



НОВІ ЗАСОБИ КІБЕРНЕТИКИ, ІНФОРМАТИКИ, ОБЧИСЛЮВАЛЬНОЇ ТЕХНІКИ ТА СИСТЕМНОГО АНАЛІЗУ

УДК 303.444

В.С. КОМАРОВ

Науково-дослідний інститут військової розвідки, Київ, Україна, e-mail: komarvlad@ukr.net.

О.А. ІЛЛЯШОВ

Науково-дослідний інститут військової розвідки, Київ, Україна,
e-mail: aleksandr.ilyashov@gmail.com.

В.В. ОЛЕКСІЮК

Науково-дослідний інститут військової розвідки, Київ, Україна, e-mail: voleksiyk@ukr.net.

МЕТОДИЧНИЙ ПІДХІД ДО ІДЕНТИФІКАЦІЇ ОБ'ЄКТІВ МОНІТОРИНГУ

Анотація. За результатами проведених теоретичних та прикладних досліджень розроблено методичний підхід до ідентифікації об'єктів моніторингу. Його практичне застосування дає змогу розрахувати коригувальні коефіцієнти інтегрального критерію визначення важливості та інформативності моніторингових ознак і визначити найбільш ефективний варіант комплексу технічних засобів моніторингу відповідних джерел (об'єктів) моніторингу.

Ключові слова: моніторингова ознака, технічний засіб моніторингу, сигнатура, пріоритетність моніторингових ознак.

ВСТУП

Загалом одним із найважливіших завдань моніторингу єдиного інформаційного простору є виявлення, розпізнавання та класифікація джерел та об'єктів моніторингу (ДОМ). При цьому під час функціонування моніторингова структура передусім має застосовувати найефективніші методи здобування та оброблення інформації, а саме структурно-системний (ССМ) та сигнатурно-системний (СГСМ) методи [1]. У них використовується множина моніторингових ознак (МО) і сигнатур — сукупність доступних для виявлення, аналізу та оцінювання кількісних і якісних параметрів ДОМ. Головною особливістю моніторингового процесу є те, що він здійснюється в умовах часткової або повної невизначеності вихідних даних про ДОМ.

Під час розпізнавання інформація обробляється відповідними методами і приймається рішення про належність ДОМ тому чи іншому класу або про його фазовий стан [1–4]. Зрозуміло, що чим меншою є похибка апостеріорного параметра МО та чим він є більш наближеним до еталона, тим меншим є ступінь невизначеності та більшою є ймовірність правильного розпізнавання ДОМ. Якість розпізнавання і класифікації ДОМ залежить від двох основних факторів. Першим фактором є наявність похибок вимірювання, обчислення та оцінювання параметрів ДОМ, що визначаються потенційними і реальними можливостями апаратури технічних засобів (систем і комплексів) моніторингу (ТЗМ). Другим фактором є неповнота МО і сигнатури, що відображають стан або належність ДОМ, і поверховий опис властивостей відповідних об'єктів.

© В.С. Комаров, О.А. Ілляшов, В.В. Олексіюк, 2022

У зв'язку з цим виникає потреба у забезпеченні максимальної величини ймовірності правильного розпізнавання ДОМ. Автори виконали теоретичні та прикладні дослідження з відповідними розрахунками, що надають змогу побудувати методичний підхід до розв'язання такої задачі, який є предметом розгляду у цій роботі.

АНАЛІЗ ЛІТЕРАТУРИ ТА ПОСТАНОВКА ЗАДАЧІ

Постійне вивчення та удосконалення процесу виявлення, розпізнавання і класифікації ДОМ є предметом багатьох наукових досліджень і дискусій. Воно є безумовно важливим теоретичним етапом пошуку технічних рішень для побудови систем і комплексів моніторингу з відповідними технічними засобами та програмним забезпеченням, що ґрунтуються на перевірених методах. Починаючи з робіт з теорії розпізнавання та класифікації за прецедентами, а також шляхом застосування розділювальних функцій до задачі класифікації Фішера та розв'язання задачі перевірки гіпотез Вальда, науковці здійснювали поступовий перехід до нейромережових моделей розпізнавання (перцептрон Розенблата), статистичних моделей тощо. Результати аналізу наявного науково-методичного апарату, а саме методів потенціальних функцій (М.А. Айзерман), «узагальненого портрета» (В.М. Вапник та А.Я. Червоненкіс), комітетів (В.Д. Мазуров), групового урахування аргументів (О.Г. Івахненко), алгоритмів таксономії та аналізу знань (М.Г. Загоруйко), логічних методів розпізнавання та пошуку залежностей (Г.С. Лбов), тестового алгоритму розпізнавання за обмеженої кількості навчальних прецедентів (Ю.І. Журавльов), загальної теорії проблемно-орієнтованого алгебраїчного синтезу коректних алгоритмів, апарату оптимізації моделей класифікації, колективного розв'язання задач кластерного аналізу, асимптотично-оптимальних логічних алгоритмів, алгоритмів зваженого статистичного розпізнавання, логічних алгоритмів розпізнавання, дерева розв'язків свідчать про існування певних недоліків у разі їхнього застосування для розв'язання задач розпізнавання у сучасній постановці [2–14].

Нині одним з перспективних напрямів розв'язання задачі розпізнавання та класифікації є розроблення програмних систем аналізу даних і прогнозу за прецедентами. Це статистичні програмні продукти для оброблення даних та візуалізації (SPSS, STADIA, STATGRAPHICS, STATISTICA, SYSTAT, Forecast Expert тощо). В їхню основу покладено науково-методичний апарат математичної статистики — перевірку статистичних гіпотез, регресійний аналіз, дисперсійний аналіз тощо. Ці програмні продукти є ефективним інструментом аналізу даних передусім на початковому етапі досліджень (пошук усереднених показників, визначення статистичної достовірності гіпотез, пошук регресійних залежностей). Вони спеціально призначені для аналізу даних та орієнтовані на відповідні аспекти розв'язання практичних задач. Їхньою алгоритмічною основою є будь-яка альтернативна модель, що використовує нейронну мережу, дерево рішень, обмежений перебір тощо. Зазначені програмні системи мають істотні обмеження у разі їхнього практичного використання. Це зумовлено відсутністю універсальності відносно розмірності розв'язуваних задач, а також типом, складністю, суперечливістю та можливою низькою структурованістю вихідних даних. До того ж, вони створені та налаштовані на розв'язання відповідних задач. Це не відповідає вимогам користувачів, які потребують більш широких можливостей порівняно із закладеними у конкретному програмному продукті.

Отже, на думку авторів у сучасних умовах найбільш раціональним є шлях створення систем та комплексів моніторингу, що ґрунтуються на уніфікованих методичних підходах, які забезпечать більш якісне та точне розв'язання конкретної задачі користувача.

Спираючись на проведений аналіз наявних літературних джерел, можна зрозуміти, що розв'язання задач розпізнавання ДОМ є неоднозначною та складною проблемою. Виходячи з цього, автори вважають, що запропонований методичний підхід ідентифікації об'єктів моніторингу становить суттєвий науковий інтерес.

Актуальність цієї роботи зумовлена високою потребою у нових підходах до ідентифікації ДОМ в умовах часткової або повної невизначеності вихідних даних про ДОМ.

Метою роботи є розроблення методичного підходу ідентифікації об'єктів моніторингу шляхом застосування отриманих теоретичних та практичних результатів досліджень за цим напрямом [1, 15–18].

ЗАГАЛЬНИЙ ОПИС МЕТОДИЧНОГО ПІДХОДУ

Весь процес правильного розпізнавання ДОМ через визначення потрібної кількості важливих та інформативних МО, формування інформаційних моделей відповідних сигнатур ДОМ за визначеними МО, а також раціонального розподілу ТЗМ за відповідними об'єктами має складатися з таких етапів.

Етап 1. Формують МО та сигнатури ДОМ, застосовуючи удосконалений СГСМ (на основі методики пріоритизації МО та розробленого інтегрального критерію, який реалізує баланс між інформативністю та важливістю МО) для виявлення, аналізу та оцінювання кількісних і якісних параметрів ДОМ в умовах часткової або повної невизначеності вихідних даних про ДОМ. Інформаційною основою реалізації СГСМ у межах досліджень є бігаусівська (полігаусівська) модель, яка задає розподіли з «товстими хвостами», що свідчать про наявність аномальних похибок вимірювань в умовах, коли заводовими є широкосмугові шумоподібні сигнали і гаусівська модель стає непридатною для аналізу та синтезу вимірювачів параметрів сигналів[1].

На цьому етапі здійснюють синтез інформаційного уявлення про ДОМ системою моніторингу, а також формування реального уявлення про ДОМ для порівняння їхніх показників за визначеним критерієм схожості. На практиці задача полягає в розробленні інформаційно-аналітичних та імітаційних моделей, що дають системне відображення ДОМ через комплексні МО (сигнатури), які підлягають вимірюванню, аналізу і пов'язані з призначенням ДОМ.

Функціонал, який характеризує інформаційно-аналітичну модель ДОМ, наведено в [1]. В її основу покладено попередній детальний опис можливих умов, в яких перебуває ДОМ, якісне та кількісне оцінювання його основних показників, характер зв'язків між об'єктами та всередині них, алгоритми здобування і формування МО та відповідних сигнатур. Інформаційна модель складається з відповідних ідентифікаторів, сукупності ознак і сигнатур належності ДОМ відповідним класам, сукупності значень вагових коефіцієнтів ознак і сигнатур належності ДОМ відповідним класам.

Пріоритетність МО ґрунтується на оцінюванні їхньої інформативності та важливості. Одночасне застосування часткових (max–min)- та (max–max)-критеріїв сприяє підвищенню точності та достовірності розрахунків пріоритетів МО у разі виявлення суперечності під час досліджень. Запропонований на їхній основі інтегральний критерій мінімуму суми пріоритетів МО, отриманих за частковими (max–min)- та (max–max)-критеріями, усуває виявлену під час досліджень суперечність і збалансовує інформативність і важливість відповідної МО. Це надає змогу формувати сигнатуру відповідного ДОМ у певних фазових станах з використанням пріоритетних МО [1].

Етап 2. Здійснюють побудову математичних моделей МО та сигнатур, а також формалізацію задачі з подальшим синтезом алгоритмів прийняття рішення про відповідність обраного ДОМ одному із заданих образів вибірки значень його сигнатури [15, 16]. Реалізація цього етапу ускладнена важкістю і трудомісткістю формалізації сигнатур та МО, їхньою неоднорідною структурою, наявністю класів образів ДОМ, що перетинаються. На цьому етапі здійснюють розрахунок інформативності МО як основи математичної моделі визначення ДОМ. Інформативність МО оцінюють за рівнем ентропії $H_{\text{мож}}$, якого може досягти система моніторингу, при цьому вона буде ефективною за умови, що рівень ентропії не перевищує порогове значення $H_{\text{пор}}$. Він є достатнім для розв'язання основних задач ідентифікації і класифікації ДОМ під час моніторингу [15].

Формування сигнатури ДОМ здійснюють методами функціонального аналізу через те, що множину моніторингових ознак ДОМ розглядають як підмножину всієї множини ознак сигнатури в результаті оброблення інформаційного потоку про ДОМ. У свою чергу, складність ДОМ, їхній вплив на інші об'єкти або процеси зумовлюють інформативність їхніх МО. У випадку застосування системою моніторингу скінченної множини сенсорів процес перетворення інформації про ДОМ здійснюють у такий спосіб: виявляють, вимірюють (визначають) параметри; оцінюють параметри та на основі їхнього аналізу формують відповідні характеристики, хоча не всі вони є інформативними; будують множину інформативних моніторингових ознак ДОМ, отриманих шляхом зменшення початкової ентропії $H_{\text{поч}}$ його стану.

На цьому етапі обчислюють інформативність k -ї МО x_k будь-якого j -го ДОМ Z_j . Далі визначають ентропію рішення про Z_j з урахуванням μ -го дискретного значення МО x_k [15]. Для визначення ентропії за ознакою x потрібно обчислити суму $H(Z_j / x_{k\mu})$ j -го ДОМ з використанням всіх дискретних значень μ МО x_k з важливістю, пропорційною ймовірності появи кожного її дискретного значення. Враховуючи отримані значення початкової ентропії $H_{\text{поч}}(Z_j)$ та $H(Z_j / x_k)$, обчислюють інформативність окремої МО x_k під час розпізнавання j -го ДОМ. Результати розрахунку відносної інформативності окремої МО x_k використовують для визначення інформативності всієї сигнатури.

У загальному випадку немає функціональної залежності між похибкою розпізнавання та кількістю інформації для прийняття рішення про розпізнавання ДОМ [19, 20], тому необхідно виявити залежність між умовною ймовірністю похибки розпізнавання ДОМ та відносною інформативністю всієї сигнатури для конкретних випадків [16].

Застосування показника відносної інформативності сигнатури для розпізнавання ДОМ Z_j ($j = 1, 2$) у випадку дихотомії визначає тільки умовну ймовірність похибки розпізнавання $P_{\text{пох}}(Z / X)$, при цьому для проведення розрахунків приймається критерій максимальної апостеріорної ймовірності ($X \in Z_1$, якщо $P_{\text{пр}}(Z_1 / X) > P_{\text{пр}}(Z_2 / X)$ та $X \in Z_2$, якщо $P_{\text{пр}}(Z_2 / X) > P_{\text{пр}}(Z_1 / X)$) для випадків $p(Z_1)p(X / Z_1) > p(Z_2)p(X / Z_2)$ та $p(Z_1)p(X / Z_1) < p(Z_2)p(X / Z_2)$.

Умовну ймовірність правильного розпізнавання $P_{\text{пр}}$ за таких умов для кожного зазначеного випадку обчислюють за відповідними співвідношеннями [16].

Під час розв'язання задач ідентифікації ДОМ завдяки існуванню певних умов є можливість виявити залежність $f(I) = P_{\text{пох}}(Z / X)$ між умовною ймовірністю похибки розпізнавання ДОМ та відносною інформативністю всієї сигнатури. В інтерпретації результатів прямих вимірювань важливу роль відіграють три закони щільності розподілу ймовірності помилки [21].

В умовах фіксованої (заданої) умовної ймовірності $P_{\text{пох}}^{\text{потр}}(Z/X)$ похибки розпізнавання ДОМ під час визначення потрібної інформативності I сигнатури X можна використовувати розраховані кількісні значення у вигляді залежності $f(I) = P_{\text{пох}}(Z/X)$ [16].

Щоб отримати значення $P_{\text{пох}} < P_{\text{пох}}^{\text{потр}}(Z/X)$, потрібно забезпечити виконання співвідношення

$$H_{\text{потр}}(Z_j) - H(Z_j/X) > I(Z_j), \quad (1)$$

де

$$I(Z_j) = -H(Z_j/X)(1 + 0.5 \log_2 0.5H(Z_j/X)) - \\ - (1 - 0.5H(Z_j/X)) \log_2 (1 - 0.5H(Z_j/X)). \quad (2)$$

Отже, враховуючи інформативність $I_k(Z_j)$ окремої МО x_k та інформативність $I(Z_j)$ сигнатури X (сукупності моніторингових ознак x_k) (2), можна вважати, що пріоритетність моніторингової ознаки x_k зі складу сигнатури визначається співвідношенням

$$H(Z/X^{(K-1)}) - H(Z/X^{(K)}) > -H(Z/X^{(K)})(1 + 0.5 \log_2 0.5H(Z/X^{(K)})) - \\ - (1 - 0.5H(Z/X^{(K)})) \log_2 (1 - 0.5H(Z/X^{(K)})), \quad (3)$$

де $H(Z/X^{(K)})$ та $H(Z/X^{(K-1)})$ — умовні ентропії зміни стану Z_j , визначені через множину моніторингових ознак $X^{(K)}$ та $X^{(K-1)}$ (в якій відсутня МО x_k) відповідно.

У роботах [19, 20] точно визначено верхню та нижню межі умовної ентропії $H(Z/X^{(K)})$ для заданої ймовірності похибки розпізнавання. На основі отриманої з $X^{(K)}$ множини пріоритетних МО у визначених межах уточнюють існуючу сигнатуру або визначають нову.

Отримані результати розрахунків щодо визначення інформативності сигнатури в цілому та розрахунків ймовірності похибки розпізнавання ДОМ залежно від відносної інформативності сигнатури та апіорних ймовірностей ДОМ у конкретних умовах, але не у загальному випадку, дають змогу перейти до третього етапу.

Етап 3. Здійснюють розподіл технічних засобів за об'єктами моніторингу з урахуванням отриманих на попередніх етапах результатів. Він є дуже важливим, оскільки розпізнавання ДОМ відповідними ТЗМ ускладнено труднощами формалізації МО та сигнатур об'єктів, їхньою неоднорідною структурою та наявністю перетинних класів образів [17, 19]. Через це задача розподілу ТЗМ за ДОМ для конкретних прикладних завдань на сьогодні не розв'язана у повному обсязі. Щоб виконати раціональний розподіл ТЗМ за ДОМ, потрібно спочатку визначити фазові стани об'єктів моніторингу шляхом розпізнавання образів за вибіркою МО. Для цього запропоновано уявити впливи ДОМ на сенсори ТЗМ як лінгвістичні змінні, використовувані у теорії нечітких множин [17, 22]. На множині Ψ ДОМ визначають L образів Ψ_i , $i \in \{1, 2, \dots, L\}$, об'єктів моніторингу. Застосовуючи показник мінімальної розмірності, що забезпечує потрібний рівень інформативності, потрібно визначити перелік ознак S та на його основі розрахувати множину сигнатур X^S , кожна з яких містить еквівалентні значення ξ нечітких МО $s_j \in S$, $j \in \{1, 2, \dots, \xi\}$. Кожний з L образів характеризується відповідною сигнатурою $x_i^S = \{x'_{i1} \dots x''_{i1}, x'_{i2} \dots x''_{i2}, \dots, x'_{i\xi} \dots x''_{i\xi}\}$, $i = 1, 2, \dots, L$, $x_i^S \in X^S$, яка по суті є впорядкованою сукупністю інтервалів нечітких значень лінгвістичних ознак, представлених у цифровому вигляді. Кожному значенню

лінгвістичної ознаки у сигнатурі образу Ψ_i ставлять у відповідність інтервал значень з використанням відповідних функцій належності. На основі наявної статистичної інформації формують вибірку визначених еквівалентних значень $x = (x_1, x_2, \dots, x_j, \dots, x_\xi)$ лінгвістичних змінних, що є результатом упорядкованого оцінювання значень нечітких ознак $S_j, j \in \{1, 2, \dots, \xi\}$. Еталонний опис сигнатур являє собою сукупність інтервалів значень S_j , що задані у вигляді сукупності відповідних умовних щільностей ймовірностей значень ознак $w_{jr}(x_j, x'_j, x''_j)$. Ці щільності характеризують еталонний розподіл значень x_j j -ї ознаки S_j на кожному r -му інтервалі можливих значень x'_{jr}, \dots, x''_{jr} для кожного з образів, що розпізнаються.

За наявності еталонних описів сигнатур визначають нечіткі відношення «об'єкт моніторингу – МО», «МО – ТЗМ» та їхню композицію «об'єкт моніторингу – ТЗМ» з використанням пріоритетних МО фазового стану ДОМ. Зазначені нечіткі відношення є, так би мовити, інформаційними моделями визначення ДОМ в їхніх фазових станах з відповідними моніторинговими ознаками. Далі таким МО ставлять у відповідність наявні ТЗМ, за допомогою яких ці ознаки виявляють. Реалізація цієї композиції надає змогу розрахувати ступінь відповідності між ДОМ та наявними ТЗМ [17].

Фазові стани (їхні образи) ДОМ $\Psi_i \subset \Psi$, що розпізнаються, задані їхнім еталонним описом в області X_i^S ξ -вимірного евклідового простору еталонів X^S з осями $x_1, x_2, \dots, x_j, \dots, x_\xi$. Еталонний опис образів фазових станів надають у вигляді суми отриманих у результаті навчання та об'єднаних в інформаційні сигнатури апріорних щільностей розподілу ймовірностей $w_j(x)$ еталонних значень ознак $x_1 \dots x_\xi$ на множині Ψ_i :

$$w_j(x) = \sum_{r=1}^{R_j} p_{jr} w_{jr}(x_j^e, x_{jr}^{e'}, x_{jr}^{e''}), \sum_{r=1}^{R_j} p_{jr} = 1, \quad (4)$$

де $w_j(x_j^e, x_{jr}^{e'}, x_{jr}^{e''})$ — зважені щільності ймовірності еталонного розподілу МО x_j^e на кожному r -му інтервалі можливих значень $x_{jr}^{e'}, \dots, x_{jr}^{e''}$ j -ї ознаки для кожного образу, що розпізнається, p_{jr} — умовні апріорні ймовірності потрапляння значення МО в r -й інтервал, $r \in \{1, 2, \dots, R_{ij}\}$, j -ї ознаки в i -му образі, $\forall j \in \{1, 2, \dots, \xi\}$.

Кожний еталонний інтервал значення j -ї ознаки $x_j^e, j \in \{1, 2, \dots, \xi\}$, в одному образі відповідає одній координаті цієї ознаки, а кожна сигнатура містить ξ інтервалів значень ознак. Виходячи з цього створюють область визначення функцій щільності ймовірностей $w_j(x_j^e, x_{jr}^{e'}, x_{jr}^{e''})$ значень складного еталонного опису (4) відповідного виду [23]. При цьому висувають N гіпотез H_1, H_2, \dots, H_N про те, що спостережувана вибірка \mathbf{x} , що складається з ϑ -кратно оцінених значень ξ ознак і має розмір $(\vartheta \times \xi)$, належить одному з описаних образів Ψ_i через значення сигнатури ознак відповідного класу. У цьому випадку простір рішень складається з N елементів γ_i рішень про ухвалення гіпотези H_i . Далі обирають нерандомізоване правило δ , за яким реалізують поділ простору X на N перетинних областей

$x_i, \bigcup_{i=1}^N x_i = X$. Якщо під час розпізнавання реалізується одночасне потрапляння

кожного значення ознаки з вибірки $\mathbf{x} = (x_1, x_2, \dots, x_j, \dots, x_\xi)$ у відповідні інтервали сигнатури

$$x_i^S = \{x'_{i1} \dots x''_{i1}, x'_{i2} \dots x''_{i2}, \dots, x'_{i\xi} \dots x''_{i\xi}\}, i = 1, \dots, L, \quad (5)$$

то ухвалюють рішення γ_i .

У [23] розглянуто синтез мінімаксних алгоритмів об'єктів моніторингу у постановці задачі багатоальтернативного розпізнавання та для одноальтернативного випадку шляхом розгляду множини об'єктів розпізнавання (моніторингу) Ψ , в якій задано лише два образи $\Psi_1 \subset \Psi$ та $\Psi_2 \subset \Psi$, де Ψ_1 відповідає образу об'єкта моніторингу, а Ψ_2 йому не відповідає.

За наведеною послідовністю перевіряють лише дві гіпотези на відповідність спостережуваної вибірки x розміру $(\vartheta \times \xi)$, що складається з ϑ -кратно вимірюваних значень ξ лінгвістичних ознак S_j , а також на відповідність сигнатурі образу Ψ_1 чи Ψ_2 об'єкта моніторингу. За результатами перевірки ухвалюють рішення γ_1 або γ_2 , які відповідають коректності гіпотез H_1 або H_2 . Однак, під час формування еталонного розподілу j -ї ознаки S_j у вигляді суми умовних розподілів $w_{1j}(x_j^e)$, $w_{2j}(x_j^e)$, зважених із невідомими апіорними ймовірностями p_1 та $p_2 = 1 - p_1$, виявляється, що на практиці ймовірності p_{1jr} , p_{2jr} є апіорі невідомими, тому однозначний розв'язок цієї задачі отримати складно. Проте його можна одержати, якщо для кожної ознаки в одному образі ДОМ вважати ці ймовірності однаковими, а потім під час розпізнавання оцінити їх та внести коригування у правило ухвалення рішення. Далі на основі (4), (5) розраховують мінімаксне правило розпізнавання образів Ψ_1 та Ψ_2 , за яким порівнюють відношення правдоподібності з пороговим значенням.

Для реалізації розподілу технічних засобів за об'єктами моніторингу вихідні дані є такими: множина ДОМ $\Psi = \{Z_j\}_{j=1}^J$, де J — кількість об'єктів моніторингу; множина S фазових станів j -го ДОМ $S = \{S_j\}_{j=1}^J$, $S_j = \{C_i\}_{i=1}^{r_j}$, де r_j — кількість фазових станів j -го ДОМ; множина МО фазових станів ДОМ $F = \{R_n\}_{n=1}^N$, де N — кількість МО фазових станів ДОМ; коефіцієнт важливості λ_{ij} i -го фазового стану j -го ДОМ, коефіцієнт K_{nij} інформативності n -ї МО i -го фазового стану j -го ДОМ. Розрахунок нечіткого відношення «об'єкт моніторингу – МО» здійснюють шляхом побудови функції належності $\mu_A(Z_j, R_n)$, яка відображає ступінь належності МО фазового стану R_n протягом визначеного інтервалу часу відповідному ДОМ Z_j . Для формування цього нечіткого відношення застосовують лише пріоритетні МО, що визначені і розраховані з використанням коефіцієнтів важливості λ_{ij} i -го фазового стану j -го ДОМ та інформативності K_{nij} n -ї МО i -го фазового стану j -го ДОМ [1]. Розрахунок пріоритетності моніторингових ознак наведено в [1]. Далі розраховують нечітке відношення «МО – ТЗМ» шляхом побудови функції належності $\mu_B(R_n, K_m)$, яка відображає ступінь відповідності K_m технічного засобу моніторингу моніторинговим ознакам фазового стану ДОМ протягом визначеного інтервалу часу [17]. Побудовані та розраховані нечіткі відношення є інформаційними моделями, що дають системне уявлення про ДОМ через відповідні МО (сигнатури), які пов'язані з призначенням, можливостями, фазовими станами тощо, та про ТЗМ, потрібні для їхньої ідентифікації. З використанням побудованих нечітких відношень розраховують їхню композицію, що встановлює системне уявлення на рівні «об'єкт моніторингу – ТЗМ» через функцію належності $\mu_C(Z_j, K_m)$, яка характеризує ступінь відповідності технічного засобу моніторингу K_m об'єкту моніторингу O_j [17].

Далі обчислюють матрицю попарних переваг V , яка визначає ступінь відповідності кожного Z_j одночасно кожній комбінації двох варіантів ТЗМ K_m . За її допомогою визначають порогове значення λ відповідності Z_j як ступеня можливості виявлення всіх Z_j для всіх пар K_m технічних засобів моніторингу:

$$\lambda < \min_{m-1, m} \max_{Z_j} \min [\mu_{C_{m-1}}(Z_j, K_{m-1}), \mu_{C_m}(Z_j, K_m)]. \quad (6)$$

Маючи кількісне значення цього параметра, обчислене для кожного варіанта K_m ТЗМ, можна визначити рівневу множину об'єктів моніторингу P , ступінь відповідності яких дорівнює пороговому значенню λ можливості виявлення ДОМ варіантом ТЗМ або перевищує його:

$$P = \{Z_j | \mu_{C_m}(Z_j) \geq \min_{m-1, m} \max_{Z_j} \min [\mu_{C_{m-1}}(Z_j, K_{m-1}), \mu_{C_m}(Z_j, K_m)]\}, Z_j \in P. \quad (7)$$

Після розрахунку рівневої множини ДОМ можна перейти до наступного етапу.

Етап 4. Завдяки розрахованим можливостям певного варіанта ТЗМ можна оцінити його ефективність у процесі розпізнавання ДОМ. Якщо вона є недостатньою, слід розрахувати коригувальні коефіцієнти, як запропоновано у [18] для інтегрального критерію мінімуму суми пріоритетів МО, отриманих за частковими (max–min)- та (max–max)-ми критеріями [1], який враховує вплив інформативності МО на умови його формування. Коригувальні коефіцієнти визначають шляхом врахування інформативності МО, у тому числі пріоритетних, які використовуються у поєднанні з іншими [17, 19]. Для визначення цих коефіцієнтів під час моніторингу використовують поняття інформативності МО у поєднанні з іншими, тобто те, що моніторингова ознака $x_n, 1 \leq n \leq S$, є інформативною, якщо у разі її вилучення з множини $x^{(S)} = (x_1, \dots, x_S)$ середня умовна ентропія змінюється

$$H(V | X^{(S-1)}) - H(V | X^{(S)}) > 0, \quad (8)$$

де

$$H(V | X^{(S)}) \triangleq - \sum_{x^{(S)} \in X^{(S)}} P(x^{(S)}) \sum_{m=1, 2} P(V_m | x^{(S)}) \log P(V_m | x^{(S)}), \quad (9)$$

$$H(V | X^{(S-1)}) \triangleq - \sum_{x^{(S-1)} \in X^{(S-1)}} P(x^{(S-1)}) \sum_{m=1, 2} P(V_m | x^{(S-1)}) \log P(V_m | x^{(S-1)}), \quad (10)$$

$x^{(S-1)}$ — скорочений вектор ознак без x_n , а $X^{(S-1)} \triangleq \{x^{(S-1)} : p(x^{(S-1)}) > 0\}$.

З огляду на викладене та прикладний аспект розв'язування задач моніторингу, більш доцільним є використання величини зміни середньої ймовірності помилки $P(e)$ під час розпізнавання, а не зміни середньої умовної ентропії. Оскільки між середньою ймовірністю помилкових рішень та середньою умовною ентропією немає однозначного зв'язку [19, 20], можна встановити лише часткові випадки їхньої відповідності [18].

У розв'язуванні задач розпізнавання під час моніторингу використовують дискретні ознаки, які мають кількість градацій $\mu > 2$. За таких умов точна нижня межа середньої умовної ентропії $H(V | X)$ досягається для будь-якого значення середньої ймовірності помилкових рішень $P(e)$, $0 \leq P(e) \leq \min \{P(V_1), P(V_2)\}$. Тому незалежно від кількості градацій ($\mu > 2$) та за виконання умов $P(e|x) = 0 \forall x \in X'$, $P(e|x) = 0.5 \forall x \in X''$, середня умовна ентропія $H(V | X)$ досягає своєї точної нижньої межі, що збігається з її точною верхньою межею, для двох крайніх значень $P(e)$:

$$P(e) = 0 \text{ для } X'' \neq \emptyset;$$

$$P(e) = 0.5 \text{ для } X' = \emptyset \text{ (} p(x|V_1) \equiv p(x|V_2) \text{ для } p(V_1) = p(V_2) = 0.5 \text{)}.$$

Практичний інтерес становить визначення умов, виконання яких забезпечить $H(V | X) = \inf H(V | X)$ для $0 < P(e) < 0.5$ та за обмежень $X' = \emptyset$, $X'' \neq \emptyset$.

У випадку $\mu > 2$, який є найбільш цікавим для задач розпізнавання, для одних значень $x \in X'$ справджуються співвідношення $p(x|V_1) > 0$, $p(x|V_2) = 0$, а для інших значень $x \in X''$ — співвідношення $p(x|V_1) = 0$, $p(x|V_2) > 0$. При цьому середня ймовірність помилкових рішень $P(e)$, для кількісного значення якої

$H(V|X) = \inf H(V|X)$, одночасно задовольняє співвідношення $P(e) < P(V_1)$ та $P(e) < P(V_2)$. Отже, $P(e) < \min \{P(V_1), P(V_2)\}$.

Тому $p(x|V_1)$ та $p(x|V_2)$ для $x \in X''$ з урахуванням обмежень $p(x|V_1)P(V_1) = p(x|V_2)P(V_2)$ можна задати у такий спосіб, щоб значення $P(e)$, яке розраховується за співвідношенням

$$P(e) = P(V_1) \sum_{x \in X''} p(x|V_1) + P(V_2) \sum_{x \in X''} p(x|V_2),$$

дорівнювало заздалегідь визначеному α .

Розрахунки свідчать про те, що значення $P(e)$ розташовуються вище значень, отриманих за співвідношенням $\inf H(V|X) = 2P(e)$ [18–20]. За рахунок варіювання параметром α можна уточнити інтервали значень $\inf H(V|X)$ залежно від $H(V|X)$.

Отже, на основі результатів визначення інформативності МО у поєднанні з іншими можна обчислити коригувальні коефіцієнти для інтегрального критерію та внести зміни у перелік пріоритетних МО, які матимуть безпосередній вплив на коефіцієнт важливості та будуть застосовуватися на етапах побудови математичних моделей МО та сигнатури, а також визначення варіанта комплекту технічних засобів моніторингу для відповідних ДОМ.

ВИСНОВКИ

Застосування сигнатурно-інформаційних технологій збору, накопичення та оброблення інформації у розробленому методичному підході до розпізнавання об'єктів моніторингу на основі раціонального розподілу технічних засобів моніторингу за об'єктами з використанням інформаційних моделей їхніх ознак та сигнатур забезпечить підвищення ефективності процесу розпізнавання ДОМ в цілому.

СПИСОК ЛІТЕРАТУРИ

1. Льяшов О.А., Комаров В.С. Удосконалення сигнатурно-системного методу із застосуванням пріоритетних ознак у складі сигнатур об'єктів моніторингу для забезпечення максимальної ймовірності правильного розпізнавання. *Кибернетика и системный анализ*. 2020. Т. 56, № 4. С. 68–78.
2. Бишоп К.М. Распознавание образов и машинное обучение. СПб.: Диалектика, 2020. 960 с.
3. Дубовой В.М., Кветний Р.Н., Михальов О.І., Усов А.В. Моделювання та оптимізація систем. Вінниця: ПП «ТД «Едельвейс», 2017. 804 с.
4. Ротштейн А.П. Интеллектуальные технологии идентификации. Вінниця: Універсум-Вінниця, 1999. 302 с.
5. Айзерман М.А., Браверманн Э.М., Розоноэр Л.И. Метод потенциалных функций в теории обучения машин. Москва: Наука, 1970. 384 с.
6. Вапник В.Н., Червоненкис А.Я. Теория распознавания образов (статистические проблемы обучения). Москва: Наука, 1974. 415 с.
7. Донской В.И., Башта А.И. Дискретные модели принятия решений при неполной информации. Симферополь: Таврия, 1992. 166 с.
8. Дуда Р., Харт П., Распознавание образов и анализ сцен. Москва: Мир, 1976. 511 с.
9. Загоруйко Н.Г., Елкина В.Н., Лбов Г.С. Алгоритмы обнаружения эмпирических закономерностей. Новосибирск: Наука, 1985. 110 с.
10. Ивахненко А.Г. Системы эвристической самоорганизации в технической кибернетике. Киев: Техніка, 1971. 372 с.

11. Лбов Г.С. Методы обработки разнотипных экспериментальных данных. Новосибирск: Наука, 1981. 160 с.
12. Фукунага К. Введение в статистическую теорию распознавания образов. Москва: Наука, 1979. 367 с.
13. Симаков В.С., Луценко Е.В. Адаптивное управление сложными системами на основе теории распознавания образов. Краснодар: КГТУ, 1999. 318 с.
14. Білокур М.О. Функціональне відображення значень ваг в штучній нейронній мережі визначених властивостей при оцінюванні альтернативних зразків озброєння. *Озброєння та військова техніка*. 2020. № 2(26). С. 20–31.
15. Льяшов О.А., Комаров В.С. Підхід до побудови математичної моделі визначення об'єктів моніторингу із застосуванням інформативності їхніх моніторингових ознак. *Кибернетика и системный анализ*. 2020. Т. 56, № 5. С. 95–99.
16. Льяшов О.А., Комаров В.С., Юрченко О.В., Дейнега О.В. Підхід до побудови математичної моделі сигнатури визначення об'єктів моніторингу із застосуванням інформативності її моніторингових ознак. *Кибернетика и системный анализ*. 2020. Т. 56, № 6. С. 115–119.
17. Льяшов О.А., Комаров В.С., Олексіюк В.В. Формалізація порядку розподілу технічних засобів за об'єктами моніторингу на основі нечітких множин. *Кибернетика та системний аналіз*. 2021. Т. 57, № 3. С. 91–96.
18. Комаров В.С., Олексіюк В.В., Льяшов О.А. Вплив інформативності моніторингових ознак на умови формування інтегрального критерію визначення їхньої пріоритетності. *Кибернетика та системний аналіз*. 2021. Т. 57, № 4. С. 131–135.
19. Файнзильберг Л.С. Математические методы оценки полезности диагностических признаков. Киев: Освіта України, 2010. 152 с.
20. Ковалевский В.А. Методы оптимальных решений в распознавании изображений. Москва: Наука, 1976. 328 с.
21. Вапник В.Н. Восстановление зависимостей по эмпирическим данным. Москва: Наука, 1979. 448 с.
22. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений. Москва: Мир, 1976. 165 с.
23. Льяшов О.А. Методика синтезу алгоритмів розпізнавання уразливостей web-ресурсів за сигнатурами нечітких лінгвістичних ознак. *Кибернетика и системный анализ*. 2017. Т. 53, № 3. С. 85–92.

V. Komarov, O. Liashov, V. Oleksiuk

METHODOLOGICAL APPROACH TO IDENTIFICATION OF MONITORING OBJECTS

Abstract. A methodological approach to identification of monitoring objects, based on the results of theoretical and applied research, is developed. Its practical application makes it possible to calculate the adjustment factors of the integrated criterion for determining the importance and informativeness of the monitoring features and to determine the most efficient version of the set of technical means of monitoring of the relevant sources (objects) of monitoring.

Keywords: monitoring feature, technical monitoring tool, signature, monitoring features priority.

Надійшла до редакції 08.06.2022