

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ СІКОРСЬКОГО»

ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ

Лабораторна робота 2

Реалізація алгоритмів генерації ключів гібридних криптосистем

Підгрупа 2А

Виконали:

Галіца О.О.

Паршин О.Ю.

Литвиненко Ю.С.

ФІ-22мн

Перевірила:

Байденко П.В.

1 Теоретичні відомості

1.1 Критерії згоди та їх основні характеристики

Статистичною гіпотезою називається будь-яке твердження про вид чи властивості розподілу випадкових величин, що спостерігаються в експерименті. Звичайно її називають основною чи нульовою гіпотезою і позначають символом H_0 . Якщо формулюється тільки одна гіпотеза, то правило, згідно з яким перевіряється, погоджуються наявні статистичні дані з цією гіпотезою, чи вони її спростовують, називаються *критерієм згоди*. Якщо гіпотеза H_0 однозначно фіксує розподіл спостережень, то її називають *простою*, у протилежному випадку — *складною*. Для побудови критерію шукають статистику $T(x)$ (тобто функцію від випадкових результатів спостережень), що характеризує відхилення емпіричних даних від гіпотетичних значень у випадку справедливості гіпотези H_0 .

Припустимо, що така статистика і її розподіл при гіпотезі H_0 знайдені. Нехай Ω — множина, якій належать можливі значення величин, що спостерігаються в експерименті, і визначимо множину $\tau = \{t \mid t = T(x), x \in \Omega\}$ — множина усіх можливих значень статистики T ; визначимо для фіксованого заздалегідь достатньо малого числа $\alpha > 0$ підмножину $\tau_{1\alpha} \subset \tau$ так, щоб імовірність здійснення події $\{T(x) \in \tau_{1\alpha}\}$ у випадку справедливості гіпотези H_0 задовольняла наступній умові $P\{T(x) \in \tau_{1\alpha} | H_0\} \leq \alpha$.

Правило перевірки гіпотези H_0 можна сформулювати в такий спосіб. Нехай x — реалізація випадкової величини X , що спостерігалася, $t = T(x)$ — відповідне значення статистики T . Якщо виявиться, що $t \in \tau_{1\alpha}$, то в припущенні справедливості гіпотези H_0 відбулася малоімовірна подія і ця гіпотеза повинна бути відкинута як така, що суперечить статистичним даним. У протилежному випадку немає підстав відмовлятися від прийнятої гіпотези і варто вважати, що спостереження не суперечать гіпотезі (чи погоджуються з нею).

В описаній методиці статистику T називають *статистикою критерію*, а підмножину її значень $\tau_{1\alpha}$ — *критичною областю* для гіпотези H_0 . Число α називають *рівнем значимості* критерію, і його можна вважати імовірністю помилкового відкидання гіпотези H_0 , коли вона вірна. У конкретних задачах її величину вибирають зазвичай рівною 0.1; 0.05; 0.01 і т.д.

Кожен критерій будується для того, щоб визначити, чи мають місце ті чи інші відхилення від основної гіпотези. Характер таких відхилень може бути різним, тому треба мати критерії як універсального типу (такі, що „уловлюють”, будь-які відхилення від основної гіпотези), так і призначені для виявлення відхилень тільки певного типу.

Будь-який розподіл $F_X = F$ спостереженої випадкової величини X , що може виявитися правдивим (тобто допустимим в даній ситуації), але такий, що відрізняється від гіпотетичного (тобто розподілу при основній гіпотезі H_0), називають *альтернативним розподілом* чи *альтернативною*. Сукупність всіх альтернативних розподілів називають *альтернативною гіпотезою* і позначають символом H_1 .

1.2 Критерії і тести для перевірки якості випадкових двійкових послідовностей

В цьому розділі спочатку наведено три критерії, що перевіряють статистичні властивості псевдовипадкових послідовностей: рівноімовірність знаків, незалежність сусідніх знаків, однорідність послідовності. Всі ці критерії є критеріями хі-квадрат Пірсона для перевірки відповідним чином сформульованих гіпотез. Після цього описано ще два критерії, які НІСТ вважає обов'язковими під час перевірки якості згенерованої послідовності: спектральний тест (з використанням дискретного перетворення Фур'є) та універсальний статистичний тест Мауера. Кожен з цих тестів було реалізовано, результати можна переглянути у розділі 2.

Розглянемо послідовність $\{Y_j\}$, $j = 1 \dots m$, де кожна Y_j є випадковою величиною, що приймає набір значень із алфавіту A . Всі величини Y_j мають однаковий розподіл та розглядаються як вихідні значення деякого генератору.

1.2.1 Критерій перевірки рівноімовірності знаків

Послідовність Y_j задовольняє умові рівноімовірності знаків, якщо кожна Y_j розподілена рівноімовірно на A . Таким чином, кожне значення із A повинно зустрічатись у довільній реалізації даної послідовності однакову кількість разів.

Тест на виконання умови рівноімовірності не відрізняється великою чутливістю, однак він доволі швидкий. В практичних задачах його рекомендується застосовувати в першу чергу, оскільки якщо послідовність не пройде цей тест, то немає рації застосовувати до неї інші. Тепер розглянемо власне алгоритм за яким потрібно діяти.

Тест

1. Розглядається байтова послідовність $\{Y_j\}$, $j = 1 \dots m$. Сформулюємо гіпотезу H_0 , що полягає в тому, що всі байти послідовності рівноімовірні.
2. За значеннями послідовності $\{Y_j\}$, $j = 1 \dots m$ обчислюється статистика

$$\chi^2 = \sum_{j=0}^{255} \frac{(\nu_j - n_j)^2}{n_j}$$

де ν_j — число байтів j , що спостерігається у послідовності, n_j — очікуване число байтів j у послідовності за умови, що вірна гіпотеза H_0 , тобто в даному випадку $n_j = \frac{m}{256}$ для всіх $j = 1 \dots 255$

3. Обчислюється граничне значення $\chi^2_{1-\alpha}$, що відповідає рівню значимості α при $l = 255$, за формулою: $\chi^2_{1-\alpha} = \sqrt{2l}Z_{1-\alpha} + l$, де $Z_{1-\alpha}$ — квантиль стандартного нормального розподілу.
4. Якщо $\chi^2 \leq \chi^2_{1-\alpha}$, то гіпотеза H_0 не суперечить експериментальним даним, в іншому випадку гіпотеза H_0 відкидається.

1.2.2 Критерій перевірки незалежності знаків

Послідовність Y_j задовольняє умові *незалежності знаків*, якщо імовірність прийняти деяке значення для Y_j не залежить від того, які значення прийняли Y_1, \dots, Y_{j-1} . Однак перевірка такої умови зазвичай вкрай важка, тому часто розглядають більш послаблені вимоги — наприклад, значення Y_j не повинно залежати від значення Y_{j-1} (незалежність від попереднього знаку).

Тест

1. Байти послідовності $\{Y_j\}$, $j = 1 \dots m$ розглядаються парами (Y_{2i-1}, Y_{2i}) , $i = \overline{1, n}$, де $n = \lfloor \frac{m}{2} \rfloor$. Сформулюємо гіпотезу H_0 , що полягає в тому, що всі байти послідовності незалежні від попереднього значення.
2. За значеннями послідовності $\{Y_j\}$, $j = 1 \dots m$ обчислюється статистика за умови, що вірна гіпотеза H_0 про незалежність байтів:

$$\chi^2 = n \left(\sum_{i,j=0}^{255} \frac{\nu_{ij}^2}{\nu_i \alpha_j} - 1 \right)$$

де ν_{ij} — кількість появи пари (i, j) , $n = \sum_{i,j=0}^{255} \nu_{ij}$, $\nu_i = \sum_{j=0}^{255} \nu_{ij}$ — кількість появи байта i на першому місці в парі, $\alpha_j = \sum_{i=0}^{255} \nu_{ij}$ — кількість появи байта j на другому місці в парі.

3. Обчислюється граничне значення $\chi_{1-\alpha}^2$, що відповідає рівню значимості α при $l = 255^2$, за формулою: $\chi_{1-\alpha}^2 = \sqrt{2l} Z_{1-\alpha} + l$, де $Z_{1-\alpha}$ — квантиль стандартного нормального розподілу.
4. Якщо $\chi^2 \leq \chi_{1-\alpha}^2$, то гіпотеза H_0 не суперечить експериментальним даним, в іншому випадку гіпотеза H_0 відкидається.

1.2.3 Критерій перевірки однорідності двійкової послідовності

Послідовність $\{Y_j\}$ задовольняє умові однорідності, якщо для довільної реалізації вибіркового розподілу, одержаний на всій послідовності, буде співпадати із вибірковою розподілом, одержаним на довільній її підпослідовності достатньої довжини; іншими словами, на довільному фрагменті послідовності веде себе однаково. Зауважимо, що для виконання умови однорідності не важливо, який саме розподіл будуть мати Y_j . Зокрема, цей розподіл не обов'язково повинен бути рівномірним.

Перевірка умови однорідності в повному обсязі також доволі складна, тому на практиці перевіряють більш слабкі форми даної умови — наприклад, розбивають послідовність на окремі інтервали та перевіряють, чи співпадають вибіркові розподіли на цих інтервалах.

Тест

1. Послідовність $\{Y_j\}$, $j = 1 \dots m$ розбивається на r відрізків довжиною $m' = \lfloor \frac{m}{r} \rfloor$, де m — загальне число байтів, $n = r \cdot m'$ — загальне число байтів, що використовуються. Формулюється гіпотеза H_0 про вибір байтів з того самого розподілу.

2. За значеннями послідовності $\{Y_j\}$, $j = 1 \dots m$ обчислюється статистика за умови, що вірна гіпотеза H_0 про вибір з одного розподілу:

$$\chi^2 = n \left(\sum_{i=0}^{255} \sum_{j=0}^{r-1} \frac{\nu_{ij}^2}{\nu_i \alpha_j} - 1 \right)$$

де ν_{ij} — кількість появи байта i у відрізку j , $n = \sum_{i,j=0}^{255} \nu_{ij}$, $\nu_i = \sum_{j=0}^{r-1} \nu_{ij}$, $\alpha_j = \sum_{i=0}^{255} \nu_{ij} = m'$ (зауважимо, що α_j дорівнює довжині j -того відрізка; в нашому тесті всі відрізки мають однакову довжину, але це не обов'язково).

3. Обчислюється граничне значення $\chi_{1-\alpha}^2$, що відповідає рівню значимості α при $l = 255 \cdot (r-1)$, за формулою: $\chi_{1-\alpha}^2 = \sqrt{2l} Z_{1-\alpha} + l$, де $Z_{1-\alpha}$ — квантиль стандартного нормального розподілу.
4. Якщо $\chi^2 \leq \chi_{1-\alpha}^2$, то гіпотеза H_0 не суперечить експериментальним даним, в іншому випадку гіпотеза H_0 відкидається.

1.2.4 Спектральний тест

Цей тест зосереджується на висотах піків у дискретному перетворенні Фур'є послідовності $\{Y_j\}$, $j = 1 \dots m$. Метою цього тесту є виявлення періодичних особливостей у досліджуваній послідовності, які б вказували на відхилення від припущення про випадковість. Мета полягає в тому, щоб виявити, чи кількість піків, що перевищують поріг у 95 %, є значною, чи ні.

Тест

1. Послідовність $\{Y_j\}$, $j = 1 \dots m$ перетворюється на послідовність $\{X_j\}$, $j = 1 \dots m$ з $X_j = \pm 1$ наступним чином: $X_j = 2Y_j - 1$. Фактично отримуємо наступне правило: $0 \rightarrow -1$, $1 \rightarrow 1$.
2. Застосовуємо дискретне перетворення Фур'є до послідовності $\{X_j\}$:

$$S_k = \sum_{n=0}^{m-1} X_n e^{-\frac{2\pi i}{m} kn}$$

В результаті чого отримали послідовність комплексних змінних.

3. Обчислити $M = \text{modulus}(S')$, де S' — підрядок, що складається з перших $n/2$ елементів у S , а функція *modulus* створює послідовність висот піків.
4. Обчислити $T = \sqrt{\left(\log \frac{1}{0.05}\right) n}$ — порогове значення. Згідно з припущенням про випадковість, 95% значень, отриманих у результаті тесту, не повинні перевищувати T .
5. Обчислити $N_0 = 0.95n/2$. N_0 — очікувана теоретична кількість піків (за припущенням випадковості), менших за T .
6. Обчислити N_1 — фактична спостережувана кількість піків, менших за T .

7. Обчислити $d = \frac{(N_1 - N_0)}{\sqrt{n \cdot 0.95 \cdot 0.05/4}}$
8. Обчислити $P - value = erfc\left(\frac{|d|}{\sqrt{2}}\right)$, де $erfc(z) = \frac{2}{\sqrt{\pi}} \int_z^{\infty} e^{-t^2} dt$ — функція помилки (фактично це спеціальна інтегральна сімгоїда). Якщо $P - value \geq \alpha$, то гіпотеза H_0 не суперечить експериментальним даним, в іншому випадку гіпотеза H_0 відкидається.

1.2.5 Універсальний тест Мауера

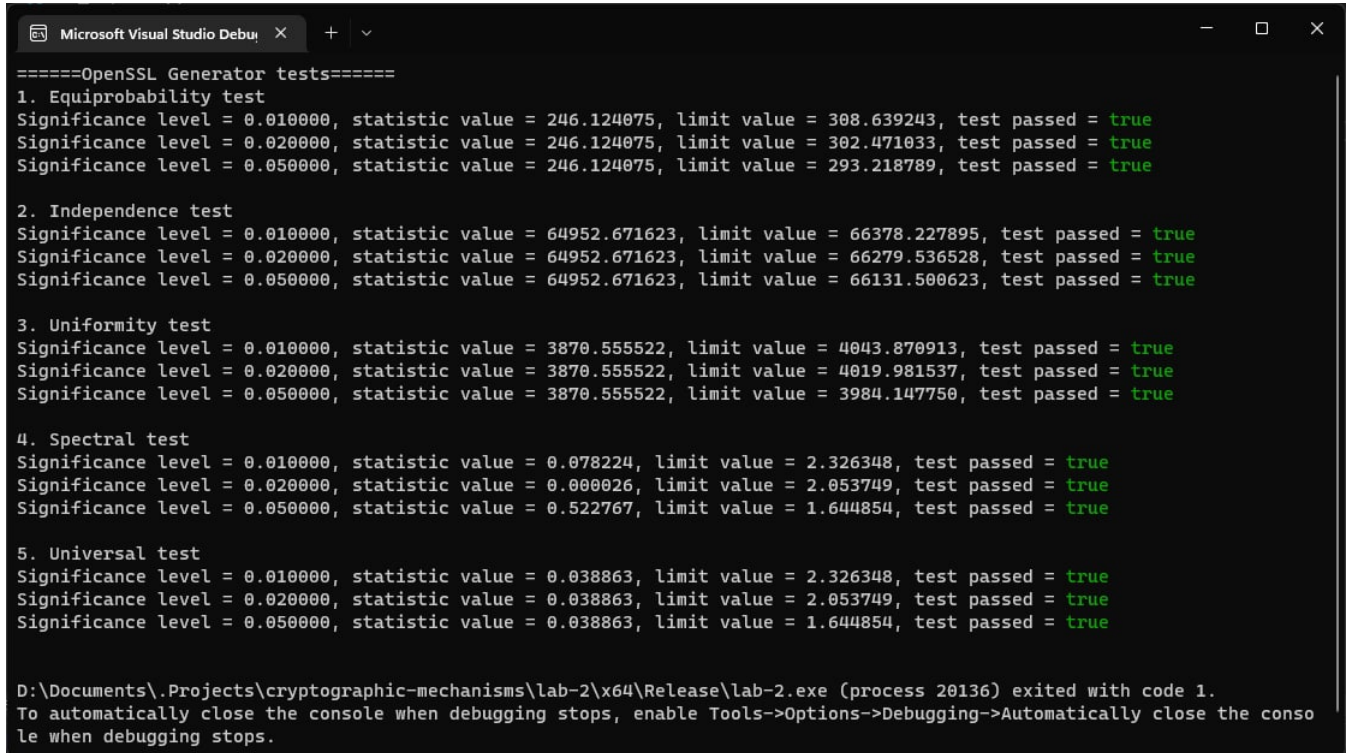
Мета тесту — визначити, чи можна послідовність значно стиснути без втрати інформації. Якщо послідовність можна значно стиснути, то вона вважається не випадковою.

Тест

1. Розбити m -бітну послідовність $\{Y_j\}$ на дві частини: послідовність ініціалізації складається з перших Q L -бітних блоків, які не перетинаються; тестова послідовність складається з наступних K L -бітних блоків, які не перетинаються. Біти, які залишились в кінці та не ввійшли у блоки, відкидаються.
 2. Використовується послідовність ініціалізації та за нею складається таблиця, де індексами є всі можливі L -бітові значення (позначимо їх T_j), а в комірках стоїть останній номер входження відповідного L -бітового блоку у послідовність.
 3. Розглядають тестову послідовність та для кожного L -бітового блоку виконуються наступні кроки:
 - (а) Обчислити відстань від поточного L -бітового блоку з номером i , до останнього такого блоку, який зустрівся раніше. Позначимо результат $x = i - T_j$.
 - (б) Замінити відповідне значення останнього входження цього блоку у таблиці на поточний номер блоку ($T_j = i$).
 - (в) Обчислити значення $sum = sum + \log_2(x)$.
 4. Обчислити статистику $f_m = \frac{1}{K} \sum_{i=Q+1}^{Q+K} \log_2(i - T_j) = \frac{sum}{K}$
 5. Обчислити $P - value = erfc\left(\left|\frac{f_n - expectedValue(L)}{\sqrt{2}\sigma}\right|\right)$. Якщо $P - value \geq \alpha$, то гіпотеза H_0 не суперечить експериментальним даним, в іншому випадку гіпотеза H_0 відкидається.
- **expectedValue* — це теоретичні оцінки значень заданої статистики, таблицю можна глянути у книзі “Handbook of Applied Cryptography”.

2 Результат виконання програми

Для усереднення результатів, тест проводився 10 разів (тобто було згенеровано 10 випадкових послідовностей і для кожної з них були отримані деякі результати, які потім усереднювалися). На зображенні можна побачити рівні значущості, значення статистик та критичні значення для кожного з тестів.



```
Microsoft Visual Studio Debug Console
=====OpenSSL Generator tests=====
1. Equiprobability test
Significance level = 0.010000, statistic value = 246.124075, limit value = 308.639243, test passed = true
Significance level = 0.020000, statistic value = 246.124075, limit value = 302.471033, test passed = true
Significance level = 0.050000, statistic value = 246.124075, limit value = 293.218789, test passed = true

2. Independence test
Significance level = 0.010000, statistic value = 64952.671623, limit value = 66378.227895, test passed = true
Significance level = 0.020000, statistic value = 64952.671623, limit value = 66279.536528, test passed = true
Significance level = 0.050000, statistic value = 64952.671623, limit value = 66131.500623, test passed = true

3. Uniformity test
Significance level = 0.010000, statistic value = 3870.555522, limit value = 4043.870913, test passed = true
Significance level = 0.020000, statistic value = 3870.555522, limit value = 4019.981537, test passed = true
Significance level = 0.050000, statistic value = 3870.555522, limit value = 3984.147750, test passed = true

4. Spectral test
Significance level = 0.010000, statistic value = 0.078224, limit value = 2.326348, test passed = true
Significance level = 0.020000, statistic value = 0.000026, limit value = 2.053749, test passed = true
Significance level = 0.050000, statistic value = 0.522767, limit value = 1.644854, test passed = true

5. Universal test
Significance level = 0.010000, statistic value = 0.038863, limit value = 2.326348, test passed = true
Significance level = 0.020000, statistic value = 0.038863, limit value = 2.053749, test passed = true
Significance level = 0.050000, statistic value = 0.038863, limit value = 1.644854, test passed = true

D:\Documents\Projects\cryptographic-mechanisms\lab-2\x64\Release\lab-2.exe (process 20136) exited with code 1.
To automatically close the console when debugging stops, enable Tools->Options->Debugging->Automatically close the console when debugging stops.
```