

ПІДХІД ДО ПОБУДОВИ МАТЕМАТИЧНОЇ МОДЕЛІ СИГНАТУРИ ВИЗНАЧЕННЯ ОБ'ЄКТІВ МОНІТОРИНГУ ІЗ ЗАСТОСУВАННЯМ ІНФОРМАТИВНОСТІ ЇЇ МОНІТОРИНГОВИХ ОЗНАК

Анотація. У статті розглянуто задачу використання інформативності моніторингових ознак сигнатури об'єктів моніторингу для конкретних умов розв'язання прикладних задач розпізнавання. Досліджено запропонований підхід до побудови математичної моделі сигнатури визначення об'єктів моніторингу шляхом варіювання величиною різниці між імовірностями розпізнавання об'єктів моніторингу.

Ключові слова: моніторингова ознака, сигнатура, інформативність моніторингової ознаки.

ВСТУП

На сьогодні актуальними завданнями є розроблення моделей, які нададуть змогу здійснювати формування моніторингових ознак та сигнатур потрібного джерела (об'єкта) моніторингу (ДОМ), а також формалізація задачі про відповідність обраного ДОМ одному із заданих образів вибірки значень його інформаційної сигнатури з подальшим синтезом алгоритмів прийняття рішення.

Формалізація сигнатур та моніторингових ознак ускладнена їхньою неоднорідною структурою та наявністю класів образів, що перетинаються, тому завдання формалізації еталонного опису ДОМ та їхнього розпізнавання за відповідними моніторинговими ознаками (МО) не розв'язані у повному обсязі [1–3].

ПОСТАНОВКА ЗАДАЧІ

Неоднорідний, нечіткий характер МО та побудованих на їхній основі сигнатур, а також недостатність навчальної вибірки суттєво ускладнюють розроблення еталонного опису (сигнатури) з використанням імовірнісно-статистичних моделей [3]. Застосування інформаційного підходу до опису сигнатур ДОМ надає змогу отримувати МО з певними похибками (втратами), що призводить до збільшення ентропії, втрат інформативності ознак і зменшення ймовірності правильного розпізнавання ДОМ. Проте у деяких випадках сукупність неінформативних МО забезпечуватиме потрібний рівень класифікації ДОМ [4]. Запропонований підхід щодо використання лише пріоритетних МО [3] надасть змогу певною мірою нівелювати в окремих випадках визначені вище труднощі, пов'язані з процесом формування сигнатури з потрібною відносною інформативністю МО, які її складають.

Оскільки у загальному випадку немає функціональної залежності між похибкою розпізнавання та кількістю інформації для прийняття рішення про розпізнавання ДОМ [4, 5], виникає потреба в побудові підходу до розв'язання прикладних задач у конкретних умовах, які надають змогу виявити залежність між умовною ймовірністю похибки розпізнавання ДОМ та відносною інформативністю всієї сигнатури.

**ПОРЯДОК ПОБУДОВИ СИГНАТУРИ НА ОСНОВІ УТОЧНЕННЯ ІНФОРМАТИВНОСТІ
МОНІТОРИНГОВИХ ОЗНАК У ВІДПОВІДНИХ УМОВАХ**

Інформативність моніторингової ознаки x_k з усієї множини моніторингових ознак $X = \{x_1, x_2, \dots, x_k, \dots, x_N\}$ визначається співвідношенням

$$I_k(Z_j) = 1 - \left(\sum_{\mu=1}^{M_k} \sum_{i=1}^{N_j} p(Z_{ji}, x_{k\mu}) \log_2 p(Z_{ji}, x_{k\mu}) + \sum_{\mu=1}^{M_k} \sum_{i=1}^{N_j} p(Z_{ji}, x_{k\mu}) \sum_{i=1}^{N_j} \log_2 p(Z_{ji}, x_{k\mu}) \right) / \sum_{i=1}^{N_j} p(Z_{ji}) \log_2 p(Z_{ji}). \quad (1)$$

За допомогою наведеного співвідношення (1) можна розрахувати відносну інформативність лише окремої моніторингової ознаки x_k , наприклад, під час розпізнавання двох джерел (об'єктів) моніторингу.

Інформативність сигнатури, яка складається з певної сукупності моніторингових ознак, характеризується імовірністю похибки розпізнавання $P_{\text{пох}}$ (відповідно ймовірність правильного розпізнавання $P_{\text{пр}} = 1 - P_{\text{пох}}$).

Застосування показника відносної інформативності сигнатури для розпізнавання двох ДОМ Z_j , $j=1,2$, визначає тільки умовну ймовірність похибки розпізнавання $P_{\text{пох}}(Z/X)$ та за умови, що 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 $P_{\text{пр}}(Z_2/X) > P_{\text{пр}}(Z_1/X)$) отримаємо:

$$P_{\text{пох}}(Z_j/X) = \frac{1}{1 + [p(Z_1)/p(Z_2)][p(X/Z_1)/p(X/Z_2)]}, \quad j=1,2, \quad (2)$$

для $p(Z_1)p(X/Z_1) > p(Z_2)p(X/Z_2)$,

$$P_{\text{пох}}(Z_j/X) = \frac{1}{1 + [p(Z_2)/p(Z_1)][p(X/Z_2)/p(X/Z_1)]}, \quad j=1,2, \quad (3)$$

для $p(Z_1)p(X/Z_1) < p(Z_2)p(X/Z_2)$.

Умовна ймовірність правильного розпізнавання $P_{\text{пр}}(Z/X)$ за умов (2) та (3) розраховується за співвідношеннями:

$$P_{\text{пр}}(Z_1/X) = \frac{p(Z_1)p(X/Z_1)}{p(Z_1)p(X/Z_1) + p(Z_2)p(X/Z_2)}, \quad (4)$$

для $p(Z_1)p(X/Z_1) > p(Z_2)p(X/Z_2)$,

$$P_{\text{пр}}(Z_2/X) = \frac{p(Z_2)p(X/Z_2)}{p(Z_1)p(X/Z_1) + p(Z_2)p(X/Z_2)}, \quad (5)$$

для $p(Z_1)p(X/Z_1) < p(Z_2)p(X/Z_2)$.

Варто зазначити, що у загальному випадку не існує функціональної залежності між похибкою розпізнавання $P_{\text{пох}}$ та кількістю інформації (ентропією $H(Z_j/X)$, $j=1,2$) для прийняття рішення про розпізнавання ДОМ [4]. Однак, під час розв'язання прикладних задач існують умови, які надають змогу на основі (1)–(3) 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: рівномірний закон розподілу, нормальний закон та закон Лапласа. Рівномірним законом щільності розподілу

ймовірностей описують помилки заокруглення вимірювальних приладів, а нормальним законом — помилки, які виникають у фізичних вимірюваннях, що проводяться в одних і тих самих умовах. Помилки фізичних вимірювань, проведених в умовах, що змінюються, описують законом Лапласа [6]. Під час розв'язання практичних задач такі вимірювання проводять в один і той самий час.

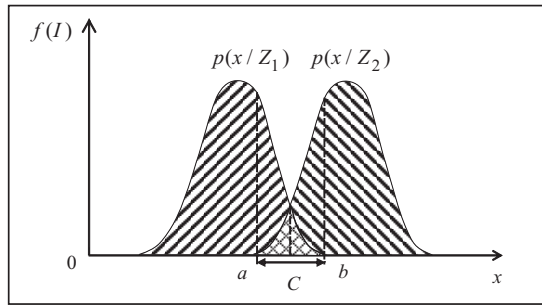


Рис. 1. Варіант розташування щільності розподілу $p(x/Z_1)$, $p(x/Z_2)$, що розглядається

Виділяють чотири варіанти розташування щільності розподілу $p(x/Z_1)$, $p(x/Z_2)$ деякої моніторингової ознаки x для двох об'єктів моніторингу [4].

Розглянемо випадок, коли різниця між імовірностями розпізнавання двох об'єктів моніторингу

$$|P(Z_1) - P(Z_2)| \geq C, \quad (6)$$

де C — наперед задане додатне число.

Це відповідає випадку, зображеному на рис. 1. Значення C є еквівалентним відстані між прямими a та b , не перевищує різницю між середніми значеннями (математичними сподіваннями) моніторингової ознаки x для двох об'єктів моніторингу і є різним для різних ДОМ.

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 $f(I) = P_{\text{пох}}(Z/X)$.

Щоб отримати значення $P_{\text{пох}} < P_{\text{пох}}^{\text{потр}}$, необхідно забезпечити співвідношення

$$I(Z_j) - W_{\text{потр}}(Z_j) > \frac{W(Z_j)}{H_0(Z_j)}, \quad j=1,2. \quad (7)$$

Знаючи, що

$$I(Z_j) = \frac{H_0(Z_j) - H(Z_j/X)}{H_0(Z_j)}, \quad I_{\text{потр}}(Z_j) = \frac{H_0(Z_j) - H_{\text{потр}}(Z_j/X)}{H_0(Z_j)}, \quad j=1,2, \quad (8)$$

отримуємо

$$H_{\text{потр}}(Z_j) - H(Z_j/X) > I(Z_j), \quad (9)$$

де

$$I(Z_j) = -H(Z_j/X)(1 + 0,5 \log_2 0,5H(Z_j/X)) - (1 - 0,5H(Z_j/X)) \log_2 (1 - 0,5H(Z_j/X)). \quad (10)$$

Отже, враховуючи інформативність $I_k(Z_j)$ окремої моніторингової ознаки x_k (1) та інформативність $I(Z_j)$ сигнатури X (сукупності моніторингових ознак x_k) (10), можна вважати, що пріоритетність моніторингової ознаки x_k зі складу сигнатури визначається співвідношенням

$$H(Z/X^{(K-1)}) - H(Z/X^{(K)}) > -H(Z/X^{(K)})(1 + 0,5 \log_2 0,5H(Z/X^{(K)})) - (1 - 0,5H(Z/X^{(K)})) \log_2 (1 - 0,5H(Z/X^{(K)})), \quad (11)$$

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 моніторингова ознака x_k) відповідно.

У роботі Ковалевського [5] визначено точні верхня та нижня межі умовної ентропії $H(Z/X^{(K)})$ для заданої ймовірності похибки розпізнавання. У випадку розпізнавання двох класів $Z_j, j=1,2$, межі умовної ентропії визначають так:

$$\begin{aligned} \inf H(Z/X^{(K)}) &= 2P_{\text{пох}}, \\ \sup H(Z/X^{(K)}) &= -P_{\text{пох}} \log_2 P_{\text{пох}} - (1-P_{\text{пох}}) \log_2 (1-P_{\text{пох}}). \end{aligned} \quad (12)$$

На основі отриманої з $X^{(K)}$ множини пріоритетних моніторингових ознак у визначених межах (12) уточнюють існуючу сигнатуру або визначають нову.

ВИСНОВКИ

У роботі запропоновано підхід до побудови математичної моделі сигнатури визначення об'єктів моніторингу на основі інформативності її моніторингових ознак із застосуванням розробленого порядку. Суть підходу полягає у варіюванні величиною різниці між ймовірностями розпізнавання двох об'єктів моніторингу для прийняття рішення на користь того чи іншого об'єкта.

Проведено дослідження можливості розрахунку інформативності сигнатури з використанням пріоритетних моніторингових ознак, одержаних за розробленим порядком.

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, є підґрунтям для подальших досліджень щодо побудови відповідних залежностей.

СПИСОК ЛІТЕРАТУРИ

1. Льяшов О.А. Методика синтеза алгоритмов распознавания уязвимостей web-ресурсов за сигнатурами значень нечітких лінгвістичних ознак. *Кибернетика и системный анализ*. 2017. Т. 53, № 3. С. 85–92.
2. Льяшов О.А. Оцінка інформативності моніторингових ознак і сигнатур та міри їх невизначеності при розпізнаванні джерел та об'єктів моніторингу в інформаційному середовищі телекомунікаційних систем. *Вісник НТУУ «КПІ». Серія Радіотехніка, Радіоапаратобудування*. 2016. № 67. С. 77–83.
3. 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.
4. Файнзильберг Л.С. Математические методы оценки полезности диагностических признаков. Киев: Освіта України, 2010. 152 с.
5. Ковалевский В.А. Методы оптимальных решений в распознавании изображений. Москва: Наука, 1976. 328 с.
6. Вапник В.Н. Восстановление зависимостей по эмпирическим данным. Москва: Наука, 1979. 448 с.

Надійшла до редакції 26.05.2020

А.А. Ильяшов, В.С. Комаров, О.В. Юрченко, А.В. Дейнега
ПОДХОД К ПОСТРОЕНИЮ МАТЕМАТИЧЕСКОЙ МОДЕЛИ СИГНАТУРЫ
ОПРЕДЕЛЕНИЯ ОБЪЕКТОВ МОНИТОРИНГА С ИСПОЛЬЗОВАНИЕМ
ИНФОРМАТИВНОСТИ ЕЕ МОНИТОРИНГОВЫХ ПРИЗНАКОВ

Аннотация. В статье рассмотрена задача использования информативности мониторинговых признаков сигнатуры объектов мониторинга для конкретных условий решения прикладных задач распознавания. Исследован предложенный подход к построению математической модели сигнатуры определения объектов мониторинга путем учета изменения величины разницы между вероятностями распознавания объектов мониторинга.

Ключевые слова: мониторинговый признак, сигнатура, информативность мониторингового признака.

O. Piashov, V. Komarov, O. Yurchenko, O. Deineha
AN APPROACH TO CONSTRUCTING A MATHEMATICAL MODEL OF THE SIGNATURE
TO DETERMINE THE MONITORING OBJECTS USING THE INFORMATIVITY OF ITS
MONITORING FEATURES

Abstract. The authors consider the use of informativity of monitoring attributes of a signature of objects of monitoring for specific conditions of solving applied recognition problems. The approach to construction of the mathematical model of a signature of definition of objects of monitoring by varying the difference between the probabilities of recognition of the objects of monitoring is proposed.

Keywords: monitoring sign, signature, information content of the monitoring sign.

Ильяшов Александр Авксентійович,

доктор військ. наук, професор, головний науковий співробітник Центрального науково-дослідного інституту Збройних Сил України, Київ, e-mail: aleksandr.ilyashov@gmail.com.

Комаров Володимир Сергійович,

доктор військ. наук, старший науковий співробітник, начальник науково-дослідного управління Центрального науково-дослідного інституту Збройних Сил України, Київ, e-mail: komarvlad@ukr.net.

Юрченко Олег Васильович,

доцент кафедри Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського», e-mail: olegyurchenko7@gmail.com.

Дейнега Александр Васильович,

доктор військ. наук, старший науковий співробітник, головний науковий співробітник Центрального науково-дослідного інституту Збройних Сил України, Київ, e-mail: davdoma@gmail.com.