DOI: 10.18372/2410-7840.19.11578
УДК 004.056.5

ЕФЕКТИВНІСТЬ ДОСКОНАЛО СТІЙКОЇ КРИПТОСИСТЕМИ ІЗ ЗБІЛЬШЕНОЮ ВІДСТАНЮ ЄДИНОСТІ

Євген Самойлик

Для захисту критично важливої для держави інформації доцільно використовувати досконало стійкі криптосистеми з теоретично доведеною ідеальною теоретико-інформаційною стійкістю. Проте існуючі досконало стійкі криптосистеми мають обмежену область використання, перш за все, через жорстке обмеження щодо неперевищення під час шифрування так званої відстані єдиності за ключем. Відносно невеликі значення відстані єдиності при шифруванні повідомлень, складених із символів алфавіту будь-якої із природних мов, обумовлюють необхідність частої зміни ключової інформації, що є проблемою для багатьох прикладних застосувань. Показана можливість збільшення відстані єдиності за рахунок синтезу штучної мови відображення прикладної області з алфавітом великої розмірності. Розглянуто ефективність досконало стійкої криптосистеми захисту текстової інформації, що поміщена у задану табличну форму, за умови, що ця текстова інформація береться із тезаурусу наперед визначеної прикладної області. Отримано математичні вирази та побудовано відповідні графіки, що визначають залежності відстані єдиності та ентропії ключа шифру від довжини повідомлення. Показано на кількісному рівні, що ефективність методу побудови досконало стійкої криптосистеми з укрупненим алфавітом мови відображення текстової інформації є суттєво вищою у порівнянні із ефективністю інших методів забезпечення режиму досконалої секретності.

Ключові слова: *захист текстової інформації, досконало стійка криптосистема, відстань єдиності, синтез мови відображення інформації, укрупнення алфавіту.*

Вступ. Інформаційна безпека нашої держави у певній мірі залежить від стійкості засобів захисту критично важливої державної інформації (зокрема, таємної або цілком таємної), у т.ч. і від атак з боку розвідувальних служб потужних організаційних структур. Існує певний спектр прикладних завдань, де вкрай бажано забезпечити надійний захист від порушень конфіденційності цієї інформації за умов, коли вона передається через відкриті канали зв'язку [1]. Практично стійкі засоби «сильної криптографії» не гарантують формальну, теоретично доведену, неможливість їхнього злому [2]. Що породжує обґрунтовану недовіру щодо можливостей їхнього застосування в умовах України. Логічно обрати за цих умов в якості об'єктів використання досконало стійкі криптосистеми з теоретично доведеною ідеальною теоретико-інформаційною стійкістю [3], що гарантують неможливість нелегального відновлення критично важливої державної інформації, що характеризується високими рівнями секретності. Проте ці об'єкти мають обмежену область застосування, перш за все, через необхідність дотримання основної умови ідеальної досконалості криптосистеми – не перевищення під час шифрування так званої відстані єдиності за ключем [4]. Тобто, необхідно слідкувати, щоб поточна сумарна довжина зашифрованої текстової інформації у процесі шифрування не перевищувала довжину ключа шифру.

Численні дослідження багатьох авторів вказують на відносно невеликі значення відстані єдиності при шифруванні повідомлень, складених із символів алфавіту будь-якої із природних мов [5]. Це призводить до необхідності частої зміни ключової інформації, що є проблемою для багатьох застосувань. Окрім того, необхідно забезпечувати випадковість й однакову ймовірність вибору варіантів реалізації ключа. Так що створення нових методів побудови досконало стійких криптосистем, що забезпечують більш великі значення відстані єдиності, являє актуальне завдання.

Аналіз потокових шифрів з рівно ймовірними ключами показує, що існує можливість збільшення відстані єдиності за рахунок синтезу штучної мови відображення прикладної області з алфавітом великої розмірності. У цьому випадку відкривається можливість створення більш ефективної досконало стійкої криптосистеми, яка не потребує визначення семантичних зв'язків між лінгвістичними конструкціями синтезованої мови.

У даній роботі розглядається ефективність досконало стійкої криптосистеми захисту текстової інформації, що поміщена у задану табличну форму, за умови, що ця текстова інформація береться із тезаурусу наперед визначеної прикладної області. Тобто, перш ніж здійснювати шифрування, використовуючи будь-який відомий досконало стійкий шифр, необхідно створити семантичний словник, лінгвістичні одиниці котрого у повній мірі відображають мовний простір цієї прикладної області.

Текстова інформація, що підлягає захисту, розглядається як потік текстових повідомлень (зокрема, слів або словосполучень) $M_{j,\nu}$, де $j = 1,2, \ldots, L$ – індекс, що означає порядковий номер повідомлення у потоці, $\nu = 1, 2, \ldots, N$ – індекс, що означає смисловий варіант повідомлення із наперед визначеного словника повідомлень S.

За цих умов використано **наступні позначення**.

$\boldsymbol{L}$ – довжина (обсяг) тексту, що підлягає захисту (кількість повідомлень, що передається на протязі одного сеансу зв'язку);

$\boldsymbol{S}$ – словник, що відображає певну множину $\boldsymbol{M}$ можливих у тексті повідомлень $M_{j,i}$;

M – повідомлення як елемент словника $\boldsymbol{S}$;

N – кількість смислових варіантів повідомлень (тобто, розмір словника $\boldsymbol{S}$);

$H(M)$ – ентропія повідомлення, узятого із словника $\boldsymbol{S}$;

n – довжина повідомлення (кількість символів);

r – ентропія мови, за допомогою якої відображаються повідомлення із $\boldsymbol{M}$ (тобто, це середня кількість інформації, що міститься в одній літері цієї мови і залежить від n);

R – максимальна ентропія мови (тобто, максимальне число бітів, яке може бути передано кожним символом абетки цієї мови за умови рівноймовірності виникнення усіх послідовностей символів);

D – надлишковість (інформаційна) мови ($D = R - r$);

$H(K)$ – ентропія системи захисту як розмір простору ключів шифру, що залежить від кількості можливих для використання ключів шифру;

K – кількість можливих для використання ключів шифру у системі захисту ($H(K) = \log(K)$);

U – відстань єдиності (що також називають точкою єдиності) - такий приблизний розмір шифрованого тексту, для якого сума ентропії відкритого тексту та ентропії ключа шифру дорівнює числу бітів, що міститься у цьому шифрованому тексті ($U = H(K)/D$);

E – множина можливих шифрограм, що утворюються шляхом застосування оператора перетворення T_i до відкритих повідомлень M із множини M, а індекс i відповідає конкретному ключу, що був при цьому застосований;

T_i – оператор перетворення M у E.

Постановка завдання. Припустимо, що критично важлива текстова інформація поміщена у табличну форму розміром s x L_2, що задана у вигляді таблиці, де s – кількість стовпців таблиці, а L_2 – кількість рядків цієї ж таблиці (див. табл. 1).

Таблиця 1

Задана таблична форма

	Найменування стовпця №1	Найменування стовпця №2	---	---	Найменування стовпця № s
Найменування рядка №1					
Найменування рядка №2					

Найменування рядка №L_2					

Вміст кожної клітини таблиці будемо розглядати як окреме повідомлення (слово або словосполучення) M із множини $\boldsymbol{M}$. От же, дані таблиці - це текстова інформація, яка підлягає захисту, і при передаванні через канал зв'язку розглядається як потік текстових повідомлень $M_{j,\nu}$, де $j = 1, 2, \ldots, L$ – індекс, що означає порядковий номер повідомлення у потоці, $\nu = 1, 2, \ldots, N$ – індекс, що означає смисловий варіант повідомлення із наперед визначеного словника повідомлень $\boldsymbol{S}$.

Дані таблиці необхідно передати через відкритий канал зв'язку у зашифрованому вигляді таким чином, щоб забезпечити досконало стійкий захист від порушень конфіденційності шляхом синтезу криптосистеми, що реалізує метод збільшення відстані єдиності (і, тим самим, спрощує систему управління ключами шифру).

Ураховуючи критичну важливість текстової інформації, що поміщена у таблицю, досконалість теоретико-інформаційної стійкості синтезованої криптосистеми захисту має бути теоретично доведеною. Що означає необхідність формального визначення обмежень, за яких забезпечується стовідсоткова гарантія неможливості однозначного відновлення відкритої інформації, що поміщена у таблицю, навіть за умов, коли у розпорядженні крипто аналітика знаходяться зразки зашифрованих даних скільки завгодно великої сумарної довжини, а крипто аналітик має необмежений час та

необмежені обчислювальні ресурси для дешифрування перехоплених криптограм.

Укрупнення алфавіту як спосіб збільшення відстані єдиності. Уведемо поняття *семантичний словник таблиці* як множину $\boldsymbol{M}$ усіх можливих семантичних варіантів повідомлень $M_{j,i}$, що можуть знайти своє відображення у таблиці заданої форми [6]. Розмір словника $\boldsymbol{S}$ це є розмір множини $\boldsymbol{M}$, що дорівнює N.

Будемо вважати, що на практиці у більшості випадків семантичні зв'язки між інформацією різних стовпців таблиці не спостерігаються. Тому доцільно припустити, що семантичний словник даної табличної форми складається із s семантичних підсловників відповідно до кількості стовпців у таблиці. Текстові повідомлення, що вносяться у перший стовпець, являють собою семантичні одиниці першого підсловника, у другий стовпець вносяться семантичні одиниці другого підсловника і т.д.

Синтезуємо мову відображення текстової інформації, що поміщається у задану табличну форму. Для цього визначимо в якості літери алфавіту цієї форми один рядок табл. 1. Іншими словами, уявімо, що кожен рядок табл. 1 є літерою певним чином визначеного алфавіту цієї табличної форми. У цьому разі кількість літер у такому алфавіті B визначиться за наступною формулою:

$$B = \prod_{i=1}^{s} S_i \,, \qquad (1)$$

де s – кількість підсловників у словнику табличної форми, S_i – кількість семантичних елементів (тобто, елементарних повідомлень M) в і-ому підсловнику.

Припустимо, що повідомлення, які поміщаються у клітини заданої табличної форми, є однакової довжини n. Визначимо кількість можливих смислових варіантів повідомлень N (тобто, розмір словника $\boldsymbol{S}$) довжиною n, що можуть бути внесені у клітини цієї таблиці. Розглянемо випадок, коли усі клітини таблиці є заповненими елементарними повідомленнями, узятими із підсловників. У цьому випадку будемо мати максимально можливу довжину послідовності повідомлень, що відображають інформацію у таблиці і мають бути передані на протязі одного сеансу зв'язку. Розмір цієї послідовності дорівнює кількості літер у визначеному алфавіті табличної форми. Якщо припустити, що літери із алфавіту даної табличної форми під час формування даних таблиці будуть не повторюватися, тобто кожен рядок у таблиці буде зустрічатися тільки один раз, то кількість можливих варіантів повідомлень N довжиною n розраховується за наступною формулою:

$$N = (S_1 S_2 \ldots S_s)\,(S_1 S_2 \ldots S_s - 1)\,\ldots\,(S_1 S_2 \ldots S_s - (s-1)), \qquad (2)$$

де S_i – розмір і-го підсловника, і = 1, 2, 3, …, s.

Знаючи N, використаємо ланцюг широко відомих математичних виразів щодо:

– ентропії повідомлення

$$H(M) = \log_2 N, \qquad (3)$$

що вимірюється у бітах;

– ентропії мови заданої табличної форми

$$r = H(M)/n; \qquad (4)$$

– абсолютної ентропії мови заданої табличної форми

$$R = \log_2 B; \qquad (5)$$

– надлишковості мови заданої табличної форми

$$D = R - r; \qquad (6)$$

– ентропії ключової системи

$$H(K) = \log_2 K, \qquad (7)$$

де K - кількість ключів у СЗІ.

І, як результат, визначимо відстань єдиності для розроблюваної ключової системи за формулою [4]:

$$U = H(K)/D. \qquad (8)$$

Аналізуючи результати розрахунків за наведеними вище виразами, легко побачити, що надлишковість штучної мови відображення інформації D, що поміщена у задану табличну форму, є надзвичайно малою, що, у свою чергу, згідно (8) визначає суттєво великі значення відстані єдиності U.

Приклад. Припустимо, що маємо таблицю із чотирма рядками та чотирма стовпцями. Відповідно для обраної табличної форми маємо $S_1 = 8$, $S_2 = 100$, $S_3 = 8$, $S_4 = 8$.

За формулою (1) кількість літер у алфавіті даної табличної форми визначається як $B = 8 \cdot 100 \cdot 8 \cdot 8 = 51200$. Так що розмір тезаурусу табличної форми визначається як $B = 51200$ [літер].

Розрахуємо кількість можливих повідомлень N довжиною 4 згідно з виразом (2). Для нашого прикладу

$$N = (S_1 \cdot S_2 \cdot S_3 \cdot S_4) \cdot (S_1 \cdot S_2 \cdot S_3 \cdot S_4 - 1) \cdot$$
$$\cdot (S_1 \cdot S_2 \cdot S_3 \cdot S_4 - 2) \cdot (S_1 \cdot S_2 \cdot S_3 \cdot S_4 - 3).$$

Отже, у цьому випадку $N = 6871142396067532800$.

За формулою (3) визначимо ентропію повідомлення $H(M)$:

$$\log_2 6871142396067532800 = 62.5752556894213,$$

тобто, $H(M) = 62.5752556894213$ [біт].

Визначимо ентропію мови r заданої табличної форми за формулою (4):

$$\frac{62.5752556\,894213}{4} = 15.6438139\,223553,$$

тобто, $r = 15{,}6438139223553$ [біт/літера].

Визначимо абсолютну ентропію R мови заданої табличної форми за формулою (5):

$$\log_2 51200 = 15.6438561\,897747,$$

тобто, $R = 15{,}6438561897747$ [біт/літера].

Надлишковість мови D заданої табличної форми визначимо за формулою (6):

$$15.6438561\,897747 - 15.6438139\,223553 = 0.00004226\,7419412,$$

тобто $D = 0{,}000042267419412$ [біт/літера].

Визначимо ентропію розроблюваної ключової системи $H(K)$ за формулою (7). Слід зазначити, що при розрахунку ентропії кількість ключів обираємо, виходячи із міркувань необхідності побудови системи із досконалою секретністю. Отже, нехай кількість ключів K дорівнює N – кількості можливих повідомлень довжиною n, тоді

$$\log_2 6871142396\,067532800 = 62.5752556\,894213,$$

тобто, $H(K) = 62.5752556894213$ [біт].

І, на кінець, визначимо відстань єдиності для розроблюваної ключової системи за формулою (8):

$$\frac{62.5752556\,894213}{0.00004226\,7419412} = 1480460.75\,583801,$$

тобто, $U = 1480460{,}75583801$ [літер].

Виберемо для нашого прикладу довжину ключового слова за формулою

$$k = \log_x H(M), \qquad (9)$$

де x – кількість літер в алфавіті, що використовується для складання ключового слова, $H(M)$ – ентропія повідомлення.

Якщо для ключового слова використовувати лише літери української мови (33 літери), тоді довжина ключового слова k буде дорівнювати:

$$\log_{33} 6871142396\,067532800 = 12.405,$$

тобто $k \approx 13$[літер].

Якщо ж, для побудови ключового слова використовувати ще й цифри (0-9), тоді довжина ключового слова повинна буди не меншою за 12 літер. $\log_{43} 6871142396\,067532800 = 11.53$, тобто $k \approx 12$[літер].

Порівнюючи отримані значення відстані єдиності та довжини ключового слова, бачимо, що на відміну від відомих методів забезпечення режиму досконалої стійкості, обсяги текстової інформації, що потребують досконало стійкого захисту, у даному випадку можуть суттєво перевищувати довжину ключового слова.

Залежність відстані єдиності та ентропії ключа шифру від довжини повідомлення. Підставляючи вирази (4) і (5) у вираз (8), отримаємо наступне:

$$U = \frac{H(K)}{\log_2(B) - \frac{H(M)}{n}}. \qquad (10)$$

Виходячи з властивостей досконало стійкої криптосистеми системи [4], кількість ключів шифру K повинна дорівнювати N - кількості смислових варіантів повідомлень довжиною n. Таким чином, за умови

$$H(K) = H(M) = \log_2 N, \qquad (11)$$

підставляючи (11) у (10), вираз для визначення відстані єдиності можна записати наступним чином:

$$U = \frac{\log_2 N}{\log_2(B) - \frac{\log_2 N}{n}}. \qquad (12)$$

Тепер знайдемо залежність N – кількості можливих варіантів повідомленнь від n – довжини повідомлення. При визначені N слід мати на увазі те, що кожен рядок у таблиці (тобто, кожна літера у повідомленні, якщо використати розглянутий вище спосіб укрупнення алфавіту мови відображення інформації у заданій табличній формі) зустрічається тільки один раз (тобто, літери не повторюються). У цьому випадку максимально можлива довжина повідомлення (тобто, максимальна кількість символів у текстовій інформації, що відображають зміст таблиці і передаються через канал зв'язку) дорівнює кількості літер в алфавіті мови відображення інформації, що синтезована для даної табличної форми. Отже, з урахуванням (1) та (2) запишемо вираз для визначення N – кількості можливих значень повідомлення при різних n:

$$N = \prod_{s=1}^{n}\left[B - (s-1)\right]. \qquad (13)$$

Таким чином, при $B = const$ і при $H(K) = H(M)$ (умова забезпечення режиму досконало стійкої криптосистеми) можна записати наступне (підставивши вираз (13) у (12)):

$$U(n) = \frac{\log_2 \prod_{s=1}^{n}\left[B - (s-1)\right]}{\log_2(B) - \frac{\log_2 \prod_{s=1}^{n}\left[B - (s-1)\right]}{n}}. \qquad (14)$$

Вираз (14) визначає залежність значень відстані єдиності U від довжини повідомлення n, що представлена на рис. 1.

Вираз, що визначає залежність ентропії ключа шифру *H(K)* від довжини повідомлення *n*, з урахуванням (13) можна записати наступним чином:

$$H(K) = \log_2 \prod_{s=1}^{n} \left[B - (s-1) \right]. \qquad (15)$$

Графіки залежності $U = f(n)$, $H(K) = f(n)$ показано на рис. 1 та рис. 2 відповідно.

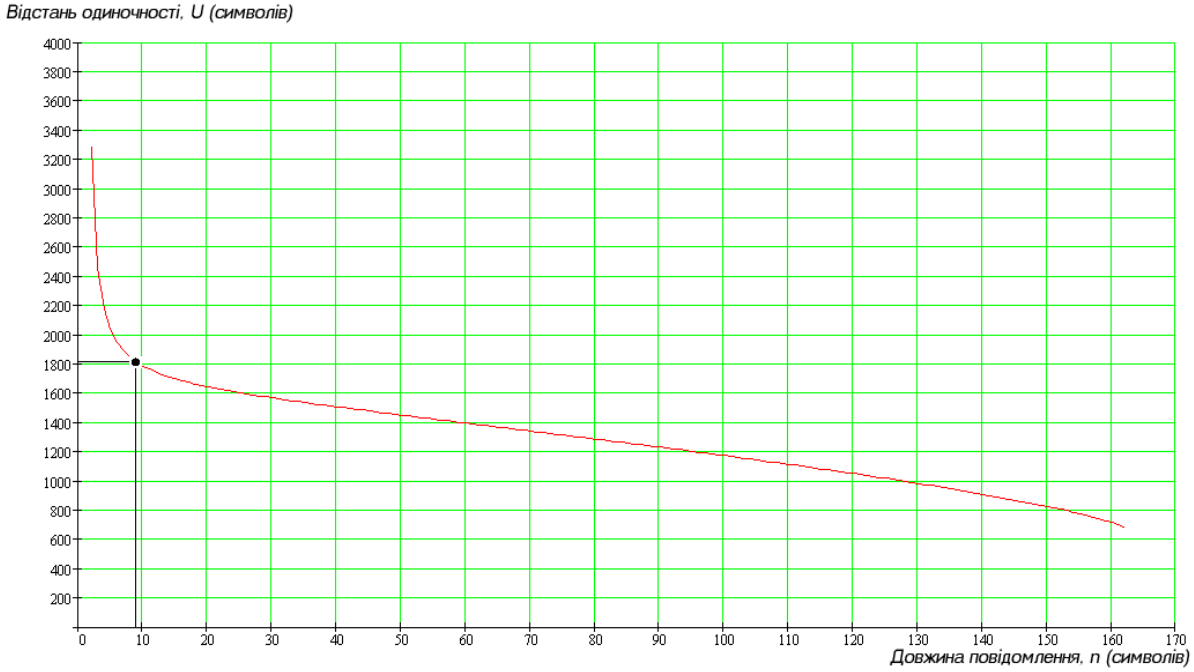

Рис. 1. Графік залежності відстані єдиності *U* від довжини повідомлення *n*

Приклад. Нехай маємо семантичний словник заданої табличної форми, що складається із чотирьох підсловників, тобто *s=4*. Ці підсловники мають певну кількість лінгвістичних одиниць, а саме $S_1=3$, $S_2=6$, $S_3=3$, $S_4=3$.

Визначимо кількість літер в алфавіті мови відображення інформації у даній табличній формі відповідно до формули (1). Отримаємо у даному прикладі *B=162*. У той же час величина *B* визначає максимальну довжину обсягу текстової інформації, що підлягає захисту повідомлення на протязі одного сеансу зв'язку, оскільки кожна літера у повідомленні зустрічається тільки один раз.

З графіку на рис. 1 видно, що навіть при максимальній за цих умов довжині повідомлення *n=162*, відстань єдиності дорівнює $U \approx 680$ символів.

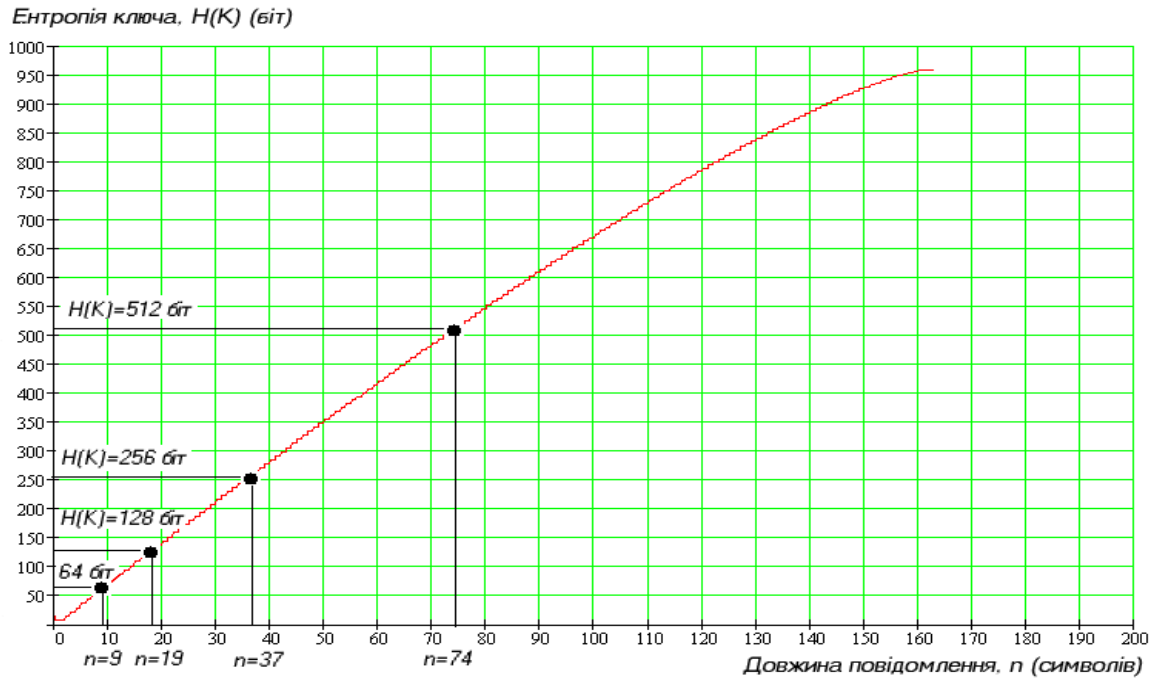

Рис. 2. Графік залежності ентропії ключа шифру *H(K)* від довжини повідомлення *n*

Отже, шляхом укрупнення алфавіту мови відображення інформації у таблиці заданої форми можемо після відповідних перетворень мати шифротексти, які значно коротші за відстань єдиності. Так що, згідно теорії К. Шеннона [3], такі шифротексти можна дешифрувати декількома способами, причому кожен з них може бути коректним, і таким чином забезпечувати досконалий захист текстової інформації, поставивши аналітика перед вибором правильного повідомлення із множини кількох можливих правильних повідомлень.

Графіки, що зображені на рис. 1 та рис. 2, доцільно використовувати для обчислення максимально можливої кількості сеансів зв'язку без зміни ключа шифру при дотриманні умов, коли не порушується ознака досконало стійкої криптосистеми. Зокрема, користуватися цими графіками пропонується наступним чином.

Користувач криптосистемою задається сталим розміром ключа шифру, наприклад 64 біти. З графіку на рис. 2 видно, що така довжина ключа шифру відповідає довжині повідомлення у 9 символів (тобто, має бути 9 рядків у таблиці заданої форми). З графіку, що зображений на рис. 1, видно, що відстань єдиності, що відповідає цій довжині повідомлення дорівнює $U=1812$ символів. Таким чином, для того, щоб криптосистема зберігала властивості досконало стійкої, при використанні ключа шифру довжиною *64* біти, користувач (передавальна сторона) повинен передавати повідомлення довжиною 9 символів (рядків таблиці) не більше ніж $1812/9 \approx 200$ разів. Іншими словами, користувач повинен змінювати ключ шифру після передавання кожних 200 повідомлень.

Залежність показника ефективності захисту від довжини повідомлення та кількості символів алфавіту. Ефективність методу захисту текстової інформації з укрупненням алфавіту мови відображення цієї інформації доцільно порівнювати із ефективністю методу одноразових блокнотів, оскільки обидва вищезазначені методи захисту можуть бути реалізованими у рамках досконало стійких криптосистем.

У разі застосування методу одноразових блокнотів для захисту інформації у режимі досконалої секретності довжина ключа шифру має дорівнювати довжині повідомлення [4]. Для визначення ентропії такого ключа використовують наступний вираз:

$$H_1(K) = \log_2(B_1^{n_1}), \qquad (16)$$

де n_1 – довжина повідомлення [*літер*], що записане мовою людського спілкування (українська, російська, англійська і т.д.), що має алфавіт B_1 [*літер*].

У той час, як ентропія ключа, що застосовується для захисту інформації при використанні укрупненого алфавіту мови відображення цієї інформації, обчислюється за виразом (15).

Отже, в загальному випадку для визначення показника виграшу у довжині ключа (за інших рівних умов) у разі застосування у криптосистемі укрупненого алфавіту у порівнянні із методом одноразових блокнотів слід використати наступний вираз:

$$Z = \frac{H_1(K)}{H(K)}. \qquad (17)$$

Підставивши вираз (16) в чисельник, а вираз (15) в знаменник виразу (17) отримаємо наступне:

$$Z = \frac{\log_2(B_1^{n_1})}{\log_2 \prod_{s=1}^{n}[B-(s-1)]}. \qquad (18)$$

При цьому слід пам'ятати, що B – алфавіт табличної форми, тобто кількість можливих комбінацій табличних рядків семантичного словника, що штучно створюється в результаті статистичного та семантичного аналізу предметної області, а n – кількість рядків табличної форми.

Приклад. Нехай вихідна таблиця з відкритою інформацією, що підлягає шифруванню, має форму у вигляді три рядки х чотири стовпця. Так що семантичний словник складається з чотирьох підсловників, тобто $s=4$.Ці підсловники мають наступну кількість лінгвістичних одиниць: $S_1=3$, $S_2=6$, $S_3=3$, $S_4=3$. От же маємо: $n=3$, $B=162$, $n_1=180$, $B_1=50$ (літери української мови, цифри *0-9,* такі знаки як *"-", ",", ".",* "пробіл", "»", "«").

За таких вихідних даних відповідно до виразу (18) значення показника виграшу у довжині ключа $Z = 46{,}2$. Тобто, при застосуванні досконало стійкої криптосистеми з укрупненим алфавітом довжина ключа шифру може бути у *46,2* рази менша за довжину ключа шифру при застосуванні методу одноразових блокнотів.

Побудуємо графіки залежності виграшу у довжині ключа від довжини повідомлення $Z = f(n_1)$ (див. рис. 3). Порівнюється ефективність застосування криптосистеми з укрупненим алфавітом (щодо довжини ключа шифру) у порівнянні із ефективністю застосування методу одноразових блокнотів.

Із графіків на рис. 3 видно, що виграш у ефективності застосування криптосистеми з укрупненим алфавітом у порівнянні з криптосистемою, що реалізує метод одноразових блокнотів, лінійно залежить від довжини повідомлення.

Розглянемо графіки залежності виграшу в ефективності застосування криптосистеми з укрупненим алфавітом (у порівнянні із ефективністю методу одноразових блокнотів) від кількості символів алфавіту звичайної мови людського спілкування, тобто визначимо залежність $Z = f(B_1)$ (див. рис. 4).

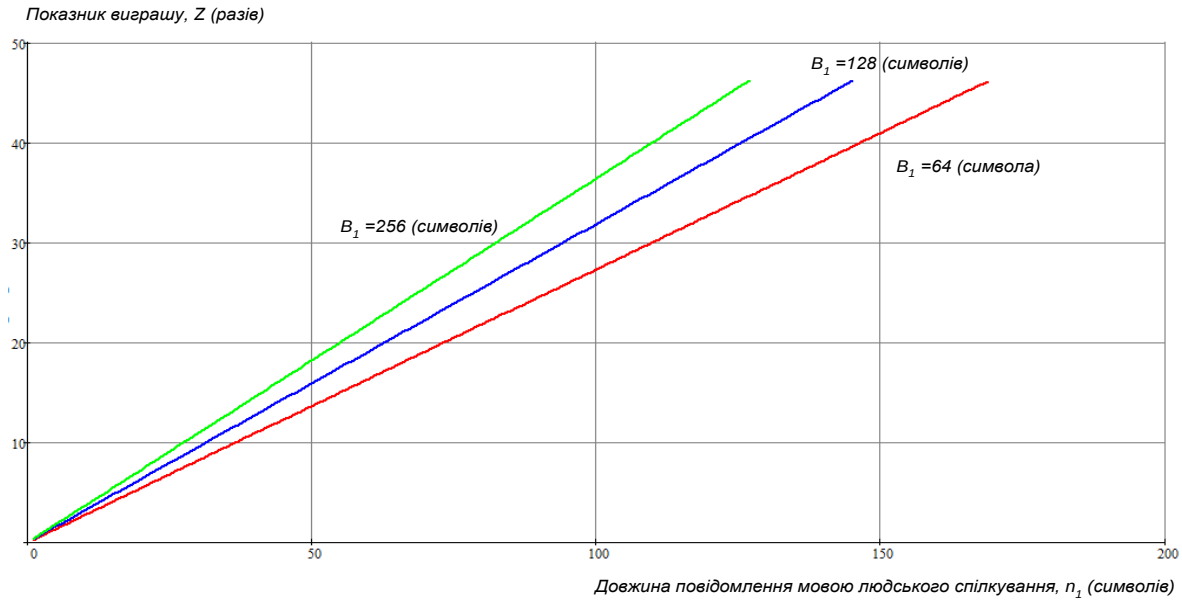

Рис. 3. Графік залежності виграшу в ефективності застосування криптосистеми з укрупненим алфавітом (щодо довжини ключа шифру) від довжини повідомлення $Z = f(n_1)$ у порівнянні із ефективністю застосування методу одноразових блокнотів

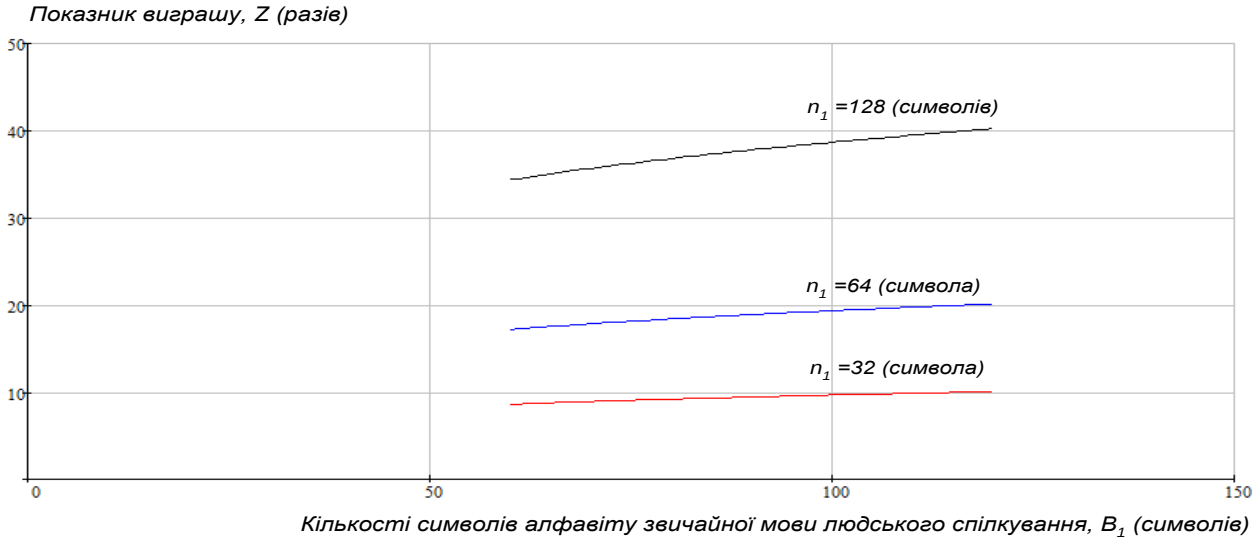

Рис. 4. Графік залежності виграшу Z від кількості символів алфавіту звичайної мови людського спілкування $Z = f(B_1)$

Із графіків на рис. 4 видно, що виграш в ефективності застосування криптосистеми з укрупненим алфавітом у порівнянні з методом одноразових практично не залежить від алфавіту використаної мови людського спілкування, а суттєво залежить від довжини повідомлень, що представлені на цій мові.

Висновки.

1. Запропоновано метод побудови досконало стійкої криптосистеми захисту від порушень конфіденційності текстової інформації, що береться із тезаурусу наперед визначеної прикладної області і поміщається у задану табличну форму. Метод базується на застосуванні механізму укрупнення алфавіту мови відображення текстової інформації, що використовує тезаурус прикладної області і враховує структуру табличної форми. Внаслідок укрупнення алфавіту збільшується так звана відстань єдиності, що є основним пороговим показником приналежності криптосистеми до класу досконало стійких систем захисту з теоретично дове-

деною ідеальною теоретико-інформаційною стійкістю. Невеликі значення відстані єдиності при шифруванні повідомлень, складених із символів алфавітів природних мов, обумовлюють необхідність частої зміни ключової інформації, що є проблемою для багатьох прикладних застосувань. Даний метод суттєвою мірою усуває цю проблему і дозволяє розширити області використання досконало стійких криптосистем.

2. Отримано математичні вирази та побудовано відповідні графіки, що визначають залежності відстані єдиності та ентропії ключа шифру від довжини повідомлення. Отримані результати доцільно використовувати для обчислення максимально можливої кількості сеансів зв'язку без зміни ключа шифру при дотриманні умов, коли не порушується ознака досконало стійкої криптосистеми.

3. Ефективність методу захисту текстової інформації з укрупненням алфавіту мови відображення цієї інформації порівняно із ефективністю методу одноразових блокнотів, оскільки обидва вищезазначені методи захисту можуть бути реалізованими у рамках досконало стійких криптосистем. Ефективність у даному випадку розуміється як виграш у довжині ключа (за інших рівних умов) у разі застосування у криптосистемі укрупненого алфавіту у порівнянні із методом одноразових блокнотів. Отримано математичний вираз для визначеного показника виграшу. Показано, що показник виграшу лінійно залежить від довжини повідомлень, що представлені на синтезованій мові, практично не залежить від алфавіту використаної мови людського спілкування, а суттєво залежить від довжини повідомлень, що представлені на цій мові.

4. Ефективність запропонованого методу побудови досконало стійкої криптосистеми, якщо в якості критерію ефективності обрана відстань єдиності, характеризується суттєво вищими рівнями у порівнянні, із ефективністю інших методів (зокрема, методу одноразових блокнотів) забезпечення режиму досконалої секретності.

ЛІТЕРАТУРА

[1]. А. Зубов, *Совершенные шифры: Вступительное слово чл.-корр. РАН Б.А. Севастьянова.* М.: Гелиос АРВ, 2003, 160 с.

[2]. Ван Тилборг, *Основы криптологии. Профессиональное руководство и интерактивный учебник*. – М.: Мир, 2006, 471 с.

[3]. C. Shannon, "A Mathematical Theory of Communication", *Bell System Technical Journal*, том 27, № 4, С. 379-423, 623-656, 1948.

[4]. С. Сушко, Г. Кузнєцов, Л. Фомичова, А. Корабльов, *Математичні основи криптоаналізу*. Дніпропетровськ: Національний гірничий університет, 2010. -465 с.

[5]. C. Shannon, "Predication and Entropy in Printed English" *Bell System Technical Journal*, том 30, №1, С. 50-64, 1951.

[6]. В. Широков, О. Бугаков, Т. Грязнухіна *Корпусна лінгвістика*. К.: Довіра, 2005, 471 с.

REFERENCES

[1]. A. Zubov, *Perfect codes: the Introductory word of member-correspondent RAS Sevastyanov B.A.* M.: GELIOS ARV, 2003, 160 p.

[2]. Van Tilborg, *Encyclopedia of cryptography and security*. New York: Springer, 2005, 684 p.

[3]. C. Shannon, "A Mathematical Theory of Communication". *Bell System Technical Journal*, vol. 27, no. 4, pp. 379-423, 623-656, 1948.

[4]. S. Sushko, G. Kuznetsov, L. Fomichova, A. Korablyev, *Encyclopedia of cryptanalysis*. Dnipropetrovsk: National Munity University, 2010, 465 p.

[5]. C. Shannon, "Predication and Entropy in Printed English", *Bell System Technical Journal*, vol. 30, no. 1. pp. 50-64, 1951.

[6]. V. Shyrokov, O. Bugakov, T. Gryaznukhina, *Cabinet-type linguistics*. K.: Dovira, 2005, 471 p.

ЭФФЕКТИВНОСТЬ СОВЕРШЕННО СТОЙКОЙ КРИПТОСИСТЕМЫ С УВЕЛИЧЕННЫМ РАССТОЯНИЕМ ЕДИНСТВЕННОСТИ

Для защиты информации, критически важной для государства, целесообразно использовать криптосистемы с теоретически доказанной идеальной теоретико- информационной стойкостью. Однако существующие совершенно стойкие криптосистемы имеют ограниченную область применения, прежде всего, из-за жесткого ограничения в процессе шифрования на параметр непревышения так называемого расстояния единственности по ключу. Относительно небольшие значения расстояния единственности при шифровке сообщений, составленных из символов алфавита любого из естественных языков, обусловливают необходимость частой смены ключевой информации, что является проблемой во многих приложениях. Показана возможность увеличения расстояния единственности путём синтеза искусственного языка отображения прикладной области с алфавитом большой размерности. Определена эффективность совершенно стойкой криптосистемы защиты текстовой информации, которая помещена в заданную табличную форму, при условии, что эта текстовая информация берется из тезауруса заранее выбранной прикладной области. Получены математические выражения и построены соответствующие графики, которые определяют зависимость расстояния единственности и энтропии ключа шифра от длины сообщения. Показано на количе-

ственном уровне, что эффективность метода построения совершенно стойкой криптосистемы с укрупненным алфавитом языка отображения текстовой информации является существенно выше в сравнении с эффективностью других методов обеспечения режима совершенной секретности.

Ключевые слова: защита текстовой информации, расстояние единственности, синтез языка отображения информации, укрупнение алфавита.

EFFICIENCY OF QUITE PROOF CRYPTOSYSTEMS WITH MEGASCOPIC DISTANCE OF UNICITY

For defence critically of important for the state information it is expedient to use to perfection proof cryptosystems with the ideal information the oretical firmness well-proven in theory. However existent to perfection proof cryptosystems have limit area of application, foremost, through hard limitation in relation to unexceeding during enciphering of the so-called distance of unicity after the key. The relatively small values of distance of unicity at enciphering of the reports made from the symbols of alphabet any of human languages stipulate the necessity of frequent change of key information that is a problem for many applied applications. It is shown that possibility of increase of distance of unicity is due to the synthesis of artificial language of reflection of application area with the alphabet of large dimension. Efficiency is examined to perfection proof cryptosystems of defence of text information that is placed in the set table form, on condition that this text information undertakes from to the thesaurus of beforehand certain application area. Mathematical expressions are got and corresponding graphic arts that determines dependences of distance of unicity and entropy of the code key on length of report are built. It is shown at quantitative level, that efficiency of method of construction to perfection proof cryptosystems with the large-sized alphabet of language of reflection of text information is substantially higher in comparing to efficiency of other methods of providing of the mode of perfect secrecy.

Keywords: defence of text information, distance of unicity, synthesis of language of reflection of information, enlargement of alphabet.

Самойлик Євген Олександрович, пошукач кафедри телекомунікаційних систем Національного авіаційного університету.
E-mail: samojlikzhenya@gmail.com
Самойлик Евгений Александрович, соискатель кафедры телекоммуникационных систем Национального авиационного университета.
Samoylik Yevgen, graduate student of Department of Telecommunication Systems, National Aviation University.