DOI 10.20535/2411-1031.2025.13.1.328982
УДК 004.056.53

СЕРГІЙ ІВАНЧЕНКО
ВАДИМ ЯРОЩУК

ВПЛИВ НАДЛИШКОВОСТІ НА БЕЗПЕКУ ТЕХНІЧНИХ КАНАЛІВ ВИТОКУ ІНФОРМАЦІЇ ТА НАБЛИЖЕНЕ КОРИГУВАННЯ ІМОВІРНОСТІ ПОМИЛКИ В КАНАЛІ

У статті розглядається вплив надлишковості на безпеку технічних каналів витоку інформації та наближене коригування імовірності помилки в каналі за умови відсутності відомостей про походження надлишковості. Досліджуються механізми внесення надлишковості в сигнали, що сприяють покращенню завадостійкості шляхом виправлення помилок, а також створюють ризики зниження конфіденційності переданої інформації.

Розглянуто два основні типи надлишковості: штучну, яка формується через кодування та контрольні символи, і природну, що є наслідком особливостей джерела інформації. Принципи її декодування суттєво відрізняються. Це унеможливлює безпосереднє застосування методів коригування імовірності помилки, розроблених для штучної надлишковості, до природної. Штучна надлишковість має чітку структуру і застосовується для підвищення завадостійкості, тоді як природна надлишковість виникає внаслідок особливостей джерела інформації, проявляючись через кореляції між символами, повторювані шаблони та статистичні закономірності.

Розглянуто особливості мовних і візуальних каналів, де природна надлишковість відіграє ключову роль у виправленні помилок. В мовних каналах вона сприяє підвищенню достовірності прийому інформації через суб’єктивність вимови та сприйняття, а у візуальних — через закономірності розподілу пікселів у зображеннях.

Запропоновано новий підхід до декодування, що базується на впорядкуванні кодових комбінацій за вагою. Використано методику поділу можливих 7-бітних комбінацій на групи, що забезпечує ефективніше виправлення помилок. Також розглянуто традиційний підхід до виправлення помилок у коді Хеммінга (7,4) та його обмеження при високих рівнях завад.

Отримані результати можуть бути використані для вдосконалення методів підвищення достовірності передачі даних та зменшення ризиків витоку інформації через технічні канали. Важливо, що цей метод не тільки покращує коригування помилок, але й відкриває нові можливості для адаптивного кодування в складних умовах передачі інформації. Його універсальність дозволяє застосовувати підхід у різних сферах, від цифрового зв’язку до мовної обробки сигналів.

Ключові слова: надлишковість, імовірність помилки, коригувальні коди, інформаційна безпека, оптимальне приймання, код Хеммінга, інформаційно-комунікаційні системи.

Постановка проблеми. Надлишковість, яка, як правило, притаманна смисловим джерелам інформації, є фактором щодо покращення передачі інформації каналами зв’язку та, водночас, погіршення інформаційної безпеки щодо витоку інформації технічними каналами. Завдяки її внесенню в канали мають місце методи виявлення та виправлення помилок на прийомі. Прикладами таких методів, що пов’язані з внесенням надлишковості в каналах передачі даних, є наявність перевірочних символів завадостійкого кодування та повтори сеансів передачі, які здійснюються за окремими запитами передачі. Всі різновиди надлишковостей джерел можуть покращувати якість передачі даних та погіршувати захищеності даних в технічних каналах витоку.

Однак, у повідомленні вона може бути і відомого походження, якщо вноситься штучно, і невідомого, якщо є проявом особливостей джерела інформації.

За надлишковістю відомого походження або штучної надлишковості існують способи декодування, які включають в себе механізми виправлення помилок та дозволяють розрахунок скоригованої імовірності помилки [1]. Це здійснюється посередництвом представлення таблиці істинності коду у виді стандартного розташування комбінацій у суміжних класах.

Для надлишковостей невідомого походження або природньої надлишковості смислових джерел (мовних, візуальних тощо), яка також дозволяє виправлення помилок, можливість такого представлення комбінацій відсутня [1]. Відсутня також і можливість розрахунку скоригованої імовірності помилки, яка враховує можливість виправлення завдяки надлишковості.

Як правило, для врахування виправляючої здатності природньою надлишковістю використовують емпіричні оцінки якості каналу, які, в свою чергу, враховують особливості джерел та приймачів, розкид параметрів сигналів за спектром, за динамічним діапазоном тощо.

В каналах передачі мовної інформації це обґрунтовується суб'єктивністю походження джерела інформації і приймача. Саме суб'єктивність сприяє особливостям формування надлишковості при мовленні, яка виражається через розкид параметрів по статистиці розподілу літер у тексті, по їх спектральним характеристикам при озвученні, та розподілу динамічного діапазону. Цей розкид параметрів по різному притаманний для мовних джерел різних мов, однак він є завжди присутнім. Як правило, особливість цих джерел є корисною для спілкування, покращено забезпечує різницю поміж звуками, що відповідають літерам алфавіту, та компенсується суб'єктивністю вимовлення голосовим апаратом людини та сприйняття її слуховим апаратом.

В каналах для візуальної інформації зображення має об'єктивне походження, але сприймається людиною-суб'єктом через її зоровий апарат. Як правило, зображення це об'єкт чи декілька об'єктів на фоні, які можуть мати різну детальність: крупну, де для зображення деталей використовується багато пікселів, та дрібну, де зображення деталей використовує невелику, а то й, наскільки це можливо, мінімальну кількість пікселів. Деталі зображення мають таку особливість, що вони відрізняються від фону, на якому вони зображені, та інших деталей. Пікселі однієї деталі зазвичай мало чим відрізняються, а то і зовсім між собою не відрізняються. Наприклад, жирна крапка, товста лінія, зображення літер тексту, як правило використовують піксель одного кольору. Очевидно, що виправлення помилкових пікселів серед їх великої кількості, що відносяться до однієї крупної деталі, здійснюється набагато простіше, ніж серед малої кількості пікселів дрібнодетального зображення.

Все це є прикладом природньої надлишковості, яка дозволяє підвищувати достовірність прийому інформації. Завдяки цьому можливе виправлення помилок. Як правило, це виправлення здійснюється інтуїтивно завдяки природній пристосованості людини-приймача. Оскільки таке виправлення принципово можливе, то можливе існування відповідних технологій для автоматизованої обробки даних та відповідного підвищення достовірності прийому та зниження захищеності інформації від витоку технічними каналами завдяки природній надлишковості джерела.

Якщо для штучної надлишковості практично завжди є можливість представлення правила кодування та декодування, як правило, у виді таблиці істинності, то для природньої надлишковості таке представлення не завжди є можливим, а то і зовсім неможливим, особливо коли це стосується великих масивів даних. Якщо для штучної надлишковості не складно скоригувати імовірність помилки за таблицею істинності, то для природньої це завдання потребує окремих наукових вирішень.

Аналіз останніх досліджень і публікацій. Слід зазначити, що вирішенню зазначеного питання було приділено багато праць вчених світу [2]–[10]. Так, у роботах [2], [3]

досліджується застосування методів машинного навчання для адаптивної корекції помилок із використанням надлишковості даних. Ці підходи дозволяють оптимізувати параметри завадостійкого кодування і знижувати ймовірність помилок у каналі шляхом аналізу статистичних характеристик інформації. Роботи [4], [5] присвячені огляду сучасних підходів, що орієнтовані на штучне введення надлишковості з метою створення корегувальних кодів.

Значна увага приділяється також дослідженням методів виправлення помилок у каналах зв'язку з високим рівнем завад [6]. Розглядаються класичні та сучасні алгоритми завадостійкого кодування, зокрема коди БЧХ, LDPC та турбокоди, що забезпечують ефективне виправлення помилок при передачі цифрових даних. У роботі [7] розглянуто використання комбінованих підходів, що поєднують кодування з виправленням помилок та адаптивну обробку сигналів, дозволяючи покращити ефективність передачі інформації в умовах змінних характеристик каналу.

Крім того, у дослідженні [8] розглядається питання підвищення достовірності прийому інформації за допомогою технологій штучного інтелекту. Використання нейронних мереж та алгоритмів глибокого навчання дозволяє покращити роботу декодерів шляхом адаптивного вибору оптимальних параметрів коригувального кодування. Це сприяє зниженню ймовірності помилок та підвищенню надійності передачі інформації навіть у складних умовах.

Роботи [9], [10] розглядають можливості використання контрольних сум та циклічних кодів для детектування та виправлення помилок. Такі методи широко застосовуються у цифрових комунікаціях та зберіганні даних, однак їх ефективність обмежена у випадку складних атак на інформаційні системи. Це підтверджує той факт, що сучасні методи коригування помилок орієнтовані переважно на покращення достовірності прийому, а не на забезпечення захисту інформації від витоку технічними каналами.

Однак, всі ці праці присвячені дослідженням коригувальних властивостей штучної надлишковості, яка має забезпечувати ефективність передачі інформації в каналах зв'язку. Застосування ж штучного інтелекту та методів машинного навчання направлені, знову ж таки, на підвищення достовірності прийому, а не на забезпечення захисту інформації від витоку технічними каналами з відповідними оцінюваннями показників.

Вплив же надлишковості на безпеку технічних каналів витоку інформації та наближене коригування імовірності помилки в каналі залишається поза межами наукової уваги вчених. В результаті аналізу інших досліджень і публікацій не виявлено матеріалів, які присвячені врахуванню надлишковості джерел у захищеності інформації від витоку технічними каналами, тим більше матеріалів, які представляють методи чи методики коригування імовірності помилки в каналі.

Таким чином, має місце актуальність завдання щодо врахування надлишковості джерел у захищеності інформації від витоку технічними каналами та коригування імовірності помилки в каналі.

Метою статті є обґрунтування розрахункового апарату щодо коригування імовірності помилки в технічному каналі витоку для надлишкових джерел інформації, який дозволив би для захисту інформації від витоку автоматизовано враховувати надлишковість джерела незалежно від її походження.

Виклад основного матеріалу дослідження. Для досягнення мети та вирішення поставленого в статті завдання має місце доцільність огляду механізмів, правил кодування та декодування з виправленням помилок в каналі, а також підходів щодо коригування імовірності помилки.

Найпростішим прикладом завадостійкого кодування є лінійний систематичний код.

Нехай задано канал, який використовує завадостійкий систематичний (n, k) код з породжувальною матрицею G, що виправляє помилки (див. рис. 1):

$$G = [I \mid C], \tag{1}$$

де I – одинична діагональна матриця;

C – матриця, що формує перевірочні символи коду та складається з лінійно незалежних рядків та стовпців.

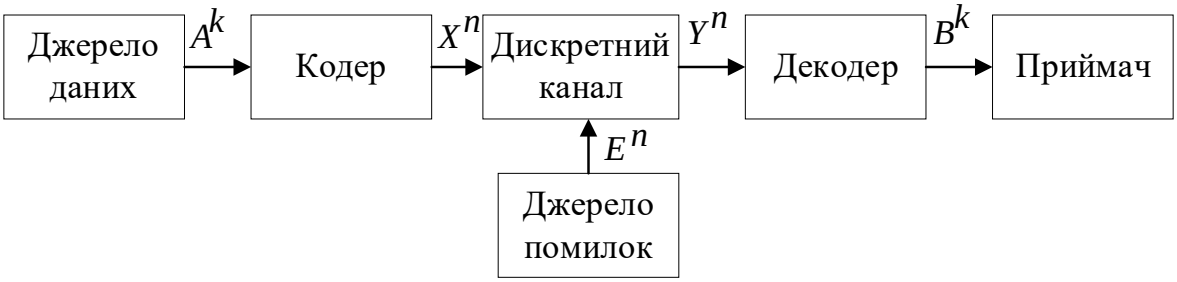

Рисунок 1 – Канал передачі дискретних даних з використанням завадостійкого кодування

Джерело виробляє послідовність дискретних даних $A_i^k = (\alpha_1, \alpha_2, ..., \alpha_k)$ довжиною k, де $\alpha = \{0,1\}$;

i – номер комбінації, $i = 1, 2, ..., 2^k$.

В кодері здійснюється кодування $A_i^k \to X_i^n$, яке вносить в повідомлення надлишковість $n > k$:

$$X^n = A^k \times G^{n \times k}, \tag{2}$$

де X – матриця розмірністю n, елементами якої є значенні послідовності X^n;

A – матриця розмірністю k, що відповідає послідовності A^k;

G – породжувальна матриця завадостійкого коду розмірністю $n \times k$.

Сформована після кодера послідовність $X_i^n = (x_1, x_2, ..., x_n)$ довжиною n, де $x = \{0,1\}$ потрапляє на вхід каналу.

Вихідна послідовність каналу:

$$Y_{gl}^n = (y_1, y_2, ..., y_n) = (X_i^n + E_s^n) \bmod 2^n = (x_1 \oplus e_1, x_2 \oplus e_2, ..., x_n \oplus e_n), \tag{3}$$

де E_n – послідовність помилок $E_s^n = (e_1, e_2, ..., e_n)$ довжиною n, на яку відрізняється вихідна послідовність Y_{gl}^n від вхідної Y_i^n, $l = 1, 2, ..., 2^k$, $g = 1, 2, ..., 2^{n-k}$;

s – номер комбінації E_s^n, $s = 1, 2, 3, \ldots, 2^n$;

$\oplus$ – операція додавання за модулем 2.

При декодуванні здійснюється перевірка на синдром S_g^{n-k}, $g = 1, 2, ..., 2^{n-k}$, який відповідає деяким найбільш імовірним комбінаціям помилок $E_s'^n = (e_1, e_2, ..., e_n)$ довжини n $s = 1, 2, ..., 2^{n-k}$:

$$S^{n-k} = Y^n \times [H^{n \times (n-k)}]^T, \tag{4}$$

де H – перевірочна матриця коду розмірністю $n \times (n-k)$.

За синдромом S^{n-k} знаходять відповідну комбінацію помилок:

$$S^{n-k} \leftrightarrow E_s'^n. \tag{5}$$

Від отриманої на виході каналу послідовності Y^n віднімають $E_s'^n$:

$$Y_s'^n = Y^n \oplus E'^n = X^n \oplus E^n \oplus E'^n. \tag{6}$$

Здійснюється перетворення з використанням оберненої до породжувальної матриці коду:

$$B^k = Y^n \times G^{-1} = [X^n \oplus E^n \oplus E'^n] \times G^{-1} = X^n \times G^{-1} \oplus [E^n \oplus E'^n] \times G^{-1} = \\ = A^k \oplus [E^n \oplus E'^n] \times G^{-1}, \quad (7)$$

де B – матриця, що відповідає послідовності B_l^k довжиною k, $B_l^k = (\beta_1, \beta_2, ..., \beta_k)$, $\beta = \{0,1\}$.

Вище зазначене кодування можна представити в суміжних класах. Це здійснюється наступним чином (див. рис. 2).

Кожному суміжному класу має відповідати синдром S_s^{n-k}, за яким і знаходиться E'^n.

Оскільки код лінійний та перша комбінація A_1^k складатиметься з нулів, то з нулів складатиметься і X_1^k як результат добутку (2). Це означає, що в суміжних класах $Y_{g1}^n = E_s'^n$, де $g = 1, 2, ..., 2^{n-k}$ та $s = 1, 2, ..., 2^{n-k}$.

Таким чином, комбінацію Y^n, що на виході каналу, перевіряють на синдром. Віднімають від отриманої Y^n лідер відповідного суміжного класу та переходять в код X^n, за яким не складно знаходиться комбінація B^k, яка має відповідати переданій A^k. При цьому, що слід зазначити, зазначена операція віднімання відповідає операції додавання за модулем 2.

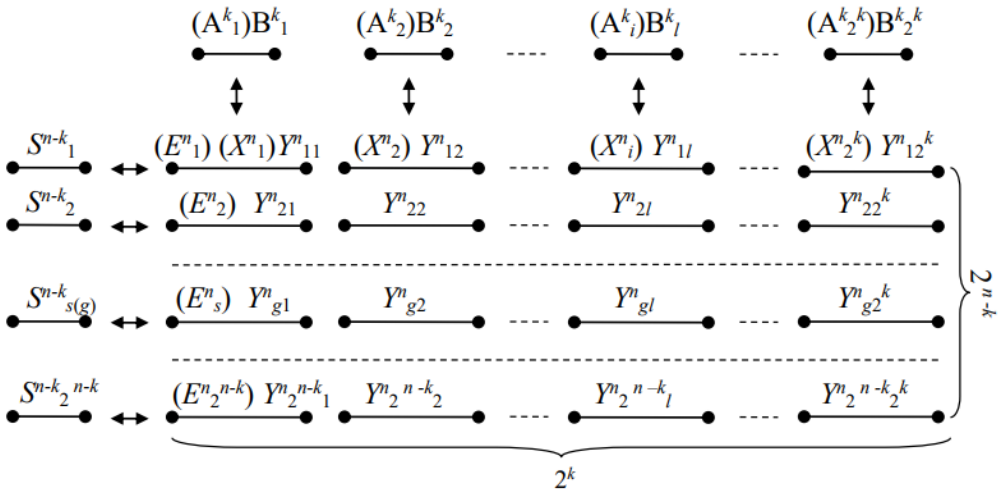

Рисунок 2 – Стандартне розташування завадостійкого коду в суміжних класах

У верхньому рядку представлені всі комбінації двійкових знаків $A_i^k = (\alpha_1, \alpha_1, ..., \alpha_k)$ довжиною k, де $\alpha = \{0,1\}$, i – номер комбінації, $i = 1, 2, ..., 2^k$, які може виробляти джерело та які потрапляють на вхід кодера. Цими ж комбінаціями і є вихідні комбінації знаків з декодера $B_l^k = (\beta_1, \beta_1, ..., \beta_k)$, $\beta = \{0,1\}$, $l = 1, 2, ..., 2^k$ тієї ж довжини.

Суміжні класи формуються шляхом додавання до кодових слів X_i^n комбінацій помилок $E_s'^n$, де $s = 1, 2, ..., 2^{n-k}$. Як правило комбінації $E_s'^n$ вибираються як найменшої ваги, оскільки це забезпечує їх найбільшу імовірність та оптимізує виправляючу здатності коду. Таким чином, суміжні класи відрізняються між собою на один і той же випадковий вектор $E_s'^n$:

$$Y_{si}^n = (y_1, y_2, ..., y_n) = X_i^n \oplus E_s'^n. \quad (8)$$

Кожному суміжному класу відповідає синдром S_s^{n-k} – комбінація даних (номер суміжного класу у двійковій системі обчислення), якій відповідає певний випадковий вектор $E_s'^n$.

Нехай дискретним каналом, що поміж X та Y є дискретний симетричний канал без пам'яті з імовірністю помилки $p < 0,5$. Тоді існує імовірність того, що із всіх 2^n комбінацій помилок E_s^n знайдуться такі, що не належатимуть 2^{n-k} лідерів суміжних класів $E_s^n \neq E_s'^n$.

Це ті помилки які код не може виправляти. Ними є випадкові $E_s^n \neq E_s'^n$, де $i \neq 1$. Помилки ж $E_s^n = E_s'^n$ є комбінаціями нулів та одиниць, що виправляються кодом. При передачі даних по каналу $X \to Y$ на виході з'являються послідовності:

$$Y_{hl}^n = (y_1, y_2, ..., y_n) = X_i^n \oplus E_{sg}^n, \tag{9}$$

де $g = 1, 2, ..., 2^{n-k}$, $l = 1, 2, ..., 2^k$, $j = 1, 2, ..., 2^k$.

В результаті цього X_i^n перейде не обов'язково в Y^n свого i-го стовпця на рис. 2, а в будь-який інший Y_{gl}^n. При цьому виправлення помилки, що здійснюватиметься при декодуванні, буде пов'язане з переходом Y_{gl}^n в Y_{1l}^n та B_l^k, які можуть не співпадати з X_i^n та A_i^k. А це означає, що помилка не виправлена, тобто $A_i^k \neq B_l^k$.

Тому, аналогічним чином, як для каналу $X \to Y$, опис якого виражено формулою (9), можна описати й канал $A \to B$:

$$B_l^k = A_i^k \oplus \Gamma_j^k, \tag{10}$$

де Γ_j^k – послідовність помилок довжини k, яка відрізняє A_i^k від B_l^k, $\Gamma_j^k = (\gamma_1, \gamma_2, ..., \gamma_k)$, $\gamma = \{0,1\}$, $j = 1, 2, ..., 2^k$.

Як очевидно, імовірність помилки p, що має місце в каналі $X \to Y$, $p = p(e = 1)$, в сукупності з виправляючими властивостями коду, що вносить надлишковість, визначають кориговану імовірність помилки $p_{\text{кор}}$, що має місце в каналі $A \to B$, $p_{\text{кор}} = p(\gamma = 1)$. Імовірність помилки є показником, який в середньому характеризує можливість появи цієї помилки. Тому беззаперечно очевидно, що цю імовірність можна виразити як математичне сподівання ваги $\omega t(\Gamma_j^k)$, поділеної на довжину комбінацій Γ_j^k:

$$p(\gamma = 1) = \frac{1}{k} \sum_{j=1}^{2^k} \omega t(\Gamma_j^k) p(\Gamma_j^k), \tag{11}$$

де $p(\Gamma_j^k)$ – імовірність комбінації коригованих помилок Γ_j^k.

В свою чергу не складно побачити, що

$$p(\Gamma_j^k) = \sum_{g=1}^{2^{n-k}} p(E_{gj}^n), \tag{12}$$

де $p(E_{gj}^n)$ – імовірність комбінації помилок E_{gj}^n в стандартному розташуванні коду:

$$p(E_{gj}^n) = p^{\omega t(E_{gj}^n)} (1-p)^{n-\omega t(E_{gj}^n)}, \tag{13}$$

де $\omega t(E_{gj}^n)$ – вага комбінації помилок E_{gj}^n, визначається стандартним розташуванням коду.

Підставивши співвідношення (13) в (12) та (12) в (11), отримаємо остаточну формулу для коригованої помилки $p(\alpha \neq \beta)$ відносно $p(x \neq y)$:

$$p_{\text{кор}} = \frac{1}{k} \sum_{j=1}^{2^k} \omega t(\Gamma_j^k) \sum_{g=1}^{2^{n-k}} p^{\omega t(E_{gj}^n)} (1-p)^{n-\omega t(E_{gj}^n)}. \tag{14}$$

Оскільки надлишковість може мати різне походження: вона може бути штучно впровадженою, як у випадку кодів з виправленням помилок, або виникати природним чином у процесах передавання й обробки даних, то принципи декодування для цих двох типів

надлишковості, як правило, мають відмінності, що унеможливлює пряме застосування методів коригування імовірності помилки, розроблених для штучної надлишковості, до природної.

Штучна надлишковість створюється цілеспрямовано для підвищення стійкості передавання даних до помилок, наприклад, через застосування коригувальних кодів.

Природна надлишковість, в свою чергу, виникає внаслідок статистичних або структурних особливостей джерела інформації. Вона може проявлятися у вигляді кореляцій між символами, повторюваних шаблонів або внутрішньої передбачуваності даних. Така надлишковість не є результатом спеціального кодування, а тому, зазвичай, не підпорядковується жорстким алгебраїчним правилам.

Крім того, природна надлишковість може варіюватися залежно від конкретного джерела інформації, що ускладнює створення універсальних методів її використання для виправлення помилок. Якщо в штучних кодах надлишковість є детермінованою і стабільною, то природна надлишковість може бути нерівномірно розподіленою та змінюватися у часі. Це означає, що методи, які працюють в одному контексті, можуть виявитися неефективними в іншому, і тому декодування потребує додаткового аналізу та адаптації до конкретних характеристик даних.

Щоб врахувати особливості природної надлишковості при виправленні помилок, необхідно адаптувати підхід до декодування. Як було зазначено, штучна надлишковість має чітку математичну структуру. Натомість природна надлишковість не підпорядковується жорстким алгебраїчним правилам, тому її обробка потребує іншого підходу.

Для усунення зазначених обмежень пропонується модифікований підхід до декодування, який ґрунтується на впорядкуванні кодових комбінацій за їхньою вагою. Основна ідея полягає в тому, щоб сформувати впорядковану множину всіх можливих 7-бітних комбінацій та розділити їх на групи таким чином, щоб кожна група містила кодові слова з однаковою або близькою вагою. Це дозволяє ефективніше розподіляти комбінації та мінімізувати ймовірність помилкового виправлення.

У коді Хеммінга (7,4) використовується 16 кодових слів, які формуються з 4-бітних інформаційних комбінацій. Однак, якщо враховувати всі можливі 7-бітні комбінації, то їх загальна кількість становить 128. Для їх структурованого представлення необхідно впорядкувати всі комбінації за зростанням ваги. Вага $\omega(C)$ визначається як кількість одиниць у комбінації C:

$$\omega(C) = \sum_{j=1}^{7} c_j \,, \tag{15}$$

де c_j – j-й біт комбінації, $j = \{0,1\}$.

На основі ваги $\omega(C)$ всі комбінації впорядковуються у вигляді множини:

$$C_{\text{упор.}} = \{C_1, C_2, ..., C_{128}\} \,, \tag{16}$$

де $\omega(C_i) \le \omega(C_{i+1}), \forall i$.

Всі можливі 7-бітні комбінації поділяються на групи по 8 комбінацій. Кожен блок містить:

1. Базову комбінацію без помилок $C_{k,0}$.

2. Комбінації, що містять рівно одну помилку, тобто всі можливі варіанти з однобітними змінами $C_{k,1}, C_{k,2}, ..., C_{k,7}$. Кожна наступна комбінація утворюється за правилом:

$$C_{k,i} = C_{k,0} \oplus e_i \,, \tag{17}$$

де $e_i = (0,0,...,1,...,0)$ – 7-бітний вектор з одиницею на i-й позиції.

3. Таким чином, кожен блок B_k виглядає наступним чином:

$$B_k = \{C_{k,0}, C_{k,1}, ..., C_{k,7}\}, \tag{18}$$

де $k = 1,2,...,16$ – кількість блоків.

При отриманні кодового слова на виході каналу передбачається, що помилки мають найменшу вагу. Отже, декодування відбувається за правилом мінімальної відстані:

$$C_{\text{декод.}} = \arg\min_{C \in B_k} d_H(R, C), \tag{19}$$

де $d_H(R, C)$ – відстань Хеммінга між прийнятою комбінацією R і можливими комбінаціями в блоці C.

Оскільки кожен блок містить всі можливі комбінації з однобітними помилками, прийнята комбінація R завжди буде коректно декодована до базової комбінації $C_{k,0}$, якщо помилка має кратність $t = 1$.

Результати проведення розрахунків зображені на рис. 3. Запропонований розрахунковий аппарат дозволяє ефективніше коригувати імовірності помилки в технічному каналі витоку інформації для надлишкових джерел, що позитивно впливає на якість зв'язку в каналі та збільшує ьезпеку від витоку технічними каналами.

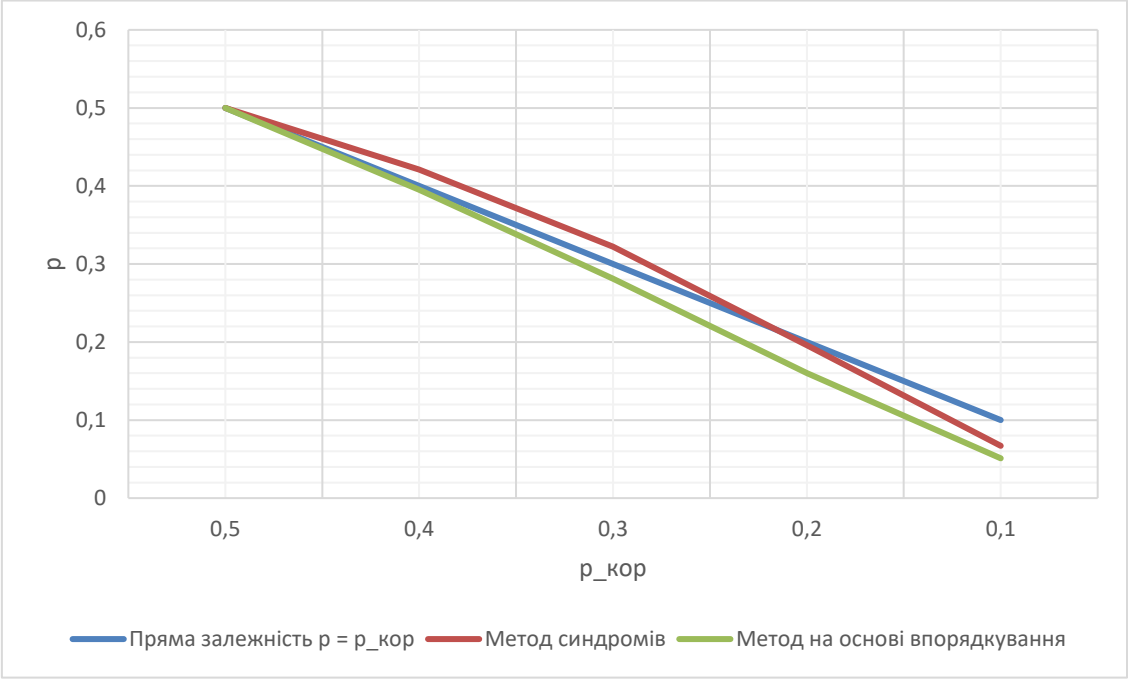

Рисунок 3 – Залежність коригованої ймовірності помилки $p_{\text{кор}}$ від помилки в каналі p

Висновки. Таким чином, здійснено огляд впливу надлишковості на безпеку технічних каналів витоку інформації та наближене коригування імовірності помилки в каналі Обґрунтовано розрахунковий апарат щодо коригування імовірності помилки в технічному каналі витоку для надлишкових джерел інформації, який дозволив би для захисту інформації від витоку автоматизовано враховувати надлишковість джерела незалежно від її походження.

У процесі розробки було досягнуто ряд результатів, що відповідають поставленим задачам, а саме: запропоновано підхід до декодування, який ґрунтується на впорядкуванні кодових комбінацій за їхньою вагою, який би автоматизовано враховувати надлишковість джерела та реалізовано метод обрахунку коригованої імовірності помилки, неважливо від походження надлишковості.

Наукова новизна дослідження полягає в розробці нового підходу до коригування ймовірності помилки, який враховує як природну, так і штучну надлишковість джерела інформації. Запропонований метод заснований на впорядкуванні кодових комбінацій за вагою, що дозволяє ефективніше знижувати рівень помилок у каналі витоку. Дана модель не вимагає попереднього накладання жорсткої штучної надлишковості.

Практична значущість роботи полягає в тому, що запропонований підхід може бути використаний у системах захисту інформації для автоматизованого врахування надлишковості незалежно від її походження. Це дозволяє зменшити ймовірність коригованої помилки, підвищити ефективність методів декодування в умовах перехоплення даних та розробляти більш стійкі алгоритми обробки інформації.

Перспективи подальших досліджень. Результати демонструють перспективи для подальшого підвищення якості безпеки технічних каналів витоку інформації та її практичного застосування. Можливість пов'язання з методами адаптивного коригування помилок із використанням машинного навчання та нейромережевих підходів для аналізу надлишковості джерела. Має місце необхідність дослідження складних моделей природної надлишковості на ефективність декодування та розширити запропонований метод на інші типи коригувальних кодів. Перспективою щодо практичного застосування є розробка адаптованих алгоритмів, що зможуть щодо змінних характеристик інформаційного потоку в реальному часі. Все це дозволить мінімізувати ризики інформаційної безпеки та здійснення надійного захисту від витоку технічними каналами витоку інформації.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] С. Іванченко, О. Гавриленко, О. Липський, та А. Шевцов, *Технічні канали витоку інформації. Порядок створення комплексів технічного захисту інформації*. Київ, Україна: ІСЗЗІ "КПІ", 2015.

[2] P. Upadhyaya, "When Machine Learning Meets Information Theory: Some Practical Applications to Data Storage", PhD dissertation, Texas A&M University, College Station, TX, USA, 2020.

[3] T. Gruber, S. Cammerer, J. Hoydis, and S. Ten Brink, "On deep learning-based channel decoding", in *Proc. 51st Ann. Con. on Inf. Scien. and Sys. (CISS)*, Baltimore, MD, USA, pp. 1-6, 2017, doi: https://doi.org/10.1109/CISS.2017.7926071.

[4] P. Upadhyaya, and A. Jiang, "Machine Learning for Error Correction with Natural Redundancy," arXiv (Cornell University), 2019, doi: https://doi.org/10.48550/arXiv.1910.07420.

[5] О. Нікуліна, В. Северин, та В. Шаров, "Розробка моделі завадостійкої передачі даних для інформаційної технології оптимізації управління динамічними системами", *Віс. НТУ "ХПІ". Серія: Сист. ан., упр. та інф. техн.*, вип. 2 (8), с. 57-62, 2023, doi: https://doi.org/10.20998/2079-0023.2022.02.09.

[6] С.П. Лівенцев та Г.Д. Созонник, "Метод адаптивного управління процесом декодування турбоподібних кодів", на *XVIII Між. наук.-техн. конф. Перспективи телекомунікацій*, Київ, 2024, с. 35-39. [Електронний ресурс]. Доступно: https://its.kpi.ua/sites/default/files/NDI%20TK%202021/Конференція%20ПТ/Збірники%20ПТ/Збірник%20матеріалів%20конференції%20ПТ%202024%20ISSN%20p.pdf. Дата звернення: Січ. 10, 2025.

[7] Ю.М. Бойко, А.І. Семенко, та І.С. П'ятін, "Особливості формування кодової надлишковості у каналах передачі інформації", *Інфоком. та комп. техн.*, т. 2, № 04, с. 12-25, 2023, doi: https://doi.org/10.36994/2788-5518-2022-02-04-01.

[8] І.Р. Цимбалюк, "Методи та засоби підвищення достовірності приймання радіосигналів із використанням нейронних мереж", дис. д-ра філософії, Нац. ун-т "Львів. Політехніка", Львів, 2024. [Електронний ресурс]. Доступно: https://lpnu.ua/sites/default/files/2024/radaphd/27733/disertaciya-cimbalyuk-rev1.pdf. Дата звернення: Лют. 12, 2025.

[9] О.І. Євдоченко, "Особливості багатопозиційних кодів ріда-соломона", на *X Всеукр. наук.-прак. конф. "Електронні та мехатронні системи: теорія, інновації, практика"*, Полтава, 2024, с. 106-107. [Електронний ресурс]. Доступно: https://nupp.edu.ua/uploads/

files/0/events/conf/2024/x-emst/zbirnik-naukovikh-prac-X-vsekrainskoi-konferencii-ems_2.pdf. Дата звернення: Січ. 10, 2025.

[10] С.В. Ільїн, Є.Л. Холод, та А.Б. Мазничко, "Організація довгострокового доступу до автентичних документів", in *Proc. II Int. Sci. Pract. Conf. Creation of new ideas of learning in modern conditions*, Бордо, 2023, с. 252-258. [Електронний ресурс]. Доступно: https://dspace.hnpu.edu.ua/server/api/core/bitstreams/a8627c6c-02af-400e-ab13-4ea4c073c4e3/content. Дата звернення: Січ. 10, 2025.

Стаття надійшла до редакції 18.03.2025

REFERENCE

[1] S. Ivanchenko, O. Havrylenko, O. Lipsky, and A. Shevtsov, *Technical channels of information leakage. The procedure for creating complexes of technical protection of information*. Kyiv, Ukraine: ISCIP "KPI", 2015.

[2] P. Upadhyaya, "When Machine Learning Meets Information Theory: Some Practical Applications to Data Storage", PhD dissertation, Texas A&M University, College Station, TX, USA, 2020.

[3] T. Gruber, S. Cammerer, J. Hoydis, and S. Ten Brink, "On deep learning-based channel decoding", in *Proc. 51st Ann. Con. on Inf. Scien. and Sys. (CISS)*, Baltimore, MD, USA, pp. 1-6, 2017, doi: https://doi.org/10.1109/CISS.2017.7926071.

[4] P. Upadhyaya, and A. Jiang, "Machine Learning for Error Correction with Natural Redundancy," arXiv (Cornell University), 2019, doi: https://doi.org/10.48550/arXiv.1910.07420.

[5] O. Nikulina, V. Severyn, and V. Sharov, "Development of a model of interference-resistant data transmission for information technology of control optimization of dynamic systems", *Bull. of NTU "KhPI". Series: Sys. An., Cont. and In. Tec.*, iss. 2 (8), pp. 57-62, 2023, doi: https://doi.org/10.20998/2079-0023.2022.02.09.

[6] S.P. Liventsev, and G.D. Sozonnik, "Method of adaptive control of the process of decoding turbo-like codes", in *Proc. 18th Int. Sci. and Tec. Conf. Telecommunications prospects*, Kyiv, 2024, pp. 35-39. [Online]. Available: https://its.kpi.ua/sites/default/files/NDI%20TK%202021/Конференція%20ПТ/Збірники%20ПТ/Збірник%20матеріалів%20конференції%20ПТ%202024%20ISSN%20p.pdf. Accessed on: Jan. 10, 2025.

[7] Y.M. Boyko, A.I. Semenko, and I.S. Pyatin, "Features of the formation of code redundancy in information transmission channels", *Infocom. and Comp. Tec.*, vol. 2, iss. 04, pp. 12-25, 2023, doi: https://doi.org/10.36994/2788-5518-2022-02-04-01.

[8] I.R. Tsymbaliuk, "Methods and means of increasing the reliability of radio signal reception using neural networks", PhD thesis, Lviv Polytechnic Nat. Univ., Lviv, 2024. [Online]. Available: https://lpnu.ua/sites/default/files/2024/radaphd/27733/disertaciya-cimbalyuk-rev1.pdf. Accessed on: Feb. 12, 2025.

[9] O.I. Yevdochenko, "Features of multi-position Reed-Solomon codes", in *Proc.10th All-Ukr. Sci. and Prac. Conf. Electronic and Mechatronic Systems: Theory, Innovation, Practice*, Poltava, pp. 106-107. [Online]. Available: https://nupp.edu.ua/uploads/files/0/events/conf/2024/x-emst/zbirnik-naukovikh-prac-X-vsekrainskoi-konferencii-ems_2.pdf. Accessed on: Jan. 10, 2025.

[10] S.V. Ilyin, E.L. Kholod, and A.B. Maznichko, "Organisation of long-term access to authentic documents", in *Proc. 2d Int. Sci. Pract. Conf. Creation of new ideas of learning in modern conditions*, Bordeaux, 2023, pp. 252-258. [Online]. Available: https://dspace.hnpu.edu.ua/server/api/core/bitstreams/a8627c6c-02af-400e-ab13-4ea4c073c4e3/content. Accessed on: Jan. 10, 2025.

SERHIY IVANCHENKO,
VADYM YAROSHCHUK

THE IMPACT OF REDUNDANCY ON THE SECURITY OF TECHNICAL INFORMATION LEAKAGE CHANNELS AND APPROXIMATE CORRECTION OF THE CHANNEL ERROR PROBABILITY

The article considers the impact of redundancy on the security of technical information leakage channels and the approximate correction of the error probability in the channel in the absence of information about the origin of redundancy. The mechanisms of introducing redundancy into signals are investigated, which contribute to improving noise immunity by correcting errors, but also create risks of reducing the confidentiality of transmitted information.

The article considers two main types of redundancy: artificial, which is formed through coding and control characters, and natural, which is a consequence of the peculiarities of the information source. The principles of its decoding differ significantly. This makes it impossible to directly apply error probability correction methods developed for artificial redundancy to natural redundancy. Artificial redundancy has a clear structure and is used to increase noise immunity, while natural redundancy arises from the characteristics of the information source, manifesting itself through correlations between symbols, repeated patterns and statistical regularities.

The article considers the features of speech and visual channels, where natural redundancy plays a key role in error correction. In speech channels, it helps to increase the reliability of information reception due to the subjectivity of pronunciation and perception, and in visual channels - due to the regularity of pixel distribution in images.

A new approach to decoding based on the ordering of code combinations by weight is proposed. The method of dividing possible 7-bit combinations into groups is used, which provides more efficient error correction. The traditional approach to error correction in the Hamming (7,4) code and its limitations at high levels of interference are also considered.

The results obtained can be used to improve methods of increasing the reliability of data transmission and reducing the risks of information leakage through technical channels. Importantly, this method not only improves error correction, but also opens up new opportunities for adaptive coding in complex information transmission conditions. Its versatility makes it possible to apply the approach in various fields, from digital communications to speech signal processing.

Keywords: redundancy, error probability, correction codes, information security, optimal reception, Hamming code, information and communication systems.

Іванченко Сергій Олександрович, доктор технічних наук, професор, професор кафедри безпеки державних інформаційних ресурсів, Інститут спеціального зв'язку та захисту інформації Національного технічного університету України "Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна, ORCID 0000-0003-1850-9596, soivanch@gmail.com.

Ярощук Вадим Дмитрович, курсант Інститут спеціального зв'язку та захисту інформації Національного технічного університету України "Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна, ORCID 0009-0002-5407-7911, vad1myar0shchuj@gmail.com.

Ivanchenko Sergii, doctor of technical sciences, professor, professor of the department of security of state information resources, Institute of special communication and information protection of National technical university of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine.

Yaroshchuk Vadym, cadet, Institute of special communication and information protection of National technical university of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine.