2 Забазпечення комп'ютерної безпеки в державних, банківських та інших інформаційних системах

УДК 681.3.06:519.248.681

ІМОВІРНІСНИЙ АЛГЕБРАЇЧНИЙ КРИПТОАНАЛІЗ ШИФРАТОРА «SFINKS» З ПЕВНИМ КЛАСОМ ФІЛЬТРУЮЧИХ ФУНКЦІЙ

Сергій Пометун
ФТІ НТУУ „КПІ"

Анотація: **Експериментально знайдено ключ потокового шифратора «SFINKS» з послабленою фільтруючою функцією за допомогою імовірнісної алгебраїчної атаки. Клас таких вразливих функцій досить широкий і містить багато функцій, стійких проти відомих неалгебраїчних методів криптоаналізу.**
Summary: **Stream cipher «SFINKS» with weakened filtering function is considered. Practical cryptanalysis was done by means of probabilistic algebraic attack. There are a number of such vulnerable functions and some of them are resistant against known non-algebraic methods of cryptanalysis.**
Ключові слова: **Потоковий шифратор, SFINKS, криптоаналіз, імовірнісна алгебраїчна атака.**

І Вступ

В даній роботі описується імовірнісна алгебраїчна атака на послаблений шифратор «SFINKS». Потоковий шифратор «SFINKS» [1], збудований на основі регістру зсуву з лінійним зворотнім зв'язком (РЗЛЗЗ), був запропонований у 2005 році на конкурс потокових шифраторів ESTREAM, що був об'явлений в рамках проекту ECRYPT1. Алгебраїчні атаки на потокові шифратори, побудовані на основі РЗЛЗЗ, були вперше запропоновані французьким криптографом Ніколасом Куртуа у 2003 році [2].

В роботі [2] пропонується два типи сценаріїв алгебраїчної атаки – детермінований та імовірнісний. Протягом наступних декількох років більш проста та явно ефективна детермінована алгебраїчна атака (або просто алгебраїчна атака) не раз удосконалювалася, модифікувалася, та була успішно застосована до ряду потокових шифраторів [3 – 7].

Імовірнісна алгебраїчна атака досліджувалася в [8], де були дані критерії стійкості ускладнюючої (фільтруючої) функції шифратора до такої атаки.

Неформально, алгебраїчна атака зводить задачу знаходження ключа до розв'язання системи рівнянь досить низького степеня над полем $GF(2)$, де невідомими є власне біти ключа. Імовірнісна алгебраїчна атака зводить цю ж задачу до розв'язання системи рівнянь ще нижчого степеня, але рівняння системи будуть істинними лише з деякою імовірністю $1-\varepsilon$. Основним показником вразливості ускладнюючої функції $f:GF(2)^k \to GF(2)$ до алгебраїчної атаки є її алгебраїчний імунітет $AI(f)$ [9], який показує найменший степінь системи рівнянь, якого можна досягти за допомогою цієї атаки. Вразливість ускладнюючої функції до імовірнісної атаки визначається розширеним алгебраїчним імунітетом $AI(f,\varepsilon)$ [8], який показує найменший степінь системи рівнянь, якого можна досягти за допомогою імовірнісної алгебраїчної атаки, при умові, що рівняння отриманої системи будуть істинними з імовірністю не менше, ніж $1-\varepsilon$.

Отже, в даній роботі описується імовірнісна алгебраїчна атака на послаблений потоковий шифратор «SFINKS», у якого оригінальна фільтруюча функція f з алгебраїчним імунітетом $AI(f)=6$ замінена на послаблену f', у якої $AI(f')=3$, та $AI(f',\varepsilon)=2$ для деякого дуже маленького ε.

Показано, що в такому випадку імовірнісна алгебраїчна атака має меншу обчислювальну складність, ніж детермінована. Була написана програма для знаходження ключа методом імовірнісної алгебраїчної атаки та успішно здійснено криптоаналіз.

Варто відмітити, що клас таких „послаблених" функцій досить широкий, і він містить багато

[1] ECRYPT (European Network of Excellence for Cryptology) – 4-х річний Європейський криптологічний науково-дослідний проект, розпочатий 1 лютого 2004 р. Конкурс потокових шифраторів ESTREAM – див. http://www.ecrypt.eu.org/stream/

рівноімовірних функцій з дуже хорошими іншими криптографічними властивостями (всіма, за виключенням алгебраїчного імунітету). Тому вразлива до імовірнісних алгебраїчних атак функція цілком могла бути використана в якомусь з шифраторів, розроблених до появи алгебраїчних атак (тобто до 2003 року).

II Імовірнісна алгебраїчна атака

Сформулюємо строго задачу аналітичного криптоаналізу типового потокового шифратора, побудованого на основі РЗЛЗЗ.

Дана система рівнянь над полем Галуа $GF(2)$

$$\begin{cases} f(P\bar{x}) = b_0 \\ f(PL\bar{x}) = b_1 \\ f(PL^2\bar{x}) = b_2 \\ \ldots\ldots\ldots\ldots\ldots \\ f(PL^{N-1}\bar{x}) = b_{N-1} \end{cases}, \qquad (1)$$

де: $b_i \in GF(2)$ – біти гами, $L: GF(2)^n \to GF(2)^n$ – лінійний оператор (описує функціонування РЗЛЗЗ), $P: GF(2)^n \to GF(2)^k$ – проекційний оператор (з n біт вибирає k аргументів $\bar{y} = (y_1, y_2, ..., y_k)$ для функції f), $f: GF(2)^k \to GF(2)$ – ускладнююча функція шифратора (часто подається у вигляді поліному від аргументів – поліному Жегалкіна), $\bar{x} = (x_1, x_2, ..., x_n) \in GF(2)^n$ - ключ.

Потрібно знайти $\bar{x} = (x_1, x_2, ..., x_n)$ – невідомі n біт ключа. При цьому всі елементи шифратора (f, L, P) та гама b_i вважається відомою, також вважається, що доступний достатньо великий обсяг гами N.

Вся проблема – у високій обчислювальній складності розв'язання системи (1).

Ідея алгебраїчної атаки – знизити степінь рівнянь системи (1), шляхом домноження їх на спеціально підібрану функцію $g(x)$ [2]

$$\begin{cases} f(P\bar{x}) = b_0 \\ f(PL\bar{x}) = b_1 \\ f(PL^2\bar{x}) = b_2 \\ \ldots\ldots\ldots\ldots\ldots \\ f(PL^{N-1}\bar{x}) = b_{N-1} \end{cases} \Rightarrow \begin{cases} g(P\bar{x}) \cdot f(P\bar{x}) = b_0 \cdot g(P\bar{x}) \\ g(PL\bar{x}) \cdot f(PL\bar{x}) = b_1 \cdot g(PL\bar{x}) \\ g(PL^2\bar{x}) \cdot f(PL^2\bar{x}) = b_2 \cdot g(PL^2\bar{x}) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ g(PL^{N-1}\bar{x}) \cdot f(PL^{N-1}\bar{x}) = b_{N-1} \cdot g(PL^{N-1}\bar{x}) \end{cases}, \qquad (2)$$

Для ілюстрації наведемо простий приклад. Якщо $f(\bar{y}) = f(y_1, y_2, y_3) = y_1y_2y_3 + y_2 + 1$, то поклавши $g(\bar{y}) = y_2 + 1$, маємо $f(\bar{y})g(\bar{y}) = (y_1y_2y_3 + y_2 + 1)(y_2 + 1) = y_2 + 1$. Степінь знижено з 3 до 1 (тут і далі додавання виконується за модулем 2). Ясно, що зі зменшенням степеня зменшується і обчислювальна складність розв'язання системи (зайвих коренів не виникає, бо для існуючих ефективних методів розв'язання потрібно $N > n$ рівнянь).

Згідно з [10, 8] імплікація (2) еквівалентна такій імплікації

$$\begin{cases} f(P\bar{x}) = b_0 \\ f(PL\bar{x}) = b_1 \\ f(PL^2\bar{x}) = b_2 \\ \ldots\ldots\ldots\ldots \\ f(PL^{N-1}\bar{x}) = b_{N-1} \end{cases} \Rightarrow \begin{cases} h_{b_0}(P\bar{x}) = b_0 \\ h_{b_1}(PL\bar{x}) = b_1 \\ h_{b_2}(PL^2\bar{x}) = b_2 \\ \ldots\ldots\ldots\ldots \\ h_{b_{N-1}}(PL^{N-1}\bar{x}) = b_{N-1} \end{cases},$$

де для функцій $h_0(\bar{y})$ та $h_1(\bar{y})$ виконується $f(\bar{y}) = 0 \Rightarrow h_0(\bar{y}) = 0$ та $f(\bar{y}) = 1 \Rightarrow h_1(\bar{y}) = 1$ відповідно (повертаючись до нашого прикладу, можна показати, що $y_1y_2y_3 + y_2 + 1 = 0 \Rightarrow y_2 + 1 = 0$). З системи (3) для розв'язання можуть вибиратися лише ті рівняння, де $b_i = 0$ або $b_i = 1$, залежно від того, степінь якої з функцій, $h_0(\bar{y})$ чи $h_1(\bar{y})$, нижчий.

Тоді ідею імовірнісної алгебраїчної атаки можна записати так [8]

$$\begin{cases} f(P\bar{x}) = b_0 \\ f(PL\bar{x}) = b_1 \\ f(PL^2\bar{x}) = b_2 \\ \ldots\ldots\ldots\ldots \\ f(PL^{N-1}\bar{x}) = b_{N-1} \end{cases} \xrightarrow{\Pr} \begin{cases} h'_{b_0}(P\bar{x}) = b_0 \\ h'_{b_1}(PL\bar{x}) = b_1 \\ h'_{b_2}(PL^2\bar{x}) = b_2 \\ \ldots\ldots\ldots\ldots \\ h'_{b_{N-1}}(PL^{N-1}\bar{x}) = b_{N-1} \end{cases}$$

де для функцій $h'_0(\bar{y})$ та $h'_1(\bar{y})$ виконується $\Pr(h'_0 = 0 \mid f = 0) = 1 - \varepsilon_0$ та $\Pr(h'_1 = 1 \mid f = 1) = 1 - \varepsilon_1$. Знову ж, з системи (4) можуть вибиратися лише ті рівняння, де $b_i = 0$ або $b_i = 1$, залежно від степенів $h'_0(\bar{y})$ чи $h'_1(\bar{y})$ та відповідних значень ε_0 та ε_1.

Значення мінімально можливого степеня функцій $h_0(\bar{y}), h_1(\bar{y}), h'_0(\bar{y}), h'_1(\bar{y})$ визначаються алгебраїчним імунітетом функції $f(\bar{y})$, а саме $\min\{\deg(h_0), \deg(h_1)\} = AI(f)$ та $\min\{\deg(h'_0), \deg(h'_1)\} = AI(f, \varepsilon)$ (тривіальні константні функції не враховуються). Ясно, що $AI(f) = AI(f, 0)$ і $AI(f, \varepsilon)$ спадає зі збільшенням ε. Отже, імовірнісна алгебраїчна атака дозволяє ще сильніше знизити степінь рівнянь системи (4) порівняно з (3), але ціною того, що ці рівняння будуть виконуватись хоча й з досить великою, але вже не одиничною імовірністю.

Найпростіший спосіб розв'язання системи (3) – лінеаризація (кожен терм високого степеня заміняється новою змінною, в результаті отримується лінійна система, але від значно більшої кількості змінних). Система (4) – система зі спотвореними правими частинами. В загальному випадку такі системи розв'язувати значно складніше. Простий спосіб для лінійних систем – випадковий вибір необхідної для розв'язання кількості рівнянь і спроба їх розв'язання, і так до тих пір, поки всі, в черговий раз вибрані рівняння, не виявляться істинними. І цей спосіб вважається ефективним! Ясно, що при скільки-небудь великих спотвореннях система (4) практично нерозв'язна. Втім, можливо, що для нелінійних систем будуть знайдені більш ефективні алгоритми. В нашому конкретному випадку для шифратора «SFINKS» послаблена функція підібрана так, що спотворення дуже малі, і вибрати необхідну для розв'язання кількість істинних рівнянь вдається практично з першого разу.

III Атака на послаблений «SFINKS»

Функціонування шифратора «SFINKS» [1] (без непотрібного для нашого розгляду механізму аутентифікації) представлено на рис. 1.

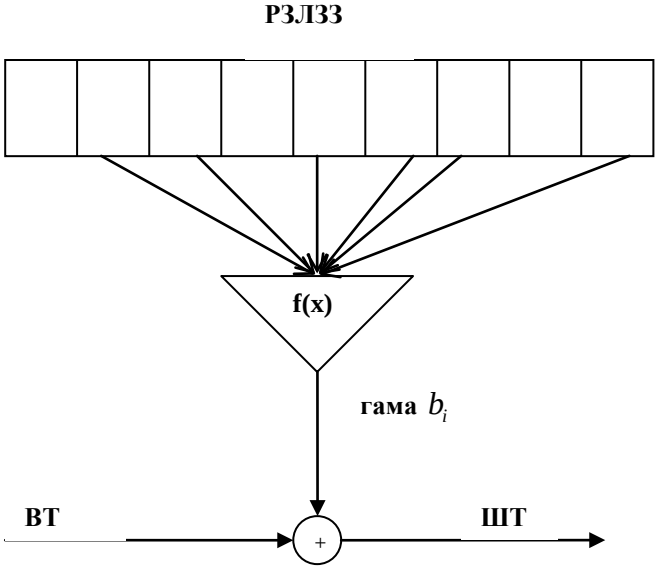

Рисунок 1 – Функціонування шифратора «SFINKS»

Де РЗЛЗЗ – регістр зсуву з лінійним зворотним зв'язком довжини $n = 256$ біт (опускаючи технічні подробиці його початкове заповнення $\overline{x} = (x_1, x_2, ..., x_n) \in GF(2)^{256}$ і вважаємо ключем, стан регістра на кожному наступному кроці визначається як дія відомого лінійного оператора $L : GF(2)^{256} \to GF(2)^{256}$ на попередній стан, оператор L підібрано так, що регістр має період $2^{256} - 1$, тобто проходить усі стани, окрім нуля); $f : GF(2)^{17} \to GF(2)$ - рівноімовірна фільтруюча функція від 17-ти аргументів (за аргументи вибираються певні 17 комірок РЗЛЗЗ, рівноімовірність означає, що функція рівно на половині аргументів дорівнює нулю, а на іншій половині - одиниці); b_i – вихід (значення) функції f на i-му кроці. ВТ – відкритий текст, що підлягає шифруванню; ШТ – шифрований текст.

Розглядається атака на ключ з відомим відкритим текстом. Тобто криптоаналітику відомі ВТ, ШТ і потрібно знайти ключ $\overline{x} \in GF(2)^{256}$ (в разі успіху, знаючи подальший ШТ та $\overline{x}$, можна було б знаходити подальший, вже невідомий ВТ).

Можна бачити, що така задача знаходження ключа шифратора «SFINKS» якраз і представляється системою рівнянь (1).

Для спрощення атаки ми замінили оригінальну фільтруючу функцію з алгебраїчним імунітетом $AI(f) = 6$ на значно слабшу f', таку що $AI(f') = 3$, та $AI(f', \frac{1}{2^{16}}) = 2$ (успішний варіант алгебраїчної атаки на повноцінний шифратор описано в [11], складність атаки становить 2^{70} операцій).

Таблиця істинності $f'(\overline{y})$ будувалася так.

Кладемо $f'_0(\overline{y}) = y_1 y_3$

В точці $\overline{y} = (y_1, y_2, y_3, ..., y_{17}) = (1,0,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0)$ заміняємо одиницю на нуль

В решті точок випадково ставимо одиниці, поки $f'(\overline{y})$ не стане рівноімовірною

В результаті виконання такої процедури була отримана певна функція $f'(\overline{y})$ 16 степеня.

Таким чином, відповідно до (4), $f'(\overline{y})$ вибрана так, що існує функція другого степеня $h'_0(\overline{y}) = y_1 y_3$, для якої виконується $f'(\overline{y}) = 0 \xrightarrow{\Pr} h'_0(\overline{y}) = 0$ з імовірністю $\frac{1}{2^{16}}$ (імплікація буде хибною лише для аргументу $\overline{y} = (1,0,1,1,1,0,0,...,0)$).

Для описаного вище шифратора аргументами y_1 та y_3 функції $f'(\bar{y})$ є перша та сьома комірка РЗЛЗЗ відповідно (на нульовому кроці це x_1 та x_7, далі – лінійні комбінації компонент $\bar{x}$). Отже, можна стверджувати, що якщо вихід функції $f'(\bar{y})$ на i-му кроці дорівнює нулю, то з великою імовірністю добуток значень першої та сьомої комірки регістра теж дорівнює нулю. Повертаючись до системи (4), виберемо з неї лише такі рівняння з номерами i_k, $k = 1, .., r$, для яких $b_{i_k} = 0$, в результаті маємо систему

$$\begin{cases} f(PL^{i_1}\bar{x}) = 0 \\ f(PL^{i_2}\bar{x}) = 0 \\ \ldots\ldots\ldots\ldots \\ f(PL^{i_r}\bar{x}) = 0 \end{cases} \xrightarrow{\Pr} \begin{cases} h'_0(PL^{i_1}\bar{x}) = 0 \\ h'_0(PL^{i_2}\bar{x}) = 0 \\ \ldots\ldots\ldots\ldots \\ h'_0(PL^{i_r}\bar{x}) = 0 \end{cases}, \qquad (5)$$

де $h'_0(\bar{y}) = y_1 y_3$ та імовірність істинності рівнянь правої частини дорівнює $1 - \frac{1}{2^{16}}$.

Таким чином, права частина системи (5) – це система рівнянь другого степеня від 256 змінних, кожне з яких може виявитись хибним з імовірністю $\frac{1}{2^{16}}$. Розв'язуючи її методом лінеаризації отримаємо систему лінійних рівнянь від $C_{256}^0 + C_{256}^1 + C_{256}^2 = 32897$ невідомих. Отже, $r = 32897$ істинних рівнянь буде досить для її розв'язання. Оскільки r приблизно вдвічі менше, ніж 2^{16}, то таку кількість істинних рівнянь вдасться вибрати навмання практично з першого разу. Оскільки лише приблизно кожне друге рівняння (4) має праву частину b_i=0, то всього для розв'язання знадобиться біля $N = 2r \approx 2^{16}$ рівнянь, відповідно потрібно 2^{16} біт=8 Кбайт відомої гами. Складність розв'язання лінійної системи визначається як куб від кількості невідомих, що становить біля $r^3 \approx 2^{45}$ бітових операцій.

Ця атака була успішно реалізована на персональному комп'ютері.

IV Експериментальні подробиці

Було написано дві програми – перша реалізує шифратор «SFINKS» (без механізму аутентифікації) з послабленою фільтруючою функцією. За допомогою неї отримувалася гама. Друга програма здійснювала пошук ключа, маючи як вхідні дані лише відрізок гами.

Обчислювання виконувались на процесорі Celeron 1.7 GHz з 640 Mb оперативної пам'яті. Побудова 50 тис. рівнянь потребувала біля 10 хвилин. роботи, час розв'язання системи з 32897 рівнянь становив 40 хвилин. При цьому об'єм матриці лінійних рівнянь становив $32895 \times 32897 \approx 1Gbit = 128Mb$. Програма використовувала до 1100 Mb оперативної пам'яті. Програми оптимізовані за швидкістю та пам'яттю.

Перші ж 33000 вибраних рівнянь виявилися істинними. При перевірці, серед 250 тис. рівнянь було 4 хибних, тобто 1 з 62500, що близько до теоретично передбаченої імовірності 1 з $\frac{1}{2^{16}}$. Зайвих коренів, як і передбачалося, не виявилося. Також примітно, що 32902 рівнянь виявилося досить для побудови матриці рангу 32895 (з позосталих чотирьох варіантів розв'язків для такої матриці легко вибирався правильний). Це важливо, бо $\Pr(h'_0(\bar{y}) = 0) = 3/4 > 1/2$, тобто кожне рівняння дає менше одного біту інформації, і був розрахунок, що може знадобитися рази в півтора більше рівнянь. Те ж, що знадобилася практично мінімальна необхідна кількість рівнянь, показує, що проблема нерівноімовірності функції $h'_0(\bar{y})$ (більш докладно див. [8]) усувається завдяки лінеаризації системи.

Розглянемо застосування до функції $f'(\bar{y})$ звичайної детермінованої алгебраїчної атаки, тобто заміни $f'(\bar{y}) = 0 \Rightarrow h(\bar{y}) = 0$, де рівняння виду $h(\bar{y}) = 0$ – завжди істинні, але вже третього степеня. В такому випадку, використовуючи метод лінеаризації, ми отримали б $C_{256}^1 + C_{256}^2 + C_{256}^3 \approx 1.3 \cdot 2^{21}$ рівнянь, що дає значно більшу складність в $C = 2^{64}$ бітових операцій.

V Висновки

Вперше застосована на практиці імовірнісна алгебраїчна атака. Об'єкт атаки - потоковий шифратор «SFINKS» з послабленою фільтруючою функцією. Були написані дві програми – для реалізація шифратора та для його криптоаналізу. Програма криптоаналізу успішно знаходить ключ. Цим доведено, що імовірнісна алгебраїчна атака може бути втілена на практиці.

Оскільки функції, вразливі до імовірнісних алгебраїчних атак, можуть мати дуже хороші всі інші криптографічні властивості, то потокові шифратори, які не розроблялися стійкими проти алгебраїчних атак (зокрема розроблені до 2003), можуть виявитися вразливими до імовірнісної алгебраїчної атаки.

Література: ***1.*** *An Braeken, Joseph Lano, Nele Mentens, Bart Preneel and Ingrid Verbauwhede. Sfinks specification and source code // April 2005, Available on ECRYPT Stream Cipher Project page, http://www.ecrypt.eu.org/stream/sfinks.html* ***2.*** *Nicolas Courtois and Willi Meier. Agebraic Attacks on Stream Ciphers with Linear Feedback // Eurocrypt 2003, Warsaw, Poland, LNCS 2656, pp. 345–359, Springer. An extended version is available at http://www.minrank.org/toyolili.pdf* ***3.*** *Nicolas Courtois. Fast Algebraic Attacks on Stream Ciphers with Linear Feedback // Crypto 2003, LNCS 2729, pp: 177-194, Springer.* ***4.*** *Nicolas Courtois. Algebraic Attacks on Combiners with Memory and Several Outputs // ICISC 2004, LNCS, to appear in Springer in early 2005. Extended version available on http://eprint.iacr.org/2003/125/* ***5.*** *Frederik Armknecht, Matthias Krause. Algebraic Atacks on Combiners with Memory // Crypto 2003, LNCS 2729, pp. 162-176, Springer.* ***6.*** *Frederik Armknecht. Improving Fast Algebraic Attacks // FSE 2004, LNCS, Springer.* ***7.*** *Philip Hawkes, Gregory Rose. Rewriting Variables: the Complexity of Fast Algebraic Attacks on Stream Ciphers // in Crypto 2004, LNCS 3152, pp. 390-406, Springer, 2004. Available from http://eprint.iacr.org/2004/081/* ***8.*** *Pometun S. Generalized Correlation and Higher Order Nonlinearity for Probabilistic Algebraic Attacks Description // http://eprint.iacr.org/2007/448* ***9.*** *Meier W., Pasalic E., Carlet C. Algebraic Attacks and Decomposition of Boolean Functions // Eurocrypt 2004, LNCS 3027, pp. 474–491, Springer, 2004* ***10.*** *Пометун С. О. Алгебраїчні атаки на потокові шифратори як узагальнення кореляційних атак. Системні дослідження та інформаційні технології 2008, у друку.* ***11.*** *Nicolas T. Courtois. Cryptanalysis of Sfinks // http://eprint.iacr.org/2005/243*

УДК 681.3

КОД УМОВНИХ ЛИШКІВ І ЦІЛІСНІСТЬ ІНФОРМАЦІЙНИХ ОБ'ЄКТІВ

Вячеслав Василенко, Олександр Юдін
Національний авіаційний університет

Анотація: **Досліджені можливості застосування в задачах забезпечення цілісності інформаційних об'єктів в телекомунікаційних мережах узагальненого завадостійкого коду умовних лишків та здійснено аналіз його можливостей.**
Summary: **Explored possibilities of application in telecommunication networks in the tasks of providing of integrity of information's holding object of the generalized ant jamming code of conditional tailings. Description of such code is offered and carried out the analysis of his possibilities.**
Ключові слова: **Викривлення, завадостійкий код, комунікаційна мережа, умовні лишки, цілісність.**

I Вступ

Відповідно до термінології нормативних документів Департаменту спеціальних телекомунікаційних систем і захисту інформації Служби безпеки України [1] під цілісністю інформації розуміється її властивість, яка полягає у тому, що інформація не може бути модифікована неавторизованим користувачем або процесом. Іншими словами, під цілісністю інформації розуміється відсутність в ній будь-яких викривлень (модифікацій), які не були санкціоновані її власником, не залежно від причин або джерел виникнення таких викривлень.

Викривлення інформації, тобто порушення її цілісності, можливі на будь-якому етапі її циркуляції у обчислювальних мережах: при зберіганні, передачі або обробці. Причини таких викривлень можуть бути випадковими або навмисними. У свою чергу, випадкові викривлення можуть бути як природними, пов'язаними з дією природних чинників, так і штучними. До числа природних чинників відносяться атмосферні електромагнітні розряди, іскріння контактів в автомобілях, електротранспорті, недостатня надійність електронних елементів і елементів електричних ланцюгів, порушення реєструючого шару магнітних або оптичних носіїв і багато що інше. І випадкові і штучні мають своїм наслідком викривлення