УДК 621. 391.519.2:519.7

Ю.І. ГОРБЕНКО, канд. техн. наук

МОДЕЛІ ОЦІНКИ ЙМОВІРНОСТЕЙ ПОВТОРНОГО ВИКОРИСТАННЯ КЛЮЧА ПРИ СИМЕТРИЧНИХ ПЕРЕТВОРЕННЯХ

Вступ

Безумовним є той факт, що ключі та ключова інформація повинні генеруватись випадково, рівно ймовірно та незалежно[1 – 3]. При виконанні вказаних вимог можна сподіватись, що при їх застосуванні для симетричних криптографічних перетворень може надіятись певна, задекларована, криптографічна стійкість. Вище вжито термін «надіятись» у тому сенсі, що криптографічна стійкість також залежить від багатьох факторів – алгоритму шифрування, параметрів криптографічного перетворення, в першу чергу довжини початкового ключа та довжини блока, при блоковому криптографічному перетворенні, чи періоду повторення при потоковому криптографічному перетворенні тощо. Як правило для генерування ряду типів ключів використовують генератори випадкових послідовностей (ГВП) та генератори псевдовипадкових послідовностей(ГПВП). В основному відомі ГПВП будуються на основі використання блокових симетричних шифрів, функцій гешування та перетворень у групі точок еліптичних кривих [1]. При цьому, досягаємий рівень криптографічної стійкості суттєво визначається реальною ентропією ключа (початкового ключа). Генерація або введення ентропії в ГПВП з використанням небезпечного методу може призвести до порушень вимог гарантій безпеки. Більш детальними вимогами до ключів є забезпечення вимог по критеріям необоротності, непередбачуваності та нерозрізнюваності [1]. У зв’язку з наведеним значення ключа має генеруватися й оброблятися за умов виконання таких вимог.

1. Склад (структура) початкового значення – початкове значення має включати вхідні дані ентропії і дані персоналізації.

2. Використання початкового значення – ГПВП можуть використовуватися для генерації як секретної, так і відкритої інформації. У всіх випадках початкове значення має тримати-ся в секреті.

3. Ентропія початкового значення – вхідні дані ентропії для початкового значення мають містити достатню ентропію для рівня безпеки, що вимагається.

4. Розмір початкового значення – мінімальний розмір початкового значення, що залежить від обраного ГПВП та стійкості захисту, яка потрібна для застосування.

5. Конфіденційність початкового значення – початкові значення мають бути оброблені способом, що є сумісним із рівнем безпеки, який вимагається для цільових даних. Наприклад, якщо ДГВБ використовується для формування ключів у криптографічній системі, то початкові значення, що використовуються для генерації ключів, мають розглядатися нібито вони є ключами.

6. Період використання початкового значення – початкове значення повинне мати вказаний обмежений термін дії та періодично оновлюватися.

7. І одна з проблемних вимог – ключі у кожного користувача, а інколи і у всіх користувачів мережі, не повинні повторюватись.

Мета статті – розробка математичних моделей оцінки можливосте появи одного і того ж ключа і одного користувача при застосуванні його для блочних криптографічних перетворень(БКП) та розробка відповідних оцінок, обмежень і рекомендацій.

1. Обґрунтування вихідних даних та критерій оцінки криптографічної стійкості від повторного використання ключа

Нехай ключ застосовується для криптографічного захисту інформації з використанням БКП. Нехай також кожен ключ генерується у відповідності з викладеними у вступі вимога-

ми. При цьому нехай також забезпечуються необхідні властивості кожного ключа відносно вимог до необоротності, непередбачуваності (вперед та назад) та нерозрізнюваності [1]. Тобто, в загальному випадку можна говорити, що кожен ключ дійсно генерується випадково, рівно ймовірно, незалежно та одно рідно. Нехай також довжини блоків та ключів приймають такі значення:

$$128/128,\ 128/256,\ 128/512,\ 256/256,\ 256/512.\ 512/512 \qquad (1)$$

Проведений аналіз існуючих, зрозуміло доступних джерел, не дозволив знайти метод вирішення задачі оцінки криптографічної стійкості при здійсненні повторного використання ключа. В подальшому під ключем будемо розуміти початковий ключ, який вводиться в засіб криптографічного захисту інформації (КЗІ) і визначає початкову ентропію його стану. Якраз на його основі потім розгортаються циклові (раундові) ключі, що використовуються, наприклад, при ітеративних блокових перетвореннях. Тому важливою є задача розробки математичні моделі та отримати оцінки ймовірностей повторного використання одного і того ж ключа. На наш погляд, вказана задача перше за все може бути вирішеною на основі використання математичного апарату «узагальненого парадоксу» про день народження, точніше на основі методів Полларда, що на ньому ґрунтуються [4 – 5].

Розгляд та вирішення задачі проведемо з урахування таких обмежень:

1) генерування кожного ключа здійснюється на двох різних станціях і вони постачають відповідному користувачеві ключі;

2) генерування кожного ключа здійснюється на одній і тій же станції, і вона постачає відповідному користувачеві ключі.

За даних умов у відповідності з метою необхідно оцінити криптографічну стійкість кожного з засобу КЗІ від повторного використання ключа, вибравши в якості основного показника ймовірність $P_п$ появи у користувача двох однакових ключів, тобто

$$Kj = Ki, \qquad (2)$$

для випадкових довільних значень j та i. А в якості критерію прийняття рішення про перевагу приймемо критерій мінімуму ймовірності повторного використання ключа в засобі КЗІ, тобто

$$\min(Pп) \qquad (3)$$

При цьому кращою будемо вважати систему, яка забезпечує мінімальне значення (3).

2. Модель λ Полларда оцінки криптографічної стійкості від повторного використання ключа

Для цієї моделі будемо вважати, що генерування кожного ключа здійснюється на двох різних генераторах і вони постачають відповідному користувачеві ключі.

Обґрунтування методу λ-Полларда проведемо в такий спосіб [4 – 5]. Розглядаються два процеси $\rho_1(Z_i)$ і $\rho_2(Z_j)$. Причому вважається, що ключі співпали, якщо

$$\rho_1(Z_i) = \rho_2(Z_j). \qquad (4)$$

Подія співпадання може здійсниться з імовірністю $P_i = \frac{1}{n_g}$, де n_g ймовірність появи g – ключа із повної множини. Тоді ймовірність події, що

$$\rho_1(Z_i) \neq \rho_2(Z_j)$$

можна оцінити як

$$R\left(\rho_1(Z_i) \neq \rho_2(Z_j)\right) = 1 - \frac{1}{n_g}. \qquad (5)$$

Якщо розглядати k таких послідовних подій, то ймовірність того, що $\rho_2(Z_1), \rho_2(Z_2),..., \rho_2(Z_k)$ не будуть збігатися з $\rho_1(Z_i)$, можна обчислити як

$$R(\rho_1(Z_i) \neq \rho_2) = \left(1 - \frac{1}{n_G}\right)^k.$$

Далі, ймовірність того, що не відбулося жодного збігу для ρ_1 і ρ_2 (при всіх $Z_1, Z_2,..., Z_k$ для ρ_1 і ρ_2):

$$R\left(\rho_1(Z_i) \underset{\forall i,j}{\neq} \rho_2(Z_j)\right) = \left(\left(1 - \frac{1}{n_G}\right)^k\right)^k = \left(1 - \frac{1}{n_G}\right)^{k^2} \quad (6)$$

Тоді, імовірність того, що хоча б одне значення $\rho_1(Z_i)$ збігається із $\rho_2(Z_j)$ для всіх значень k:

$$R(\rho_1(Z_i) = \rho_2(Z_j)) = 1 - \left(1 - \frac{1}{n_G}\right)^{k^2} \quad (7)$$

Таким чином, співвідношення (7) у загальному випадку визначає ймовірність співпадання двох ключів, які генеровані на двох різних станціях.

Подамо (7) у вигляді, зручному для обчислень при параметрах $P_k = R(\rho_1(Z_i) = \rho_2(Z_j))$, n_g і k. З урахуванням того, що $n_g \geq 2^{128}$, з великою точністю справедливо, що [4, 5]

$$1 - \frac{1}{n_g} = e^{-\frac{1}{n_{gt}}} \quad (8)$$

$$P_k = 1 - e^{-\frac{1}{n_G}k^2} = 1 - e^{-\frac{k^2}{n_G}} \quad (9)$$

Формула (9) – наближена в тому розумінні, що в повну групу подій включено k^2 подій, що колізія не відбувається. А потім у (9), по суті, додається ще одна подія, коли відбувається колізія $(\rho_1(Z_i) = \rho_2(Z_j))$. Таким чином, подій більше, ніж може відбутися, тобто $k^2 + 1$.

Отримаємо уточнену формулу, аналогічну (7). Нехай можливі $k(k-1)$ подій, коли колізія не відбувається, й одна подія, коли відбувається одна колізія. Тоді уточнена за цієї умови формула (9) має вигляд:

$$P(\rho_1(Z_i) = \rho_2(Z_j)) = 1 - \left(1 - \frac{1}{n_G}\right)^{(k^2-1)} \quad (10)$$

Далі, з урахуванням (8), маємо:

$$P_k = 1 - e^{-\frac{1}{n_G}(k^2-1)} = 1 - e^{-\frac{k^2-1}{n_G}} \quad (11)$$

Формула (10) найбільш точна, тому надалі будемо розглядати саме її. Позначимо для подальшого, що число спроб(кроків) $I = k$, то після ряду перетворень одержимо, аналогічно трьох параметричне рівняння, яке пов'язує між собою три параметри – складність, число ключів, що можуть використовуватись в БСШ, а також ймовірність співпадання двох ключів:

$$I^2 - 1 + n_g \ln(1 - P_k) = 0. \quad (12)$$

Далі, так як $I^2 >> 1$, то (12) можна подати у вигляді

$$I^2 = -n_g \ln(1 - P_k)$$

або

$$I_\lambda = \sqrt{-n_g \ln(1 - P_k)}.$$

Також (11) в великою точністю можна подати у вигляді

$$P_k = 1 - e^{-\frac{1}{n_G}(I^2)} = 1 - e^{-\frac{I^2}{n_G}} \quad (13)$$

Таким чином, (11) – (13) зв'язують три параметри моделі обчислення повторного використання ключа – ймовірність спів падання двох ключів P_k, число можливих ключів n_g та число ключів I, які можуть бути генеровані для користувача(одного і того ж засобу КЗІ).

Відмітимо, що при $P_k = 10^{-9}$ маємо

$$I_{0,99} \approx \sqrt{-n \ln 10^{-9}} = 5.5\sqrt{n} \quad (14)$$

Формулу (14) можна також подати для $P_k = 10^{-r}$ у вигляді

$$I_{1-10^{-r}} \approx \sqrt{-n \ln 10^{-r}} \quad (15)$$

В табл. 1 наведені оцінки ймовірностей появи двох однакових ключів P_k згідно (13) у залежності від числа генерованих ключів I, що генеровані та поставлені користувачеві, та допустимого числа ключів, які можуть бути використані в системі, $n = 2^{lk}$, де lk – довжина ключа(128, 256, 512 бітів).

Таблиця 1

Ймовірності появи двох однакових ключів P_k згідно (13), I=k

№ п/п	l_k	I	$P_{mu} = 1 - \exp(-I * I/n_g)$
1	128	1.000000e+009 1.000000e+012 1.000000e+015 1.000000e+018 1.000000e+024	0 1.4432899320127035e-015 1.4693679606381238e-009 0.0014682889460021498 1
2	256	1.000000e+024 1.000000e+032 1.000000e+036 1.000000e+048	0 4.3187675657918589e-014 4.3180749546012365e-006 1
3	512	1.000000e+032 1.000000e+064	0 0

Аналіз даних табл. 1 показує, що зі збільшенням довжини блоку $l_б$ БСП та відповідно довжини ключа l_k показує, що зі збільшенням довжин блоків та ключів число ключів, які можуть бути генеровані для користувача, збільшується і при $l_k = l_б = 512$ обмеження практично відсутні, тобто колізії практично неможливі.

3. Модель ρ Полларда оцінки криптографічної стійкості повторного використання ключа

В цій моделі будемо вважати, що ключі генеруються та поставляються користувачеві з однієї станції. При розробці моделі будемо до уваги брати [1, 4, 5].

При розгляді даної моделі будемо вважати, що поява значення $\rho(Z_\nu)$ відбувається випадково та рівно ймовірно. Як і вище визначимо ймовірність $P(n,k)$ того, що серед зна-

чень $\rho(\mathrm{Z}_\nu)$ принаймні два співпадуть. Також позначимо через $R(n,k)$ імовірність того, що при k обчисленнях значень збігу не буде, тобто умова $\rho(Z_i)=\rho(Z_j)$ не виконується. Оскільки, $P(n,k)$ і $R(n,k)$ складають повну групу подій, то

$$P(n,k)+R(n,k)=1$$

та

$$P(n,k)=1-R(n,k). \tag{16}$$

Знайдемо $R(n,k)$. Для цього знайдемо число способів N_1 обчислення $\rho(\mathrm{Z}_\nu)$, коли при k експериментах збігу не буде. Обчисливши $\rho(\mathrm{Z}_1)$ маємо n значень без повторення, при $\rho(\mathrm{Z}_2)$ – $(n-1)$ значень,..., при $\rho(\mathrm{Z}_k)$ – $(n-(k-1))$ значень. Тому

$$N_1=n(n-1)(n-2)...(n-(k-1)).$$

Далі, з незалежною появою $\rho(\mathrm{Z}_\nu)$ в при k обчисленнях маємо $N_з$ – число подій, причому

$$N_з=n^k.$$

Значить

$$R(n,k)=\frac{N_1}{N_з}=\frac{n(n-1)(n-2)...(n-(k-1))}{n^k}, \tag{17}$$

Далі підставивши (17) у (16), маємо:

$$P(n,k)=1-\frac{n(n-1)(n-2)...(n-(k-1))}{n^k}. \tag{18}$$

Необхідно відзначити, що формула (18) є точною й дозволяє обчислити ймовірність колізії за відомих n та k. Але важливим є завдання розв'язання параметричного рівняння (P,n,k), коли один або два параметри змінні. Якщо необхідно обчислити $P(n,k)$ при змінюваних (n,k), враховуючи те, що n та k, як правило, не досить великі, при обчисленнях краще використати (18) у вигляді [4]:

$$\begin{aligned} P(n,k)&=1-\frac{n}{n}\cdot\left(\frac{n-1}{n}\right)\left(\frac{n-2}{n}\right)...\left(\frac{n-(k-1)}{n}\right)= \\ &=1-\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)...\left(1-\frac{k-1}{n}\right). \end{aligned} \tag{19}$$

Так як реальні значення k великі, то обчислення (19) потребує значних ресурсів. Також нас цікавить знаходження параметра k як параметра оцінки складності криптографічного аналізу методом «повного розкриття», тобто знаходження, наприклад, особистого ключа d_X.

У нашому випадку $k<<n$, тому $x\equiv\frac{k}{n}<<1$. Для цієї умови можна скористатися тим, що:

$$(1-x)\approx e^{-x}, \tag{20}$$

але, звичайно, оцінюючи похибку при заданих значеннях (n,k).

Підставивши в (19) $\left(1-\frac{i}{n}\right)=e^{-\frac{i}{n}}$, маємо

$$P(n,k)=1-e^{-\frac{1}{n}}\cdot e^{-\frac{2}{n}}\cdot...\cdot e^{-\frac{k-1}{n}}. \tag{21}$$

Вирази в показниках степеня – це члени арифметичної прогресії, перший член $\frac{1}{n}$, а число членів дорівнює $k-1$, тому [4]:

$$P(n,k)=1-e^{-\left(\frac{1}{n}+\frac{k-1}{n}\right)\frac{k-1}{2}}=1-e^{-\frac{k(k-1)}{2n}}. \quad (22)$$

Таким чином, вираз (22) пов'язує між собою три основні параметри – імовірність колізії $P(n,k)=P_k$, складність криптоаналізу методом ρ-Полларда $k=I$ та розмір простору значення базової точки – з порядком базової точки *n*.

Надалі будемо подавати (22) у вигляді

$$1-P_k=e^{-\frac{k(k-1)}{2n}}, \text{ або } \ln(1-P_k)=-\frac{k(k-1)}{2n}.$$

Після простих перетворень одержимо

$$I^2-I+2n\ln(1-P_k)=0. \quad (23)$$

Рівняння (23) дозволяє отримати наближене значення, тобто є неточним. Похибка пов'язана з використанням наближення $(1-x)\approx e^{-x}$, але вона при наших вихідних даних, коли k < < n $_g$, буде достатньо малою [1, 5].

Отримані вище співвідношення дозволяють зробити такі висновки та рекомендації.

1. Точне значення $P(n,k)$ можна отримати використовуючи (19).

2. Спів відношення (23) зв'язує три основні параметри, змінюючи одні з них можна отримати значення іншого або інших двох, а враховуючи (19) отримуємо достатньо точні значення .

3. Враховуючи те, що $I^2>>I$, можна скористатися наближенням:

$$I^2\approx -2n\ln(1-P_k) \text{ або } I_\rho\approx\sqrt{-2n\ln(1-P_k)}. \quad (24)$$

В табл. 2 наведені оцінки ймовірностей появи двох однакових ключів P_k згідно (22) у залежності від числа генерованих ключів I, що генеровані та поставлені користувачеві, та допустимого числа ключів, які можуть бути використані в системі, $n=2^{lk}$, де *lk* – довжина ключа (128, 256, 512 бітів).

Таблиця 2

Оцінки ймовірностей появи двох однакових ключів P_k згідно (22)
у залежності від числа генерованих ключів, *I=k*

№ п/п	l_k	I	Pk = 1 – exp(-I * (I-1)/(2n$_g$))
1	128	1.000000e+009 1.000000e+012 1.000000e+015 1.000000e+018 1.000000e+024	0 7.7715611723760958e-016 7.346839803190619e-010 0.00073441415507669028 1
2	256	1.000000e+024 1.000000e+032 1.000000e+036 1.000000e+048	0 2.1538326677728037e-014 2.159039808047325e-006 1
3	512	1.000000e+032 1.000000e+064	0 0

Наведені в табл. 2 дані, отримані з використанням ρ Полларда моделі, дозволяють зробити висновки, аналогічні висновкам, що наведені в табл. 1 для моделі λ Полларда. Також

отримані з використанням моделі λ Полларда дані близькі по значенням до даних, отриманих з використанням моделі ρ Полларда.

Висновки та рекомендації

1. Початкове значення ключа, його розмір та ентропія (тобто випадковість) мають вибиратися таким чином, щоб мінімізувати ймовірність колізії послідовностей, що вироблена на основі різних початкових значень, та зменшити ймовірність вгадування початкового значення. Довжина початкового значення повинна, як мінімум, дорівнювати числу бітів для зазначеного рівня захисту, але для збільшення гарантії достатньої ентропії й усунення будь-яких проблем можливого багатократного використання початкового числа довжина початкового числа має бути більше мінімального.

2. Для забезпечення непередбачуваності необхідно приділяти особливу увагу при отриманні та обробці початкового значення ключа. Початкове значення має генеруватися випадково, рівно ймовірно та не залежно.

3. На основі математичного апарату узагальненого парадоксу про день народження запропоновано дві моделі оцінки стійкості проти повторної появи у користувача одного і того ж ключа, тобто співпадання ключа у одного користувача. Як основні запропоновані моделі, коли ключ генерується двома різними та одним генератором, відповідно співвідношення (13) та (22). Хороше співпадання оцінок для розглянутих моделей дозволяє поставляти ключі як одним так і двома різними системами.

4. Наведені в табл. 1 та 2 оцінки ймовірностей появи двох однакових ключів P_k згідно (13) та (22) у залежності від числа генерованих ключів I, що генеровані та поставлені користувачеві, та допустимого числа ключів, які можуть бути використані в системі, n =2^{lk}, де lk – довжина ключа(128, 256, 512 бітів), дозволяють обгрунтувати та вибрати параметри ключової системи в частині довжини ключа, строку дії ключа та умов їх застосування, при чому обидві моделі дають практично одинакові оцінки.

5. При довжині блоку 128 бітів реальне обмеження виникає тоді, коли число ключів, що генеруються двома генераторами, дорівнює 1.000000e+012, хоча при цьому ймовірність співпадання двох ключів не перевищує 10^{-15} (7.7715611723760958e-016). При цій же довжині блоку та застосуванні одного генератора, ймовірність співпадання можна оцінити як 7.7715611723760958e-016.

6. При довжині блоку 256 бітів реальне обмеження виникає тоді, коли число ключів, що генеруються двома генераторами, дорівнює 1.000000e+032, хоча при цьому ймовірність співпадання двох ключів не перевищує 10^{-14} (4.3187675657918589e-014). При цій же довжині блоку та застосуванні одного генератора ймовірність співпадання можна оцінити як 2.1538326677728037e-014.

7. При блоковому симетричному шифруванні з довжиною блока 512 бітів, як слідують дані табл. 1 та 2, практичних обмежень із – за співпадання ключів немає. Крім того, при збільшенні довжини блоку та відповідному збільшенні довжини ключа, стійкість БСШ «Калина» проти можливостей появи одного і того ж ключа суттєво збільшується. В цьому є також перевага БСШ, у якого довжина ключа та довжина блоку може приймати значення 512 бітів.

Список літератури: 1. *Горбенко І.Д., Горбенко Ю.І.* Прикладна криптологія. – Харків : ХНУРЕ ; Форт, 2012. – 868 с. 2. *Долгов В.И., Лисицкая И.В.* Блочные симметричные шифры. Методология оценки стойкости к атакам дифференциального и линейного криптоанализа. – Харьков : ХНУРЭ ; Форт. – 455 с. 3. *Daemen J.* Annex to AES Proposal Rijndael. http://www.nist.gov/aes. 4. *Тевяшев А.Д., Горбенко Ю.И.* Оценка опасности криптоаналитических атак методом создания коллизий // Радиотехника. – 2002. – Вып. 126. – С. 166–172. 5. *Горбенко Ю.І.* Математична модель перекриття шифру для потокових криптосистем // Радиотехника. – 2005. – Вып.141. – С. 40-54.

Харківський національний
університет імені В.Н.Каразіна

Надійшла до редколегії 14.03.2015