту, що передається.

Головна проблема теорії кодування — як поєднати надійність, швидкість передачі і зручність кодування. Іншими словами, як передавати інформацію швидко, дешево і надійно.

Захист інформації від зловмисників розпадається на 2 проблеми: захист від підміни або несанкціонованої модифікації (охорона *аутентичності* або *інтегральності* інформації) і захист від несанкціонованого доступу або відтворення (охорона *конфіденційності*, таємності, секретності інформації або прав власності). Обидві проблеми виникають як при зберіганні інформації, так і при її передачі, і мають багато дуже різних аспектів — організаційних, технічних, юридичних, тощо.

Є дуже багато різних пасивних методів захисту при зберіганні, коли сама інформація не змінюється, а лише робиться більш складним несанкціонований доступ до неї: грифи, сейфи, ключі і т.п. Один із методів пасивного захисту інформації при передачі — приховування самого факту передачі.

Предметом *криптографії* є активні методи захисту аутентичності і конфіденційності інформації при її зберіганні чи передачі *відкритими* каналами зв'язку. Це робиться за допомогою перетворення (шифрування) інформації для унеможливлення як несанкціонованого доступ до неї, так і незаконної її модифікації. При цьому вважається, що хоча у перетвореному вигляді повідомлення і може стати доступним зловмиснику, але він не зможе витягти з нього захищену інформацію.

Протягом тисячоліть криптографія була мистецтвом засекречування важливої державної інформації при передачі її по захищених каналах зв'язку, а криптоаналіз був двоїстим криптографії мистецтвом розкриття такої інформації. Тому криптологія, що об'єднує в собі криптологію та криптоаналіз історично раніше перебували майже виключно у військових і дипломатичних відомств. Однак у нинішній період здійснення у всьому світі комп'ютерної революції , коли величезна кількість персональної, фінансової, комерційної та технологічної інформації зберігається на комп'ютерних банках даних і пересилається по інформаційних комп'ютерних мережах, надзвичайно важливим є те обставина, що в суспільстві з'являється найгостріша потреба у громадянській криптографії.

Серед вимог, які висуваються до шифрів, головними є *ефективність* (шифрування і дешифрування повинні відбуватися швидко) і *надійність* (зловмисник не повинен встигнути зламати шифр за той час, доки зашифрована інформація має лишатися таємною). Позаяк одночасно і повністю ці вимоги задовольнити не вдається, при виборі конкретного шифру доводиться шукати компромісу. В одних випадках термін засекреченості інформації (наприклад, біржової) вимірюється кількома годинами чи навіть хвилинами, тому на шифрування і передачу відводяться лічені секунди. У той же час деякі державні чи комерційні таємниці повинні зберігатися десятиліттями, зате

можна не поспішати при шифруванні.

Криптографія — це галузь, яка вивчає тайнопис (криптографію) та методи її розкриття (криптоаналіз), яка за влучним висловом Рональда Ривеста, професора — Массачусетського технологічного інституту — і одного з авторів знаменитої криптосистеми RSA, є повитухою всієї «computer science» взагалі.

Побудова сучасної криптології як науки грунтується на сукупності фундаментальних понять і фактів математики, фізики, теорії інформації та складності обчислень, природно дуже складних для всебічного і глибокого осмислення навіть професіоналами. Однак, незважаючи на органічно притаманну їй складність, багато теоритичних досягнень криптології, зараз широко використовуються в нашому насиченому інформаційними технологіями житті, наприклад: в пластикових smart-картки, в електронній пошті, в системах банківських платежів, при електронній торгівлі через Internet, в системах електронного документообігу, при веденні баз даних, системах електронного голосування та ін.

З іншого боку, саме загальна потреба і широкий спектр можливостей практичного використання стимулюють теоретичні та прикладні дослідження не тільки в цій галузі знань і у відповідних галузях математики, фізики, теорії інформації та теорії обчислень, але також наполегливо спонукають до вдосконалення юридичних та правових норм і механізмів на державному, міжнародному і загальнолюдському рівні, що часто породжує відкриті обговорення в пресі або палкі дебати в парламентах різних країн, і навіть змушують обговорювати пов'язані з цим питання на наради глав великих держав.

Дана робота присвячена вивченню проблеми практичної реалізації алгоритмів криптосистем з відкритим ключем та побудови програм реалізацій алгоритмів таких криптосистем засобами середовища програмування Lazarus.

**Об'єкт** дослідження — криптографічні системи з відкритим ключем.

**Предмет** дослідження — криптографічні протоколи та математичний апарат, які дозволяють реалізовувати криптографічні алгоритми з відкритим ключем на основі поняття однонаправленої функції.

**Мета** дослідження полягає у вивченні криптографічних систем з відкритим ключем та реалізація навчальних варіантів таких систем засобами середовища програмування Lazarus.

У 1976 році американці Уітфілд Діффі та Мартін Геллман (Diffi W., Hellman M.) в статті «Нові напрямки в криптографії» запропонували новий принцип побудови криптосистем, які не вимагають не тільки передачі ключа приймаючому сповіщення, але й збереження в тайні методу шифрування. Ці

шифри дозволяють легко зашифрувати та дешифрувати текст і їх дозволяється використовувати багато разів.

В 1974 році Меркл (Merkle) винайшов механізм узгодження криптографічного ключа шляхом явних асиметричних обчислень, які отримали назву головоломка Меркла. Асиметричність головоломки Меркла полягає в тому, що її обчислювальна складність для законних учасників протоколу узгодження ключа і для перехоплювачів зовсім різна: легальні учасники легко проробляють обчислення, а нелегальні — ні. Головоломка Меркла представляє собою першу ефективну реалізацію однонаправленої функції з секретом.

Тільки зараз стало відомо, що Кокс (Cocks), британський криптограф, винайшов першу криптосистему з відкритим ключем в 1973 році. Алгоритм щифрування Кокса, який отримав назву *алгоритму з несекретним ключем шифрування*, використовує складність розкладу цілого числа на прості множники і співпадає з системою RSA. Тільки в 1997 році група з електронного захисту засобів зв'язку розсекретила алгоритм Кокса.

В 1978 році Р. Рівест, А. Шамір и Л. Адлеман (R.L.Rivest, A.Shamir, L.Adleman) запропонували приклад функції, яка має ряд гарних властивостей. На її основі була побудована реально використовувана система шифрування, ця криптоситсема отримала назву RSA (по першим літерам прізвищ авторів).

## Основні означення та зауваження

Поняття *однобічної функції* введене в 1975 Діффі і Геллманом. Під цим розуміється таке бієктивне відображення $f : X \to Y$, що значення $f(x)$ обчислюються «легко», а для випадково вибраного $y$ значення оберненої функції $f^{-1}(y)$ обчислюється «важко». Іншими словами, *однобічною* називається функція $f : X \to Y$, яка задовольняє такі дві умови:

а) існує поліноміальний алгоритм обчислення $f(x)$;

б) не існує поліноміального алгоритму інвертування функції $f(x)$ (тобто знаходження якого–небудь розв'язку рівняння $f(x) = y$ відносно $x$).

Часом вимагають більше: для майже всіх випадково вибраних $y$ «важко» знайти навіть яку-небудь часткову характеризацію $x = f^{-1}(y)$.

Більш строге означення виглядає наступним чином. Нехай $\Sigma$ — скінченний алфавіт. Для довільної функції $f : \Sigma^* \to \Sigma^*$ через $m(n)$ позначимо найменше $m$, для якого $f(\Sigma^n) \subseteq \bigcup_{i=1}^{m} \Sigma^i$. Функція $f$ називається *чесною*, якщо існує такий поліном $p(n)$, що $p(m(n)) \geqslant n$ для всіх $n$.

Чесна функція $f$ називається *однобічною*, якщо

а) існує поліноміальний алгоритм обчислення $f(x)$;

б) які б поліном $p(n)$ і поліноміальну ймовірносну машину Тьюрінґа $A$ ми не взяли, для всіх достатньо великих $n$ і випадково вибраного слова $x \in \Sigma^n$ виконується нерівність

$$\mathbf{P}\{f(A(f(x))) = f(x)\} < 1/p(n)\,.$$

Друга умова означає, що кожна поліноміальна ймовірносна машина Тьюрінґа може знайти який–небудь розв'язок рівняння $f(x) = y$ лише із знікомою ймовірністю. А «чесність» потрібна, щоб функція $f$ не занадто «стискала» вхідні дані (якщо $y$ набагато коротше за $x$, то машині може просто не вистачити часу, щоб виписати розв'язок $x$).

До нинішнього часу для жодної функції немає строгого доведення її однобічності.

**1–й конкретний приклад** «практично» однобічної функції запропонував Purdy в 1974 р. — це функція $f : \mathbb{Z}_p \to \mathbb{Z}_p$, де $p = 2^{64} - 59$, а $f(x)$ має вигляд

$$f(x) = x^{2^{24}+17} + a_1 x^{2^{24}+3} + a_2 x^3 + a_3 x^2 + a_4 x + a_5,$$

де $a_1, \ldots, a_5$ — довільні 19-цифрові числа.

Для існування стійких (з обчислювальної точки зору) криптосистем з відкритим ключем необхідне існування однобічних функцій.

Під *шифром* (або *криптосистемою*) будемо розуміти сукупність *алгоритму шифрування* $E$ і *алгоритму дешифрування* $D$ разом із певною родиною (простором) так званих *ключів*. Алгоритми шифрування і дешифрування визначають загальну структуру шифру, а ключ задає конкретні параметри цієї структури. Наприклад, у шифрі циклічного зсуву ключем є величина зсуву.

Ключ $k$ складається з двох компонент: ключа шифрування $k_1$ і ключа дешифрування $k_2$. У симетричних криптосистемах зв'язок між компонентами настільки простий (наприклад, у шифрах перестановки це дві взаємно обернені підстановки), що зазвичай вказують лише першу з них. Однак в асиметричних криптосистемах відновлення однієї компоненти за іншою є складною математичною задачею.

Якщо *криптотекст* $c$ одержується в результаті застосування до відкритого повідомлення $m$ алгоритму шифрування $E$ з ключем $k$, то пишемо $c = E_k(m)$. Аналогічно запис $m = D_k(c)$ означає, що повідомлення $m$ одержане з криптограми $c$ за допомогою алгоритму дешифрування $D$ з ключем $k$.

*Асиметричними* називаються криптосистеми, в яких для шифрування і дешифрування використовуються різні ключі, причому ключі пов'язані та-

ким чином, що визначення одного з них за відомим іншим є з обчислювальної точки зору дуже важкою задачею. Тому один із ключів можна не тримати в таємниці і зробити загальнодоступним (звідси — інша назва таких криптосистем: *криптосистеми з відкритим ключем*). Системи з відкритим ключем шифрування використовуються головним чином для передачі секретної інформації, а з відкритим ключем дешифрування — для різних протоколів ідентифікації абонентів і підтвердження аутотентичності повідомлень.

Асиметрична криптосистема визначається трьома алгоритмами. Це

а) алгоритм $K$ генерації ключів (на вхід подається випадкове число або набір символів $r$, на виході одержуємо пару $(k_1, k_2) = K(r)$, яка складається з ключа шифрування $k_1$ і ключа дешифрування $k_2$). Один з цих ключів розголошується і є відкритим, інший є секретним і тримається в таємниці;
б) алгоритм шифрування $E_{k_1}$ і
в) алгоритм дешифрування $D_{k_2}$.

Очевидно, що для будь–якого відкритого тексту $m$ має виконуватися рівність $D_{k_2}(E_{k_1}(m)) = m$. Усі три алгоритми є відкритими і загальнодоступними.

Недоліком асиметричних систем є набагато менша, ніж у симетричних, швидкість шифрування–дешифрування. Тому у випадках великих об'ємів конфіденційної інформації вони часто використовуються лише для створення чи шифрування перед передачею відкритим каналом зв'язку секретних ключів, які потім використовуються в симетричних системах.

## Основна частина

Основні задачі, для розв'язання яких використовуються криптосистеми з відкритим ключем:

1. Забезпечення *конфіденційності* інформації при її пересиланні.
2. Підтвердження *цілісності* або *аутентичності* (тобто що повідомлення не було підмінене чи сфальшоване під час пересилання).
3. *Ідентифікація* (підтвердження, що повідомлення було вислане саме вказаним адресатом).
4. Запобігання можливості відмови учасника інформаційного обміну від переданого повідомлення.
5. Обмін секретними ключами (зокрема, для користування криптосистемами із секретними ключами).
6. Жеребкування на відстані.

7. Поділ таємниці (наприклад, стартового коду балістичної ракети).
8. Доведення без розголошення.
9. Створення систем гасел для контролю повноважень при доступі до даних.

Ці завдання розв'язуються за допомогою спеціальних *протоколів*, тобто процедур, які вказують послідовність і тип повідомлень, якими обмінюються між собою адресати.

Основною метою дослідження є побудова учбових протоколів криптосистем з відкритим ключем (асиметричних систем) та програмна реалізація цих протоколів криптосистем засобами середовища програмування Lazarus. Основною задачею була побудова бібліотеки обробки цілих чисел великої розмірності, побудова арифметичних процедур обробки цілих чисел, в записі яких використовується досить значна кількість цифр.

В якості учбових протоколів криптосистем з відкритим ключем для їх вивчення були вибрані:

а) протокол обміну ключами Діффі-Геллмана;
б) протокол RSA;
в) протокол Ель-Гамаля;
г) протокол Рабіна.

Опишемо більш досконало тільки *RSA* – протокол.

Спеціалісти-комп'ютерщики із Массачусетського технологічного інституту Рональд Л. Райвест, Аді Шамір і Леонард Адлеман (R.L.Rivest, A.Shamir, L.Adleman) розробили метод, який дозволяє реалізувати систему Діффі – Геллмена на основі використання простих чисел – метод отримання цифрових підписів і комерційних криптографічних систем – *RSA* шифр.

Розглянемо основні кроки протоколу **криптосистеми *RSA***.

*Алгоритм генерування ключів.* Вибирають два досить великі прості числа $p$ і $q$. Тоді для $n = p \cdot q$ значення функції Ойлера дорівнює $\varphi(n) = (p-1) \cdot (q-1) = n - p - q + 1$. Далі випадковим чином вибирають число $e$, яке не перевищує $\varphi(n)$ і взаємно просте з ним. Для $e$, використовуючи алгоритм Евкліда, знаходять елемент $d$ такий, що $d < \varphi(n)$ і $ed \equiv 1 (\mathrm{mod}\ \varphi(n))$, тобто елемент $d$ – обернений до елемента $e$ в мультиплікативній групі $\mathbb{Z}^*_{\varphi(n)}$.

*Відкритим ключем* є числа $n$ та $e$.

*Таємним ключем* є $d$, $p$, $q$.

*Алгоритм шифрування.* Шифрування відбувається блоками так, щоб кожен блок позначав число, яке не перевищує $n$. Алгоритм шифрування $E$ по-

лягає у піднесенні $M$ до степеня $e$ за модулем $n$: $C = E(M) = M^e \pmod n$.

*Алгоритм дешифрування* $D$ полягає у піднесенні $C$ до степеня $d$ за модулем $n$, тобто $M = D(C) = C^d \pmod n$.

Для ілюстрації свого методу Райвест, Шамір і Адлеман зашифрували деяку фразу, для чого спочатку перетворили її у цифрову форму $x$, замінюючи букву a англійського алфавіту на 01, букву b – на 02 і т.д., букву z – на 26, пропуск між словами – на 00, а потім зашифрували описаним алгоритмом з

$$n = 114381625757888867669325779976146612010218296721242362562561842 9$$

$$3570693524573389783059712356395870505898907514759929002687954354 1$$

і $e = 9007$, причому було відомо, що прості числа $p$ і $q$ були, відповідно, 64 і 65-цифровими. Першому, хто дешифрує криптотекст

$$C = 9686961375462206147714092225435588290575999112457431987469512 09$$

$$3081629822514570835693147662288398962801339199055182994515781515 4$$

було обіцяно винагороду в 100 доларів США. Тільки через 17 років у 1994 р. Аткінс, Графф Ленстра і Лейленд (D.Atkins, M.Graff, A.K.Lenstra, P.C.Leyland) дешифрували цю фразу (числа $p$ та $q$ виявилися таким:

$$p = 3490529510847650949147849619903898133417764638493387843990820577,$$

$$q = 3276913299326670954996198819083446141317764296799294253979828853 3,$$

а вихідне повідомлення було досить безглуздою фразою «the magic words are squeamish ossifrade». Дешифрування зайняло 220 днів і задіяно було близько 1600 комп'ютерів, об'єднаних сіткою Internet.

Проаналізуємо питання стійкості системи *RSA*. У процесі зламання *RSA* шифру криптоаналітику необхідно розв'язувати наступну задачу: відомо числа $e$, $n$, $C$, де $n$ є добутком двох невідомих простих чисел $p$ та $q$, і НСД$(e, \varphi(n)) = 1$. Потрібно знайти таке число $x$, що $x^e \equiv C \pmod n$. На даний момент не існує ефективного алгоритму для розв'язання цієї задачі, але і не доведено, що такого алгоритму не існує. Цю задачу можна звести до наступної: за відомими $e$, $n$, де $n = pq$ ($p$ та $q$ невідомі) і НСД$(e, \varphi(n)) = 1$, знайти таке $d$, що для всіх цілих $x$ виконується $x^{ed} \equiv x \pmod n$, яка, у свою чергу, призводиться до обчислювання $\varphi(n)$. Останню задачу можна розв'язувати або використовуючи факторизацію модуля $n = pq$, або спробувати обійтися без розкладу на прості множники.

Задача розкладу натурального числа на прості множники (*задача факторизації*) еквівалентна задачі пошуку хоча б одного простого дільника числа $n$

і такого ефективного алгоритму також так і не знайдено. Звичайно можна перебирати всі прості множники числа до $\sqrt{n}$. Але, наприклад, для числа, яке записується 100 десятковими цифрами, знайдеться не менше $4 \cdot 10^{42}$ простих чисел, що не перевищують $\sqrt{n}$. Комп'ютеру, що виконує мільйон ділень за секунду (при навіть дуже грубій оцінці) потрібно буде не менше, ніж $10^{35}$ років. Відомі і більш ефективні способи розкладу цілих чисел на множники, але і вони працюють дуже повільно. Так, на сьогодні найефективніші алгоритми факторизації потребують часу $\exp c\sqrt{\ln n \ln\ln n}$. На межі сучасних можливостей є факторизація чисел із 150 десятковими цифрами. Серед останніх досягнень у цій області можна згадати про успіх Ленстра та Монассі, які розклали 155-значне число на три простих числа. Для цього вони використали 1000 об'єднаних ЕОМ та шість тижнів їх машинного часу. Обчислення здійснювалися за допомогою алгоритму англійського математика Дж. Полларда. Факторизація чисел із більшою кількістю цифр – задача майбутнього. Тому числа $p$ та $q$ вибирають:

*1)* достатньо великими (не менше 100 десяткових знаків), не дуже близькими одне до одного, але, у той же час, щоб вони не дуже відрізнялися одне від одного;

*2)* $p$ і $q$ повинні бути такими, щоб найбільший спільний дільник чисел $p-1$ і $q-1$ був невеликим, наприклад, 2;

*3)* $p$ та $q$ повинні бути сильно простими числами (натуральне число називається сильно простим, якщо число, яке більше за нього на одиницю, має великий простий дільник, число, яке менше за нього на одиницю, також має великий простий дільник, причому, якщо від цього останнього простого дільника відняти одиницю, то отримаємо число, яке також має великий простий дільник.

Якщо хоча б одна із умов не зберігається, то існують досить ефективні алгоритми факторизації числа $n$.

Спроби обійтись без факторизації призводять до наступної задачі: за відомими $e$, $n$, де $n = pq$ ($p$ та $q$ невідомі) і $\text{НСД}\,(e, \varphi(n)) = 1$, знайти таке число $d$, що $ed-1$ ділиться на $\psi(n)$, де $\psi(n) = \text{НСК}\,(p-1, q-1)$. Еквівалентність цієї і вихідної задач випливає із того, що якщо $n = pq$, де $p$, $q$ – різні прості числа, то $x^t \equiv x \pmod{n}$ для всіх цілих $x$ тоді і тільки тоді, коли $t \equiv 1 \pmod{\psi(n)}$. При знаходженні такого $d$, що $ed-1$ ділиться на $\psi(n)$, число $m = ed-1$ використовується для факторизації $n$ за допомогою так званої ймовірнісної процедури. І знову приходимо до такої ж складної обрахункової задачі, як і факторизація числа.

## Висновки

Досліджено алгоритми побудови учбових протоколів криптосистем з відкритим ключем, також досліджено криптостійкість таких протоколів. Досліджено проблеми реалізації алгоритмів обробки чисел зі значною кількістю цифр в записі.

## Література

1. *Алферов А.П.* Основи криптографії / А.П. Алферов, А.Ю.Зубов, А.С.Кузьмин, А.В. Черемушкин. — М.: Гелиос АРВ, 2002. — 480 с.
2. *Виноградов И.М.* Основы теории чисел / И.М. Виноградов. — М.: Наука, 1981. — 176 с.
3. *Ахо А.* Построение и анализ вычислительных алгоритмов / А. Ахо, Дж. Хопкорофт, Дж. Ульман. — М.: Мир, 1979. — 536 с.
4. *Василенко О.Н.* Теоретико-числовые алгоритмы в криптографии / О.Н. Василенко. — М.: МЦМО, 2003. — С. 43–48.
5. *ElGamal T.* A public key cryptosystem and a signature scheme based on discrete logaritms / T. ElGamal // IEEE Transactions on Information Theory. — Vol. 31, 1985. — P. 469–472.
6. *Merkle M.H.* Hidding information and signatures in trapdoor knapsacks / M.H. Merkle, M.E. Hellman // IEEE Transactions on Information Theory. — Vol. 24, 1978. — P. 525–530.