

Ю.Г. Войлов, д.т.н.
Східноукраїнський державний університет
Є.М. Литвинов, к.т.н., доц.
О.С. Ліхоузов, аспір.
Т.А. Ліхоузова, аспір.
Національний технічний університет України "КПІ"

**МЕТОД СКОРОЧЕННЯ ОБ'ЄМУ РОБОЧОЇ БІБЛІОТЕКИ ІДЕНТИФІКАЦІЇ
 СПЕКТРАЛЬНОГО СКЛАДУ ВИПРОМІНЮВАННЯ**

Запропонована робота пов'язана з проблемою ідентифікації радіонуклідів за відсканованим спектром. Запропоновано метод оптимізації робочої бібліотеки гамма-спектрів для вирішення конкретних задач. Для оцінки якості мінімізованої бібліотеки використано інформаційний критерій.

Проблема дистанційного контролю просторової дислокації джерел у задачі радіологічного моніторингу є основною при будь-якій екстремальній ситуації, пов'язаній з викидом радіонуклідів. Своєчасне виявлення ділянок забруднення та адекватна оцінка характеру цього забруднення, яка виконана дистанційно й за короткий час, дозволяє значно зменшити дози опромінення персоналу служби радіаційної розвідки та знизити ризик опромінення людей, що потрапили в зону забруднення [1]. Актуальність вирішення подібної задачі висока ще й через сплеск активності міжнародного тероризму й імовірності здійснення терористичних актів на ядерних технологічних установках.

Основними факторами, що характеризують середовище з джерелами іонізуючого випромінювання, є щільність потоку гамма-випромінювання й його спектральний склад у кожній точці простору. Радіологічне дистанційне дослідження доставляє інформацію про досліджуване середовище в кінцевому числі точок спостережень у заданому енергетичному діапазоні, утворюючи вихідний масив даних. Надалі ця інформація зберігається в пам'яті і може бути використана для відновлення спектра кожного джерела, локалізованого в полі зору в будь-якому обраному енергетичному вікні.

Нехай у кожній конкретній пошуковій задачі за мету пошуків обрано кінцеве число різних типів об'єктів, кожному з яких відповідає визначений спектр гамма-випромінювання. Тоді ціннісний опис можна представити у вигляді множини точок спостережень із заданим на ній відношенням еквівалентності, тому кожен клас точок, крім одного, відповідає визначеному типу (класу) цільових об'єктів. Один клас складають точки, у які не попадає жоден з цільових об'єктів. Уточнимо тепер постановку задачі.

Нехай M – кінцева множина точок спостережень, упорядкованих у просторі і пронумерованих, та нехай $X = \{X^m, m = \overline{1, L}\}$, множина векторів значень параметрів, що спостерігаються, визначена на M (m – номер точки L , – число точок у M).

Завдання полягає в тому, щоб по заданому X і деякій апріорній інформації про зв'язок між множиною X і кожним з k типів пошукових об'єктів розділити множину M на $k + 1$ клас так, щоб j -й клас складався з точок, «перспективних» для пошуку об'єктів j -го типу ($j \leq k$). Передбачається, що клас з номером $k + 1$ складається з «безперспективних точок».

Існує безліч методів рішення цієї задачі [1]. Таке застосування припускає наявність базової бібліотеки гамма-спектрів, що формується в контрольованих умовах на зразкових ізотопах або моделюється. Розмір базової бібліотеки являє собою кінцеву множину, а склад робочої бібліотеки визначається умовами конкретної задачі.

При розв'язанні задачі оперативної класифікації радіоактивного забруднення характерне проведення глобального моніторингу з метою виявлення й ідентифікації радіонуклідів за дуже обмежений час, максимальне значення якого визначається умовами радіаційної безпеки при роботі в зоні забруднення. Як наслідок, немаловажну роль грають обсяг робочої бібліотеки і число порівнюваних параметрів.

Використовуючи методику й алгоритми, запропоновані в [2, 3], можна з мінімальними втратами, в точності, істотно скоротити число порівнюваних параметрів для кожної конкретної робочої бібліотеки за рахунок порівняння характеру відносин між значеннями інтенсивності випромінювання на різних ділянках спектра.

Нехай вектор значень параметрів, що спостерігаються, складається з n змінних (тобто частин енергетичного діапазону):

$$X = \{X_i\}, i = \overline{1, n} \tag{1}$$

з Q_i рівнями квантування для кожної змінної, і обсяг робочої бібліотеки дорівнює N . Тоді можна скласти матрицю даних:

$$D = X \times N \tag{2}$$

На першому етапі мінімізації робочої бібліотеки проводиться квантування неперервних даних або угруповання дискретних у менше число рівнів, а на другому етапі – заміна груп найбільш пов’язаних (за критерієм взаємної інформації) змінних похідними змінними за допомогою алгоритмів [3].

У даній постановці задачі для першого етапу найбільш оптимальним буде метод рівноімовірного квантування [4].

Як критерій для виділення груп найбільш пов’язаних змінних будемо використовувати визначення взаємної інформації [5]:

$$I(X_i, X_j) = \sum_{(X_i, X_j)} p(X_i, X_j) \log_2 \frac{p(X_i, X_j)}{p(X_i)p(X_j)} \tag{3}$$

Процедура виділення груп найбільш пов’язаних змінних має вигляд. Знаходиться пара змінних (X_i, X_j) , для яких $I(X_i, X_j)$ є максимальною серед усіх значень взаємної інформації, розрахованих для пар змінних в X . Змінна X_i повинна залежати від X_j з визначеним рівнем значимості, що може бути перевірене на підставі χ^2 -критерію. На підставі того, що $1,3863NI(X_i, X_j)$ асимптотично розподілено як χ^2 [4], значимість цієї величини може бути оцінена за таблицями χ^2 -розподілу для ступеня волі df , що може бути визначена у вигляді:

$$df = [Q_i - 1] \times [Q_j - 1] \tag{4}$$

Якщо величина $I(X_i, X_j)$ значима, то змінні (X_i, X_j) можна об’єднати й одержати нову систему з $n - 1$ змінною. Далі весь процес повторюється з новою системою. Якщо в результаті проведених обчислень не буде знайдено значимих величин $I(X_i, X_j)$, це говорить про те, що процес виділення груп взаємопов’язаних змінних закінчений і отримана похідна система з $m \leq n$ змінними.

У тому випадку, якщо число m теж занадто велике, можна виключити з розгляду найменш значимі для визначення структури системи змінні. Це можна зробити на підставі величини взаємної інформації:

$$I(X_i, X - X_i) = \sum_{(X_i, X - X_i)} p(X_i, X - X_i) \log_2 \frac{p(X_i, X - X_i)}{p(X_i)p(X - X_i)}, \quad i = \overline{1, m} \tag{5}$$

Задавши рівень значимості величини взаємної інформації, знаходимо за допомогою χ^2 -критерію її граничне значення і відкидаємо незначущі змінні. Ступінь волі df для χ^2 -розподілу може бути визначено у вигляді:

$$df = \prod_X [Q_i - 1] \tag{6}$$

У результаті одержимо систему із $k \leq m \leq n$ змінними.

Поєднуючи вищесказане, можна сформулювати наступну методику оптимізації робочої бібліотеки:

1. Формування вихідної робочої бібліотеки на підставі апріорної інформації.
2. Квантування неперервних даних або групування дискретних у менше число рівнів.
3. Виділення і заміна груп найбільш зв’язаних (за критерієм взаємної інформації) змінних похідними змінними.

Як міру для порівняння вимірюного спектра зі спектрами робочої бібліотеки використовували інформаційну відстань, яку можна обчислити різними способами [3, 6], наприклад, із формули:

$$G({}^1f, {}^2f) = D\left({}^1f, \frac{{}^1f + {}^2f}{2}\right) + D\left({}^2f, \frac{{}^1f + {}^2f}{2}\right), \tag{7}$$

де 1f і 2f – довільні розподіли імовірностей, визначені на тій самій множині станів; $({}^1f + {}^2f)/2$ – розподіл імовірностей, отриманий узяттям середнього для кожної пари відповідних імовірностей з 1f і 2f ; D – інформаційна відстань, задана рівнянням:

$$D(f, f^h) = \frac{1}{\log_2 |C|} \sum_{c \in C} f(c) \log_2 \frac{f(c)}{f^h(c)}, \tag{8}$$

де константа $1/\log_2 |C|$ – нормуючий коефіцієнт, що забезпечує виконання співвідношення

$$0 \leq D(f, f^h) \leq 1. \tag{9}$$

Для прикладу наведемо залежності інформаційної відстані й обсягу робочої бібліотеки від кількості рівнів квантування і рівня значимості взаємної інформації при різній масі і формі матеріалу. Вихідними

данями є спектри гамма-випромінювання зразків плутонію різної маси і форми, різного ізотопного та хімічного складу ([1], додаток 3 і 4).

На рис. 1 наведена залежність інформаційної відстані від рівня значимості взаємної інформації (кількість рівнів квантування даних вимірів 4, зразки SRM-948 метал і Pu2 повторні цикли, маса зразків 100 г кульової форми, у робочій бібліотеці зразки форми «товста пластина»).

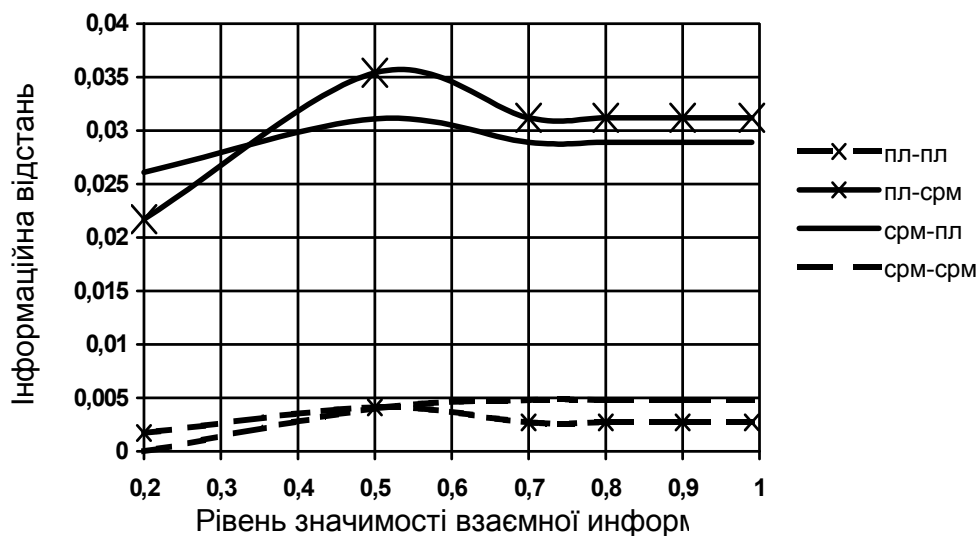


Рис. 1

На рис. 2 наведена залежність інформаційної відстані від кількості рівнів квантування (рівень значимості 0,99; маса зразків 100 г кульової форми). Виконані дослідження з оцінки впливу вибору числа рівнів квантування, результати яких добре погодяться з результатами, отриманими в [4, 7], показують, що при двійковому квантуванні спостерігається їхня нестійкість. Погіршення результатів при п'ятирівневому квантуванні пов'язано з порушенням обмеження даних [4].

На рис. 3 наведена залежність інформаційної відстані від рівня значимості при різній масі зразків (кількість рівнів квантування 4, маса зразків 100 г і 1 кг кульової форми).

На рис. 4 наведена залежність обсягу робочої бібліотеки від рівня значимості та кількості рівнів квантування.

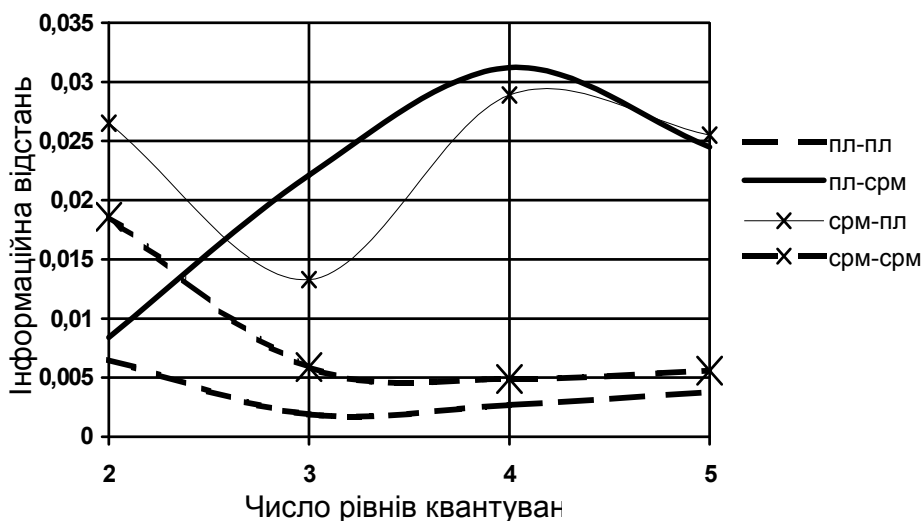


Рис. 2

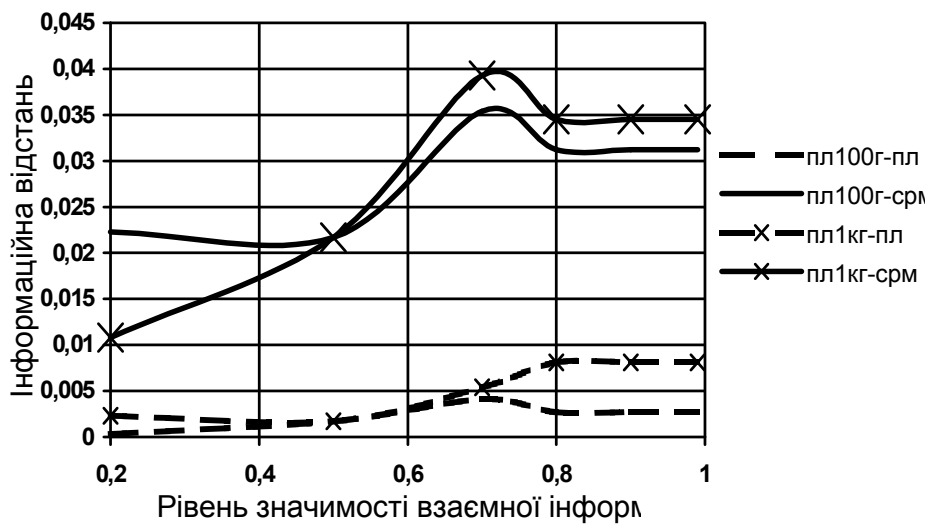


Рис. 3

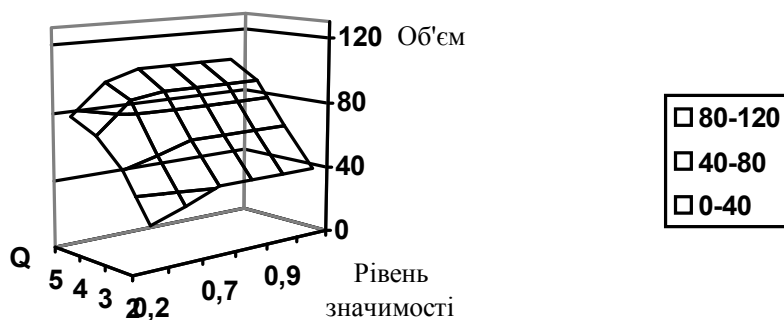


Рис. 4

Як видно з графіків, при рівні значимості взаємної інформації 0,5 обсяг робочої бібліотеки скорочується більш, ніж у два рази і при цьому спектри матеріалів, що ідентифікуються, добре розпізнаються.

ΛΙΤΕΡΑΤΥΡΑ:

1. *Войлов Ю.Г.* Системы гамма-видения. – Луганск: ВУГУ, 1999.
2. *Литвинов Е.Н., Лихоузов А.С., Лихоузова Т.А.* Методика идентификации многосвязных систем // Проблемы создания новых машин и технологий: сборник научных трудов КГПИ. Вып. № 1. – Кривой Рог, 1999.
3. *Литвинов Е.Н., Лихоузов А.С., Лихоузова Т.А.* Алгоритм идентификации сложных динамических систем // “Проблемы создания новых машин и технологий”: сборник научных трудов КГПИ, № 1. – Кривой Рог, 1999.
4. *Литвинов Е. Н., Лихоузов А. С., Лихоузова Т. А.* Модель системы данных в анализе реконструируемости сложных динамических систем // Адаптивные системы автоматического управления, Днепропетровск, 1998.
5. *Шеннон К.* Работы по теории информации и кибернетике. – М.: Издательство иностранной литературы, 1963.
6. *Клар Дж.* Системология: автоматизация решения системных задач. – М.: Радио и связь, 1990.
7. *Conant R, Valdes-Perez,* Information Loss Due to Data Quantization in Reconstructability Analysis, International Journal of General Systems, 1983. – Vol. 9. – P. 235–247.

ВОЙЛОВ Юрій Григорович – доктор технічних наук Східноукраїнського державного університету.

Наукові інтереси:

– ідентифікація складних динамічних систем.

ЛИТВИНОВ Євген Миколайович – кандидат технічних наук, доцент Національного технічного університету України “КПІ”.

Наукові інтереси:

– ідентифікація складних динамічних систем.

E-mail: litvin@tk.ntu-kpi.kiev.ua

ЛИХОУЗОВ Олексій Сергійович – аспірант Національного технічного університету України “КПІ”.

Наукові інтереси:

– ідентифікація складних динамічних систем.

ЛИХОУЗОВА Тетяна Анатоліївна – аспірантка Національного технічного університету України “КПІ”.

Наукові інтереси:

– ідентифікація складних динамічних систем.

Подано 25.09.2000

До рис. 1:

Інформаційна відстань
Рівень значимості взаємної інформації

До рис. 2:

Інформаційна відстань
Число рівнів квантування

До рис. 3:

Інформаційна відстань
Рівень значимості взаємної інформації

До рис. 4:

Об'єм
Рівень значимості