УДК 519.87:004.415.3 (045)

Юдін О.К., д-р. техн. наук
Чунарьова А.В.

МАТЕМАТИЧНІ АСПЕКТИ ВИКОРИСТАННЯ БАГАТОАЛЬТЕРНАТИВНИХ ПРАВИЛ В ЗАДАЧАХ КАНАЛЬНОГО КОДУВАННЯ ІНФОРМАЦІЙНИХ ПОТОКІВ

**Інститут інформаційно-діагностичних систем
Національного авіаційного університету**

Проведено аналіз багато альтернативних правил прийняття рішення. Побудована модель вирішальних процедур для визначення можливості використання багато альтернативних правил прийняття рішення у задачах канального кодування. Встановлені пороги прийняття рішення з метою вірогідного відновлення повної кодової конструкції

Вступ

Під цілісністю інформації розуміється її властивість, яка полягає в тому, що інформація не може бути модифікована неавторизованим користувачем або процесом. Іншими словами, під цілісністю інформації розуміється відсутність у ній будь-яких спотворень чи викривлень (модифікацій), які не санкціоновані власником інформації, незалежно від причин або джерел їх виникнення.

Спотворення інформації, тобто порушення цілісності, можливі на будь-якому етапі її циркуляції в інформаційно-комунікаційних мережах: при зберіганні, передачі або обробці.

Причини таких процесів можуть бути випадковими або навмисними. І ті, і інші дії мають своїм наслідком спотворення деякої кількості символів у цифровому поданні інформації (у інформаційній кодовій конструкції) і в цьому розумінні є загрозами функціональним властивостям захищеності інформаційних ресурсів – їхній цілісності та доступності.

Тому завдання забезпечити цілісність інформаційних ресурсів є одним із найактуальніших при розробці та експлуатації інформаційних систем та їх елементів. Ця актуальність підтверджується й вимогами, щодо загального часу на обробку та передачу даних, а також припустимої ймовірності помилок у повідомленнях, яку слід розуміти як: імовірність порушення цілісності кодових слів, які передаються по каналах зв'язку, або обробляються.

Постановка задачі

Задача завадостійкого канального кодування – забезпечення цілісності інформаційних повідомлень із застосуванням завадостійких коригувальних кодів. Завадостійким коригувальним кодуванням називається такий вид кодування, який дає змогу реалізовувати програмні, апаратні або програмно-апаратні засоби виявлення та усунення спотворень у інформаційних повідомленнях [1].

Задачею даної статті є побудова моделей вирішальних процедур статистичних правил для визначення можливості використання стандартних багато альтернативних методів прийняття рішення, а також проведення аналізу їх адекватності (визначення правил та відповідних мір кількості інформації, порогів прийняття рішення) для задач вірогідного декодування повної кодової конструкції інформаційного потоку даних.

Аналіз багато альтернативних методів прийняття рішення

В результаті проведеного аналізу та розробки нових алгоритмів, авторами виділенні основні вимоги та критерії, щодо формування сучасних методів згорточного завадостійкого декодування інформаційних потоків супутникових каналів зв'язку [2]:

− забезпечення скорочення часу на прийняття остаточного твердого рішення, щодо відновлення повної кодової послідовності;

− відмова від проходження всіх етапів решітки згорточного кодування на основі рішення багато альтернативних задач;

− зменшення надлишковості завадостійкого коду за рахунок підвищення вірогідності процедур декодування;

− використання послідовних правил прийняття рішень за рахунок накопичення достатньої кількості інформації на базі більш інформативних параметрів сигналу;

− можливість зупинки процесу декодування у разі відповідності функції правдоподібності встановленим порогам з урахуванням достатньої кількості інформації (Вальдовський підхід).

Запропоновані авторами, нові методи оптимального декодування інформаційних потоків були розроблені на основі вирішення багато альтернативної задачі з урахуванням послідовних правил прийняття рішень. Однак, данні методи сформовані на базі інформативності параметрів кодових конструкцій з використанням традиційних, статистичних правил прийняття вірогідного рішення відносно тієї, чи іншої гіпотези.

Данні критерії, характеризуються розрахунковою мірою мінімальної кількості інформації та на її основі сформованими порогами прийняття рішення, відносно представлених гіпотез. Апробація та впровадження зазначених методів, зазвичай використовувалась для вирішення задач радіотехніки, радіолокації і навігації з кількістю гіпотез, що не перевищувала $N = 3\text{-}10$ для інформаційних об'єктів різних класів.

У випадку вирішення задач згорточного декодування на основі багато альтернативних правил, процедури прийняття рішень повинні обробити інформаційні складові кодових конструкцій, можлива комбінація котрих складає 512 і більше Тобто, вирішити багато альтернативну задачу при кількості гіпотез $N = 512\text{-}1024$ з урахуванням вірогідної можливості повтору кодових конструкцій або появи кодового слова з близько схожими інформаційними параметрами.

Таким чином, необхідно провести аналіз та побудувати моделі статистичних правил для обґрунтування можливості використання стандартних багато альтернативних методів прийняття рішення та визначення їх адекватності (правил та відповідних мір кількості інформації, вірогідних порогів прийняття рішення) для задач завадостійкого декодування повної кодової конструкції інформаційного потоку при постійнім збільшенні кількості альтернативних гіпотез.

Розглянемо можливість побудови математичних моделей процедур оптимального декодування інформаційних потоків на основі вирішення багато альтернативної задач прийняття рішень. Данні моделі повинні бути побудовані з урахуванням забезпечення скорочення часу на прийняття остаточного твердого рішення, щодо відновлення повної кодової послідовності на базі інформативності параметрів.

В сучасній статистичній теорії для вирішення багато альтернативних задачах, відомі наступні математичні апарати: Котельникова, Байеса, Шеннона, Фішера, Кульбака. Данні методи надають можливість використання критеріїв й процедур ідентифікації інформаційних сигналів на базі мінімально достатньої кількості інформації $I_k^{(x)}$ з визначеними порогами прийняття рішень $V_k^{(x)}$ [3, 1] (табл. 1). В даному випадку з *табл.1* видно, що поріг прийняття рішення $V_k^{(1)}$ для кількості гіпотез від 3 до 10, змінюється в межах 0.67 – 0.2 [3]. Однак, данні правила розроблені для ідентифікації сигналів сформованих від різних класів інформаційних об'єктів з різко відмінними інформаційними параметрами (боїнг чи спортивний літак) та мають конкретні недоліки:

− по-перше, у визначених правилах не беруться до уваги сцени де можлива ситуація появи двох або взагалі десяти однакових гіпотез одночасно, а також інформаційних об'єктів з схожими параметрами (слабо відмінні);

– по-друге, визначені статистичні пороги, що сформовані на базі мінімально-достатньої кількості інформації, при збільшенні кількості альтернативних гіпотез ($N = 50\text{-}1000$) втрачають фізичний зміст та практично дорівнюють нулю.

Таблиця 1. Класичні пороги прийняття рішення при кількості гіпотез $N = 3\text{-}10$

Класичні пороги прийняття рішення \ Кількість гіпотез		3	4	5	6	7	8	9	10
Котельникова, Байеса	$V_k^{(1)} = 1 - I_k^{(1)} = = 1 - \frac{N-2}{N} = \frac{2}{N}$	0.6	0.5	0.4	0.3	0.28	0.25	0.2	0.2
Шеннона	$V_{sh}^{(3)} = 1 - I_{sh}^{(3)} = = 1 - \ln(N-1) = = \ln\left(\frac{e}{N-1}\right)$	0.3	0.1	0.4	0.6	0.7	0.9	1.1	1.2
Фішера	$V_f^{(4)} = 1 - I_f^{(4)} = = 1 - \ln\left(\frac{1}{N-1}\right) \cdot \ln\left(\frac{N-1}{N^2}\right) = = \ln(e) + \ln^2(N-1) - - 2\ln N \cdot \ln(N-1)$	0.04	0.8	1.5	2.1	2.7	3.3	3.8	4.3
Кульбака	$V_{kl}^{(5)} = 1 - I_{kl}^{(5)} = = 1 - 2\ln(N-1) = = \ln\left(\frac{e}{(N-1)^2}\right)$	0.38	1.2	1.7	2.2	2.5	2.8	3.1	3.3

Таким чином, класичні багато альтернативні правила прийняття рішення не адекватні у разі застосування для задач, що вирішуються з метою ідентифікації (відновлення) повної кодової конструкції з кількістю гіпотез більше 10, а також визначенні пороги прийняття рішення статистично завишенні і недосяжні в ситуаціях коли сцени ідентифікації інформаційних сигналів мають однакові чи схожі параметри (це стосується не тільки задач декодування кодових комбінацій, а взагалі всіх радіотехнічних задач).

Розробка методів відновлення кодових конструкцій та побудова моделей процедур статистичних правил

Задача багато альтернативності в нашому випадку носить складний характер.

Так, як багато альтернативністю гіпотез являється набір кодових конструкцій, які в загальному випадку характеризуються технічними можливостями інформаційно-телекомунікаційних систем та мереж (ІКСМ) й можливістю цифрового представлення сигналу ($2^0, 2^1, 2^2, 2^3, 2^4, 2^5 \ldots$). Під кодовими конструкціями будемо розуміти інформаційні сигнали ІКСМ, представлені кодовим словом сформованим на базі методів завадостійкого канального кодування у вигляді послідовності радіо або відео імпульсів (рис. 1). Частіше всього цифрові сигнали представляються як 16-32 бітові архітектури, а кількість гіпотез при цьому зростає до $N \to 2048$.

При даній постановці задачі, зростає ймовірність появи інформаційних сигналів з ідентичними параметрами. В цьому випадку вірогідна поява переданих по каналу зв'язку кодових слів, що мають слабо відмінні параметри конструкцій, тобто відрізняються на нуль, одну, дві чи три Хемінгові відстані (співпадаючи або на 1-3 біта різниці). Тобто, слабо відмінні чи «близько стоячі» інформаційно-альтернативні гіпотези.

В даному випадку зазначені пороги прийняття рішення по критеріям Котельникова, Байеса, Шеннона, Фішера, Кульбака статистично не досяжні та їх процедури стають не робочими.

Існуючі класичні правила прийняття рішення при реалізації послідовних багато альтернативних задач, дуже важливі, але не спроможні до розв'язання поставленої задачі. Розглянуті класичні правила прийняття рішення в свою чергу не забезпечують мінімально достатньої кількості інформації при прийнятті остаточного рішення у розв'язанні задач ідентифікації

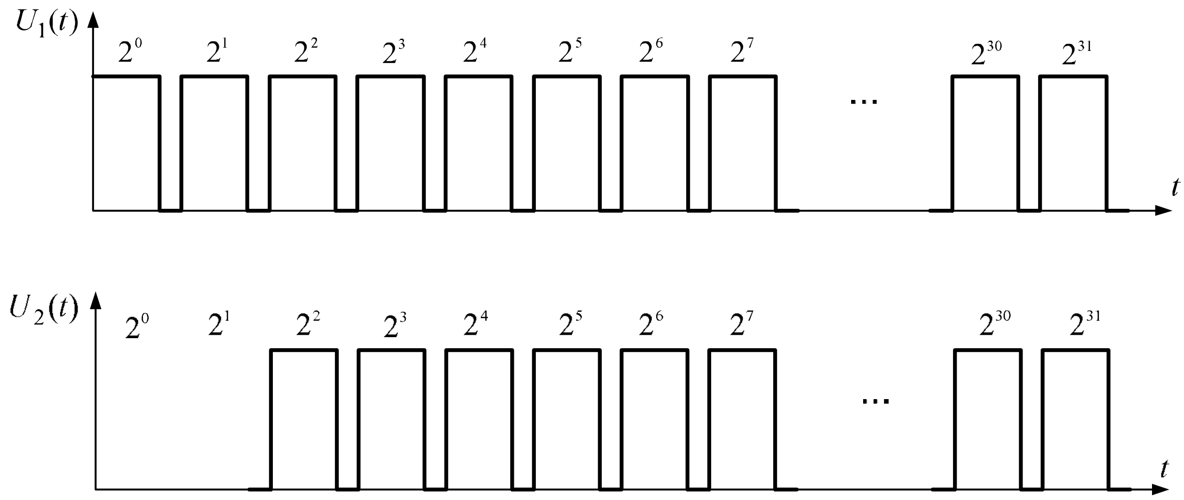

Рис. 1. Слабо відмінні кодові конструкції

слабо відмінних кодових конструкцій. Тобто потрібно щоб поріг прийняття рішення зменшувався, був таким, що забезпечить ідентифікацію та прийняття рішення про відновлення повної кодової конструкції у складній ситуації.

Вирішення встановленого протиріччя, між кількістю слабо відмінних гіпотез та величиною статистичного порогу прийняття рішення, можливо з урахуванням того, що поява всіх однакових гіпотез N одночасно: ймовірна, тобто одночасно можлива поява однієї й тієї ж кодової конструкції (граничний випадок $N = 10$).

В даному випадку при присутності під множини n, співпадаючих або слабо відмінних кодових слів, значення статистичного порога прийняття рішення не може перевищувати $V_{\min} = 1/n$. Приведений поріг є дійсно мінімальним, що характеризує мінімально-достатню кількість інформації про кодове слово.

Таким чином, можна зробити висновок, що для вирішення багато альтернативної задачі на базі мінімально – достатньої кількості інформації при відновленні повного кодового слова, існують два статистичних порога прийняття рішень:

– нижній поріг прийняття рішення, що дорівнює $V_{\min} = 1/n$ при появі в множині N кодових слів n, що мають слабо відмінні параметри, тобто кодових конструкцій, які мають мінімальну Хемінгову відстань до 4 біт в ближніх позиціях або повністю співпадають;

− верхній поріг прийняття рішення дорівнює $V_{\max}$ та встановлений згідно стандартних статистичних правил для інформаційних сигналів з різко відмінними параметрами (табл. 1).

Отже, надалі ми будемо говорити про двох порогові процедури прийняття рішення для слабо та сильно відмінних гіпотез. Загальний вид залежності статистичних порогів прийняття рішення від кількості альтернативних гіпотез $N = 3 - 10$ має вид рис. 3.

Вирішення даного питання, дозволяє нам перейти до наступного та найбільш важливого протиріччя, між зростаючою кількістю альтернативних гіпотез та величиною статистичного порогу прийняття рішення.

Вирішення даної задачі можливе на основі досягнення мінімально достатньої міри кількості інформації, що сформована з урахуванням найбільш інформативних параметрів інформаційних сигналів.

Розробимо та впровадимо статистичні моделі процедур вирішення складних багато альтернативних задач з урахуванням присутності альтернативних гіпотез з слабо відмінними параметрами, а також на базі математичного моделювання сформуємо графіки залежності статистичних порогів прийняття рішень від зростаючої кількості гіпотез.

Під інформативними параметрами сигналу будемо розуміти спектральне представлення послідовності кодових слів. Даний вид представлення параметрів сигналу є найбільш інформативним для формування порогів прийняття рішення на базі мінімально-достатньої кількості інформації. Для спектрального представлення сигналів використовуємо пряме перетворення Фур'є, тобто даний сигнал буде представлятися у вигляді суми гармонічних коливань з різними частотами.[4, 2].

Отже, за найбільш інформативний параметр інформаційного сигналу вибираємо не дискретні часові відліки сигналу, а дискретні значення енергетичних складових спектрів відповідних кодових конструкцій. Спектри сигналів зображені на рис. 2, відповідають 32 бітовим кодовим конструкціям з слабо відмінними параметрами.

Оскільки далі спектральне подання сигналу буде використатися при розрахунках умовних ймовірностей і виборі найбільш ймовірної гіпотези, то можна говорити про кількість інформації, яка відповідає кожній гіпотезі, позначене як: $I_k(x)$ у відповідності до $\{H_k\}$, де $k = 1,...,N$ [4, 2]. Тобто для кожної гіпотези $\{H_k\}$ обчислюється кількість інформації $I_k(x)$, що міститься в спектрі. В цьому разі позначимо $S_i(x_j)$ ($j = 1,...,N$, $i = 1,...,N$), як спектральне представлення прийнятого інформаційного сигналу при наявності в каналі зв'язку адитивного білого гауссового шуму.

Міра кількості інформації $I_k(x)$, яка розраховується для кожної гіпотези $\{H_k\}$, має наступний вид:

$$I_k(x) = 1 - \Psi(q_k(x)), \qquad (1)$$

де $q_k(x)$ – апостеріорна ймовірність появи інформаційного сигналу; $\Psi(q_k(x))$ – узагальнена міра невизначеності для k-го інформаційного сигналу, розрахована по апостеріорним ймовірностям.

Враховуючи, що за інформативний параметр інформаційного сигналу використовується спектральне представлення та формула (1) прийме вид:

$$I_k\left[P\left(H_k/S_i(x_j)\right)\right] = 1 - \Psi\left[P\left(H_k/S_i(x_j)\right)\right], \qquad (2)$$

де $P\left(H_k/S_i(x_j)\right)$ – апостеріорна ймовірність появи інформаційного сигналу;
$\Psi\left[P\left(H_k/S_i(x_j)\right)\right]$ – узагальнена міра невизначеності розрахована по апостеріорній ймовірності:

$$\Psi\left[P\left(H_k/S_i(x_j)\right)\right] = \varphi\left(\frac{f\left[P\left(H_k/S_i(x_j)\right)\right]}{P\left(H_k/S_i(x_j)\right)}\right).$$

У результаті того, що ми використали, як інформативний параметр спектральне представлення сигналу, то розрахунок умовної щільності ймовірності розподілу знайдемо по формулі:

$$W\left[S_i(x_j)/H_k\right]=\prod_{j=1}^{N}\frac{1}{\sqrt{2\pi\sigma_k}}e^{-\frac{(S_k(x_j)-m_{ij})^2}{2\sigma_k^2}},\qquad(3)$$

де $S_i(x_j)$ – прийнятий сигнал, за інформативний параметр якого взято спектральне представлення, j – поточний номер спектральних складових кожного $S_i(x)$ прийнятого сигналу ($j=1...N$), i – поточне значення номеру інформаційного сигналу для множини гіпотез $i=1,...,N$ ($k=1,...,N$), m_{ij} – математичне очікування.

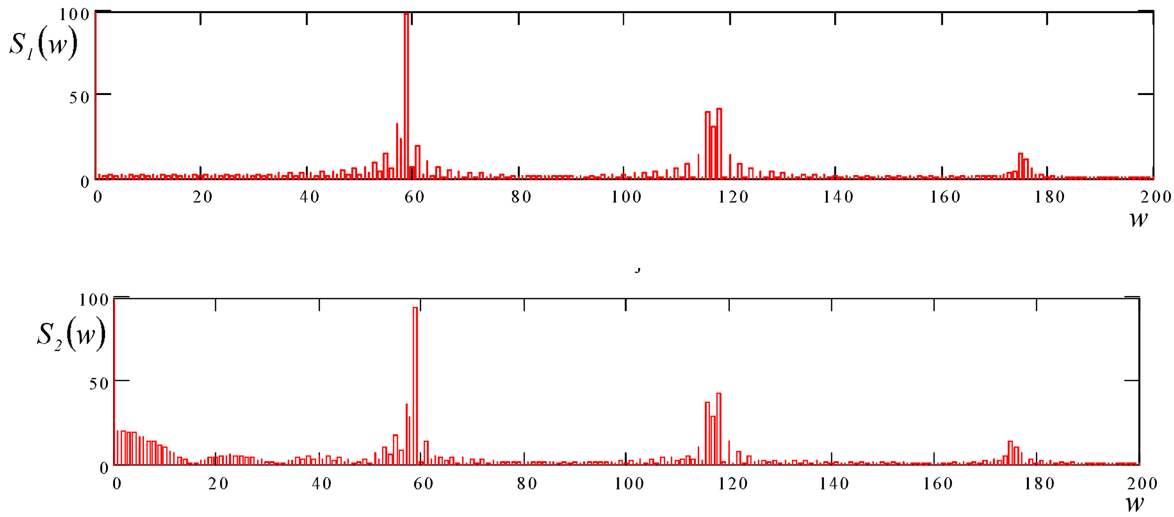

Рис. 2. Спектральне представлення слабо відмінних кодових конструкцій

З огляду на вище викладену формулу (3), одержимо загальний Байесовський вираз знаходження умовної ймовірності появи інформаційного сигналу, тобто відповідної гіпотези $\{H_k\}$ при умові прийнятого інформаційного сигналу $S_i(x_j)$:

$$P\left[H_k/S_i(x_j)\right]=\left.\frac{\prod_{j=1}^{N}\frac{1}{\sqrt{2\pi\sigma_k}}e^{-\frac{(S_k(x_j)-m_{ij})^2}{2\sigma_k^2}}}{\sum_{j=1}^{N}\prod_{j=1}^{N}\frac{1}{\sqrt{2\pi\sigma_i}}e^{-\frac{(S_i(x_j)-m_{ij})^2}{2\sigma_i^2}}}\right|_{\max},\qquad(4)$$

Для прийняття рішення про ідентифікацію прийнятої послідовності $S_i(x_j)$, можна використати наступну процедуру прийняття рішення:

$$I_k\left[P\left(H_k/S_i(x_j)\right)\right]>V_{\max},\qquad(5)$$

ідентифікація повного кодового слова з урахуванням присутності в кодових послідовностях N «слабо відмінних» кодових слів.

Нище представлений графік залежності $V_{\max}$ від кількості альтернативних гіпотез (рис. 3).

На базі впроваджених методів математичного моделювання та ґрунтуючись на проведеному аналізі встановленої закономірності, можна зробити висновок: що при зростанні кількості гіпотез більше $N>10$ та появі в множині N кодових конструкцій: n слабо відмінних інформаційних сигналів, значення статистичного порогу $V_{\max}$ для процедури прийняття рішення стає фіксованим на відповідному рівні, а саме $V_{\min}=0.25$. Статистичні пороги стають фіксованими на певному рівні при появі в множині слабо відмінних кодових конструкцій та зростанні кількості альтернативних гіпотез

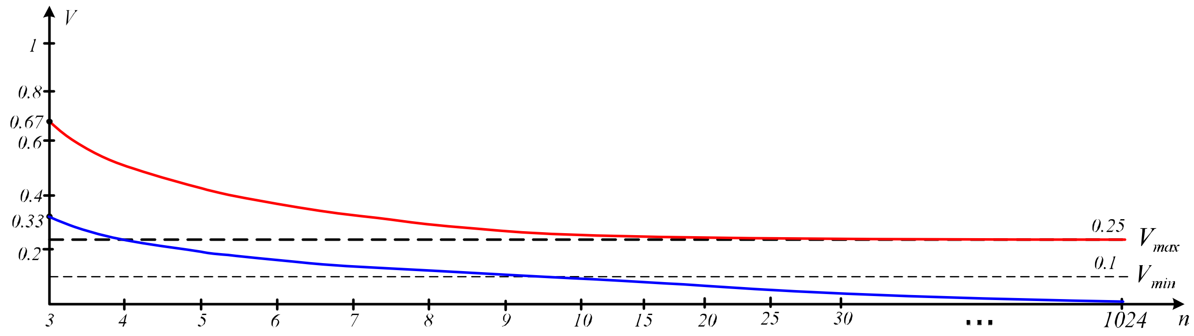

Рис. 3. Двох порогова процедура прийняття рішення при збільшені кількості гіпотез

(зазначені пороги прийняття рішення по критеріям Котельникова, Байеса, Шеннона, Фішера, Кульбака повинні спадати при збільшення кількості гіпотез, та врешті втрачають фізичний смисл).

В даному випадку, під поняття «слабо відмінні» кодові конструкції підпадають сигнали, що мають мінімальну Хеммінгову відстань різниці у словах – не більше 4 біт в ближніх позиціях: $d_{ij} \approx 4$. Тобто, при появі в множині інформаційних сигналів, таких що мають різницю не більше 4 біт, Байесовське правило прийняття рішення ідентифікує їх, як слабо відмінні гіпотези і фіксує поріг прийняття рішення на рівні $V_{\min} = 1/n \approx 0.25$.

Отже, правило Байеса спрацьовує, фіксоване значення умовної ймовірності порогів дорівнює $0.3 - 0.32$, при відсутності шумової складової та на рівні $V_{\min} = 1/4 \approx 0.25$ при співвідношенні *сигнал\шум=3\2* .

Висновки

Проведено аналіз багато альтернативних методів прийняття рішення.

Встановлено, що класичні правила та пороги прийняття рішення статистично завищені та неадекватні при розв'язання багато альтернативної задачі з урахуванням появи в множині слабо відмінних кодових конструкцій та збільшенні кількості альтернативних гіпотез.

Показано, що при збільшення кількості гіпотез $N > 10$, значення порогу прийняття рішення становить фіксоване значення.

Побудована математична модель вирішальних процедур статистичних правил для визначення можливості використання багато альтернативних методів прийняття рішення у задачах декодування повної кодової конструкції.

Відповідно до цього, проведено аналіз та визначені правила і міри кількості інформації, а також встановлені пороги прийняття рішення з метою вірогідного відновлення повної кодової конструкції інформаційного потоку даних.

Список літератури

1. *Скляр Б.* Цифровая связь. Теоретические основы и практическое применение. 2-е издание. Пер. с англ. – М.: Издательский дом “Вильямс”, 2003. –1104 с.

2. *Юдін О.К.* Кодування в інформаційно-комунікаційних мережах – Монографія. – К.: Книжкове видавництво НАУ, 2007. – 302 с.

3. *Косенко Г.Г.* Метод последовательного расширения областей принятия решений в задачах распознавания // Радиоелектроника. – М. – 1980. – №27 – С. 72–75.

4. *Юдін О.К.* Алгоритми ідентифікації сигналів та керуючих повідомлень у телекомунікаційних системах і мережах // Зб. наук. праць інституту проблем моделювання в енергетиці ім. Г.Є. Пухова «Моделювання та інформаційні технології». – №37. – К.: ІПМЕ – 2006. – С. 9–16.