DOI: 10.46299/ISG.2025.MONO.TECH.1.8.5

8.5 Науково-методичний апарат обробки різнотипних даних в автоматизованих системах управління спеціального призначення

8.5.1. Обґрунтування доцільності використання теорії штучного інтелекту для обробки різнотипних даних в автоматизованих системах управління

Метою цього розділу є обґрунтування необхідності застосування теорії нечітких графів для опису процесу обробки різнотипних даних в автоматизованих системах управління (АСУ) [306-325].

Одним із можливих способів опису моделей процесу опису процесу обробки різнотипних даних в АСУ є застосування теорії нечітких графів, основною перевагою якої є можливість адекватного представлення вихідних даних відносно вхідної інформації, що представляється слабо-структурованими (нечіткими) показниками. Дана перевага дозволяє застосувати теорію нечітких графів в задачах аналізу оперативної обстановки в умовах невизначеності [326–344].

Модель процесу обробки різнотипних даних в АСУ можна представити у вигляді матриці знань (бази знань) (таблиця 1), що містить кількісні та якісні ознаки та характеристики функціонування АСУ [345-364].

Матрицею знань [365, 366] називається таблиця, що сформована за такими правилами:

1. Розмірність матриці $(n+1)\times N$, де $(n+1)$ – кількість стовпців, а $N = k_1 + k_2 + ... + k_m$ – кількість рядків.

2. Перші n стовпців матриці відповідають вхідним змінним $i = \overline{1,n}$, а $(n+1)$-й стовпець відповідає значенням d_j вихідній змінній $y(j = \overline{1,m})$.

3. Кожний рядок матриці являє собою деяку комбінацію знань вхідних змінних, яка віднесена експертом до одного з можливих значень вихідної змінної

y. При цьому: перші k_j рядків відповідають значенню вихідної змінної $y = d_1$, другі k_2 рядків – значенню $y = d_2$, останні k_m – значенню $y = d_m$.

4. Елемент a_i^{jp}, що стоїть на перетині i-го стовпця та jp-го рядка відповідає лінгвістичній оцінці параметра x_i в рядку нечіткої бази знань з номером jp. При цьому лінгвістична оцінка a_i^{jp} обирається з терм-множини відповідної змінної x_i, тобто $a_i^{jp} \in A_i$, $i = \overline{1,n}$, $j = \overline{1,m}$, $p = \overline{1,k_j}$.

Таблиця 1 – Модель функціонування АСУ на проміжку часу

Номер вектору ознак на вході	Ознаки елементів АСУ (вхідні змінні)				Рішення оцінювання обстановки (Вихідна змінна)
	x_1	x_2	$\dots x_i \dots$	x_n	y
11	a_1^{11}	a_2^{11}	$\dots a_i^{11} \dots$	a_n^{11}	d_1
12	a_1^{12}	a_2^{12}	$\dots a_i^{12} \dots$	a_n^{12}	
…	…	…	…	…	
$1k_1$	$a_1^{1k_1}$	$a_2^{1k_1}$	$\dots a_i^{1k_1} \dots$	$a_n^{1k_1}$	
…					
$j1$	a_1^{j1}	a_2^{j1}	$\dots a_i^{j1} \dots$	a_n^{j1}	d_j
$j2$	a_1^{j2}	a_2^{j2}	$\dots a_i^{j2} \dots$	a_n^{j2}	
…	…	…	…	…	
jk_j	$a_1^{jk_j}$	$a_2^{jk_j}$	$\dots a_i^{jk_j} \dots$	$a_n^{jk_j}$	
…					
$m1$	a_1^{m1}	a_2^{m1}	$\dots a_i^{m1} \dots$	a_n^{m1}	d_m
$m2$	a_1^{m2}	a_2^{m2}	$\dots a_i^{m2} \dots$	a_n^{m2}	
…	…	…	…	…	
mk_m	$a_1^{mk_m}$	$a_2^{mk_m}$	$\dots a_i^{mk_m} \dots$	$a_n^{mk_m}$	

Представлена модель структурно складається з *m* шарів ознак (сукупності вхідних інформаційних масивів) на певних проміжках часу та можливих варіантів функціонування АСУ (сукупності рішень). Прийняття рішення здійснюється на кожному етапі з урахуванням ознак функціонування АСУ [367–391].

Ієрархічні системи нечіткого логічного висновку. Для моделювання багатомірних залежностей “входи – виходи” доцільно використовувати ієрархічні системи нечіткого висновку. У таких системах вихід однієї бази знань подається на вхід іншої, більш високого рівня ієрархії. В ієрархічних базах знань відсутні зворотні зв’язки. На рисунку 1 наведено приклад ієрархічної нечіткої системи, яка моделює залежність $y = f(x_1, x_2, x_3, x_4, x_5, x_6)$ з допомогою трьох баз знань f_1, f_2, f_3.

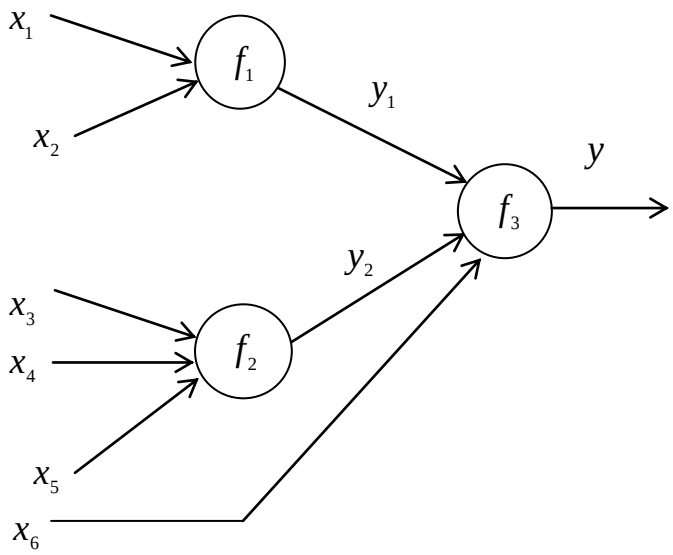

Рис. 1 – Приклад ієрархічної системи нечіткого виводу

Ці бази знань описують залежності $y_1 = f_1(x_1, x_2)$, $y_2 = f_2(x_4, x_5, x_6)$ та $y = f_3(y_1, x_3, y_2)$. Застосування ієрархічних нечітких баз знань дозволяє подолати “прокляття розмірності”. Ще однією перевагою ієрархічних баз знань – компактність. Невеликою кількістю нечітких правил в ієрархічних базах знань можна адекватно описати багатомірні залежності “входи – виходи”.

Нехай для лінгвістичної оцінки змінних використовується по п’ять термів. Тоді максимальна кількість правил для задавання залежності $y = f(x_1, x_2, x_3, x_4, x_5, x_6)$ за допомогою однієї бази знань буде складати $5^6 = 15625$. При нечіткому виводі за ієрархічною базою знань процедури дефазифікації і фазифікації для проміжних змінних y_1 та y_2 (рис. 1) не виконуються. Результат логічного висновку у вигляді нечіткої множини $\tilde{y}^* = \left(\frac{\mu_1(X^*)}{\tilde{d}_1}, \frac{\mu_2(X^*)}{\tilde{d}_2}, ..., \frac{\mu_m(X^*)}{\tilde{d}_m} \right)$ напряму передається до машини нечіткого висновку наступного рівня ієрархії. Тому для проміжних змінних в ієрархічних нечітких базах знань достатньо задати тільки терм-множини, без описання функції належностей.

8.5.2. Методика розподілу даних в автоматизованих системах управління

Сутність методики розподілу даних в АСУ полягає у раціональному розподілі даних між елементами АСУ, за рахунок відбору найбільш важливих джерел даних і раціональної кількості градацій, забезпечити задану ймовірність правильного розпізнавання типу даних, що в них циркулює.

Для вибору варіанту плану розподілу джерел даних між елементами АСУ використовуються часткові показники якості розподілу:

1. Повнота охоплення спостереженням елементів АСУ П , що розраховується як відношення суми коефіцієнтів важливості елементів АСУ Y_j, що включені в план розподілу, до суми коефіцієнтів важливості всіх елементів АСУ:

$$\text{П} = \frac{\sum_{j=1}^{\{m\}_u} Y_j}{\sum_{j=1}^{J} Y_j}, \qquad (1)$$

де $\{m\}_u$ – кількість елементів АСУ, вибрані для спостереження в u-му плану розподілу.

2. Витрати технічного ресурсу навантаження АСУ, що визначається як сума витрат технічних ресурсів АСУ $S_{заг_{АСУ}}$.

3. Ймовірність відстеження стану та характеру діяльності усієї сукупності елементів АСУ, що підлягають розподілу – $\overline{P}$.

$$\overline{P} = \frac{1}{m}\sum_{j=1}^{m} P_j, \tag{2}$$

де P_j – ймовірність відстеження стану та характеру діяльності елементів АСУ, що включені до плану розподілу.

Тоді система часткових показників якості вибору плану розподілу даних між елементами АСУ буде мати вигляд:

$$\begin{cases} \Pi \to \max \\ S_{заг_{АСУ}} \to \min. \\ \overline{P} \to \max \end{cases} \tag{3}$$

З урахуванням системи часткових показників, функціонал, що відображає якість розподілу даних між елементами АСУ за умови вибору того чи іншого варіанту плану розподілу Π, можна записати у вигляді:

$$F_{\Pi_{U_{опт}}} = \max F(\Pi_{\Pi_U}, S_{заг_{АСУ_{\Pi_U}}}, \overline{P}_{\Pi_U}). \tag{4}$$

Однак в існуючих методиках розподілу інформації за важливістю елементів АСУ розрахунок коефіцієнтів важливості здійснюється неявно (4), крім того, порядок їх розрахунку не визначено.

Отже, виникає актуальна наукова задача багатокритеріальної оптимізації процесу розподілу даних між елементами АСУ з урахуванням їх важливості для підвищення ефективності обробки даних.

Тоді (4) можна переписати у вигляді:

$$\mathrm{F}_{\Pi_{\mathrm{U}_{\mathrm{опт}}}} = \arg \max_{\mathrm{im}\in\mathrm{Im}} \mathrm{F}(\mathrm{Im}, \mathrm{S}_{\mathrm{заг}_{\mathrm{АСУ}_{\Pi_{\mathrm{U}}}}}, \overline{\mathrm{P}}_{\Pi_{\mathrm{U}}}), \tag{5}$$

де Im – вектор коефіцієнтів важливості (пріоритетності) елементів АСУ в смузі спостереження; $\overline{P}_{\Pi_U}$ – ймовірність відстеження стану та характеру діяльності елементу АСУ при виборі u -го плану розподілу.

Важливість елементу АСУ можна розглядати як неметричний критерій корисності (НКК). Основною складністю при розв'язанні поставленої задачі є представлення НКК у кількісному вигляді з метою його подальшого введення до функції корисності (ФК).

Для представлення НКК у кількісному вигляді визначені неметричні часткові критерії корисності (НЧКК), які мають характеризувати важливість елементу АСУ.

Основними НЧКК виступають ступінь пріоритетності задачі в інтересах якої ведеться розподіл, або ступінь пріоритету елементу АСУ ($X_{пріор}$); ступінь інформативної цінності ($X_{інф}$); ступінь оперативної цінності елементу АСУ ($X_{опер}$).

Позначимо $Q(X_{пріор}, X_{інф}, X_{опер})$, як функцію корисності НЧКК.

$X_{пріор}, X_{інф}, X_{опер}$ незалежні системи величин. Тоді функції корисності НЧКК можна представити системою виразів:

$$\begin{cases} \psi_{пріор} = Q(X_{пріор})\, f(X_{пріор}), \\ \psi_{інф} = Q(X_{інф})\, f(X_{інф}), \\ \psi_{опер} = Q(X_{опер})\, f(X_{опер}), \end{cases} \quad (6)$$

де $f(X_{пріор})$, $f(X_{інф})$, $f(X_{опер})$ – функції залежності корисності від метричних критеріїв.

У свою чергу функція корисності елементу АСУ матиме вираз:

$$\psi = Q(X_{пріор}, X_{інф}, X_{опер})\, f(X_{пріор}, X_{інф}, X_{опер}). \quad (7)$$

Для дослідження впливу неметричних критеріїв введемо обмеження, сутність якого полягає в тому, що вплив метричних критеріїв є рівноцінним, тобто (немає залежності від метричних критеріїв):

$$f(X_{пріор}) = f(X_{інф}) = f(X_{опер}) = 1. \quad (8)$$

Аналіз обмеження (8) показує, що показники між собою рівноцінні за метричним критерієм. В свою чергу функції залежності корисності від неметричних критеріїв змінюються за лінійним законом і визначаються нижнім і верхнім значенням прийнятих оцінок. Здійснивши операцію нормування за максимальним значенням, виходячи із прийнятої шкали будь-яка перевага одного із показників виразу (8) за неметричним критерієм призведе до домінування функції корисності відповідного показника.

Враховуючи (8), вираз (7) буде представлений, як:

$$\psi = Q(X_{пріор}, X_{інф}, X_{опер}). \quad (9)$$

З метою вибору раціонального виду функції корисності $Q(X_{пріор}, X_{інф}, X_{опер})$ зручно представити $X_{пріор}, X_{інф}, X_{опер}$ у вигляді нечітких множин [368, 369], а оцінки НЧКК, як їх елементи відповідно. Тоді $Q(X_{пріор}, X_{інф}, X_{опер})$ можна ототожнити з функцією належності набору вхідних значень показників основних НЧКК $x_{пріор}, x_{інф}, x_{опер}$ до нечітких множин $X_{пріор}, X_{інф}, X_{опер}$, відповідно. Таким чином, задачу визначення важливості елементу АСУ можна сформулювати як задачу прийняття рішення щодо важливості елементу АСУ, а результат процесу прийняття рішення можна представити у вигляді:

$$\mathrm{Im} = Q(x_{пріор}, x_{інф}, x_{опер}), \quad (10)$$

де $x_{пріор}, x_{інф}, x_{опер}$ – набір вхідних значень показників основних НЧКК; Im – рішення щодо визначення важливості елементу АСУ.

Завдання прийняття рішення щодо визначення важливості елементу АСУ полягає в тому, щоб на основі інформації про вектор вхідних показників ($x_{пріор}, x_{інф}, x_{опер}$) визначити результат Im . Необхідною умовою формального розв’язання поставленої задачі є наявність залежності (10). Для встановлення такої залежності необхідно розглядати вхідні показники (НЧКК) і вихідне рішення як лінгвістичні змінні, що задані на універсальних множинах.

Для оцінювання таких лінгвістичних змінних пропонується використовувати якісні терми, що складають терм-множину:

$X_{інф} = \{H, нС, С, вС, В\}$ – терм-множина змінної $x_{інф}$,

$X_{опер} = \{H, нС, С, вС, В\}$ – терм-множина змінної $x_{опер}$,

$X_{пріор} = \{H, С, В\}$ – терм-множина змінної $x_{пріор}$,

$\text{Im} = \{H, нС, С, вС, В\}$ – терм-множина змінної Im,

де $H, нС, С, вС, В$ – відповідно “низький”, “нижче середнього”, “середній”, “вище середнього”, “високий”;

Im – множина змінних, що характеризують важливість елементу АСУ:

$$X_{інф} = [1,5],$$
$$X_{опер} = [1,5], \quad (11)$$
$$X_{пріор} = [1,3],$$
$$\text{Im} = [1,5].$$

Для оцінювання значень лінгвістичних змінних $x_{пріор}, x_{інф}, x_{опер}$, відповідно (11) використаємо відповідну шкалу якісних термів.

У відповідності з методиками когнітивної інженерії щодо синтезу баз знань отримано бази знань, що характеризують важливість е. Використовуючи математичний апарат теорії нечітких множин базу знань перетворено в логічні рівняння:

$$\mu^{\text{Im}_j}(X_{пріор}, X_{інф}, X_{опер}) = \max_J \left\{ \min_i \left[\mu^J(x_{i(пріор)}), \mu^J(x_{i(інф)}), \mu^J(x_{i(опер)}) \right] \right\}, \quad (12)$$

де μ – функції належності відповідних лінгвістичних змінних $x_{пріор}, x_{інф}, x_{опер}, \text{Im}$, множинам $X_{пріор}, X_{інф}, X_{опер}$, Im. $J \in \{H, нС, С, вС, В\}$; $x_{i(пріор)} \in [1,2,3]$; $x_{i(інф)} \in [1,2,3,4,5]$; $x_{i(опер)} \in [1,2,3,4,5]$; $\text{Im}_i \in [1,2,3,4,5]$.

З аналізу числових результатів експерименту зроблено висновок, що прийняття рішення щодо важливості елементів АСУ визначається виразом:

$$\frac{x_{інф} + x_{опер}}{2} \cdot x_{пріор} = \text{Im}. \quad (13)$$

Мінімальне значення, що може приймати вираз (13) становить $\text{Im} = 1$, тоді максимальне $\text{Im} = 15$. Визначимо ФН до термів-множин важливості елементів АСУ.

Для цього приведемо інтервали вимірів кожної змінної до одного універсального інтервалу [305,309] за допомогою співвідношення:

$$\mu^j(\text{Im}_i) = \tilde{\mu}^j(u), u = 4\frac{\text{Im}_i - \underline{\text{Im}}}{\overline{\text{Im}} - \underline{\text{Im}}}, j = H, нС, С, вС, В, \quad (14)$$

Аналітична модель функції належності представлена виразом:

$$\mu^j(u) = \frac{1}{1 + \left(\frac{u-b}{c}\right)^2}, \quad (15)$$

де параметри b і c задаються за результатами попередньої оцінки функціонування АСУ.

Графічне зображення ФН згідно з виразом (15) наведено на рис. 2

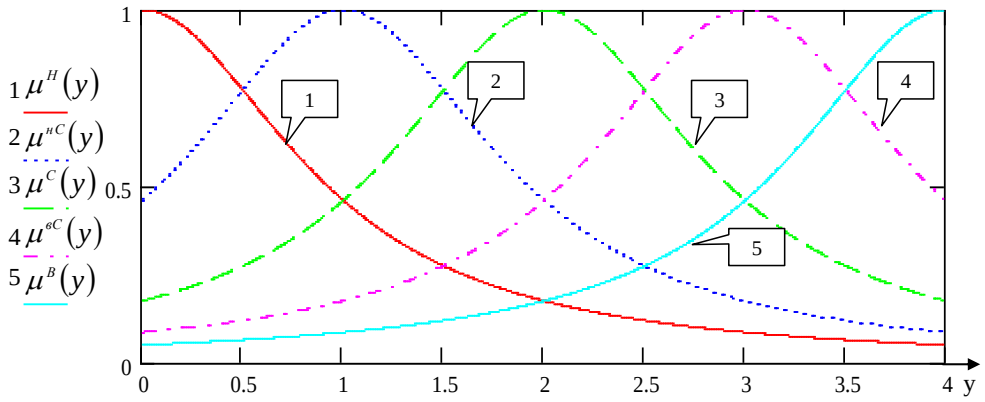

Рис. 2 – Графічне зображення ФН нечіткої множини ступенів важливості елементу АСУ

Розрахунок коефіцієнтів PI. Визначаються коефіцієнт пріоритетності $x_{пріор}$, коефіцієнт ступеню інформативної цінності $x_{інф}$, коефіцієнт ступеню оперативної цінності $x_{опер}$.

Розрахунок важливості ОР. За допомогою (13) розраховуються значення важливості відповідного елементу АСУ.

Визначення лінгвістичного значення важливості елементу АСУ.

Оптимізація вектору ознак функціонування елементу АСУ.

Відомо, що зі збільшенням кількості ознак, які характеризують процес функціонування елементу АСУ, збільшуються час необхідний для ідентифікації режиму його функціонування та інші затрати, в першу чергу апаратні, в наслідок чого знижується оперативність процесу оцінювання.

Виділення інформативних ознак у реальній обстановці є складним завданням, особливо при веденні оцінювання процесу функціонування елементів системи АСУ, коли набір ознак може бути дуже великим, а самі ознаки корельовані між собою. Тому постає завдання відбору й виділення групи найбільш інформативних ознак з метою зменшення розмірності вектору вихідних даних при одночасному знаходженні такої системи координат, у якій ймовірність правильного розпізнавання елементів АСУ буде максимальною або достатньою для прийняття рішення.

Зменшення розмірності простору ознак в умовах великої кількості елементів системи АСУ відіграє значну роль, оскільки збільшує пропускну спроможність каналів системи АСУ у цілому. Тому, що збільшення кількості ознак, що характеризують елемент системи АСУ значною мірою призводить до збільшення похибок ідентифікації.

Залежність імовірності ідентифікації процесу функціонування елементу системи АСУ від розмірності простору ознак наведено на рисунку 4.

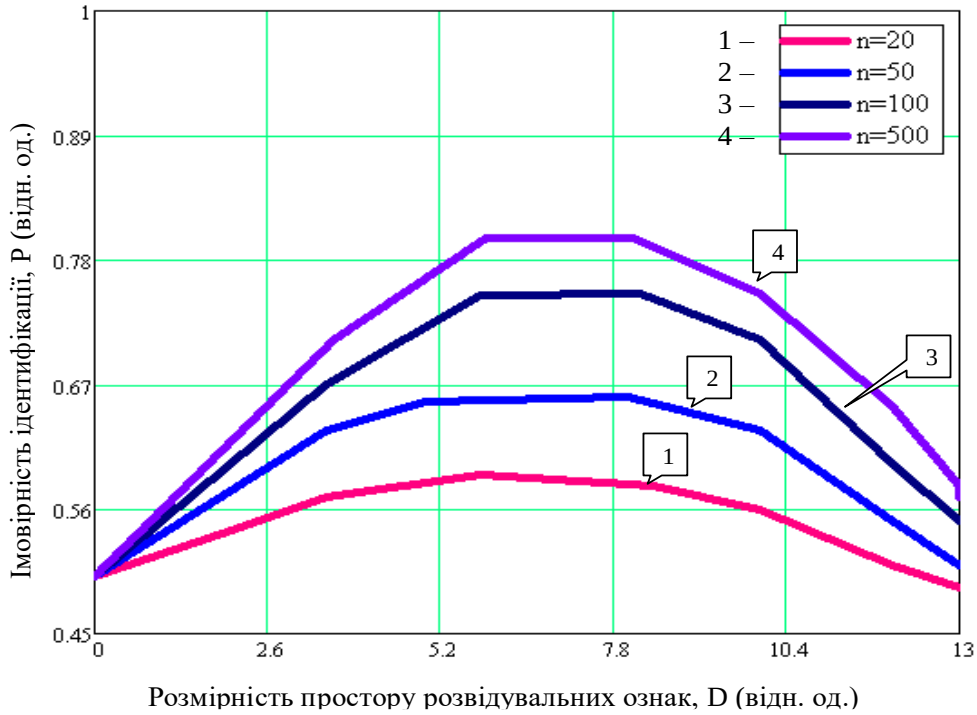

Рис. 3 – Залежність імовірності ідентифікації режиму функціонування елементу АСУ від розмірності простору ознак

З наведених графіків видно, що довільне збільшення розмірності простору ознак може призвести до погіршення ймовірності правильного розпізнавання.

Формування вектору ознак елементу АСУ математично може бути представлено у вигляді:

$$Y = AX, \tag{16}$$

де X – вектор ознак, що характеризують процес функціонування елементу АСУ; Y – вектор можливих рішень; A – матриця перетворень.

Перевірка оптимальної розмірності вектору ознак. Умовою закінчення циклу ліквідації градацій є величина порогу втрат інформативності за усіма ознаками $\left(\sum \Delta I_k\right)_{\max}$ або за конкретною ознакою. Можна також задати максимальну кількість градацій, яку необхідно залишити в процесі мінімізації.

На рис. 2 наведені криві зміни інформативності ознак залежно від кількості їх градацій. Аналіз цих кривих за усіма ознаками дає можливість мінімізувати кількість градацій в аспекті витрат пам’яті пристрою ідентифікації і сумарних витрат інформативності параметрів ідентифікації. Наведені

залежності свідчать, що кількість градацій ознак доцільно вибирати не більше 4-6 (по точці переломлення більшості кривих), що збігається з результатами наведеними на рисунку 3.

Проведення корегування параметрів системи. На підставі удосконаленої методики корегування параметрів, що заснована на використанні генетичного підходу, відбувається корегування параметрів системи.

8.5.3. Модель процесу оцінювання обробки різнотипних даних в АСУ з використанням експертної інформації.

Нехай відомі: множина рішень $D = \{d_j\}, (j = \overline{1,m})$, що відповідає результату оцінювання обробки різнотипних даних в АСУ y; множина вхідних показників $X = \{x_i\}, (i = \overline{1,n})$; діапазони кількісної зміни кожного вхідної інформації $x_i \in \left[\underline{x}_i, \overline{x}_i \right], i = \overline{1,n}$; функції належності, що дозволяють представити показники $x_i, i = \overline{1,n}$ у вигляді нечітких множин; матриця знань, визначена за правилами (Табл. 1). Графічно її можна відобразити у вигляді рис. 4.

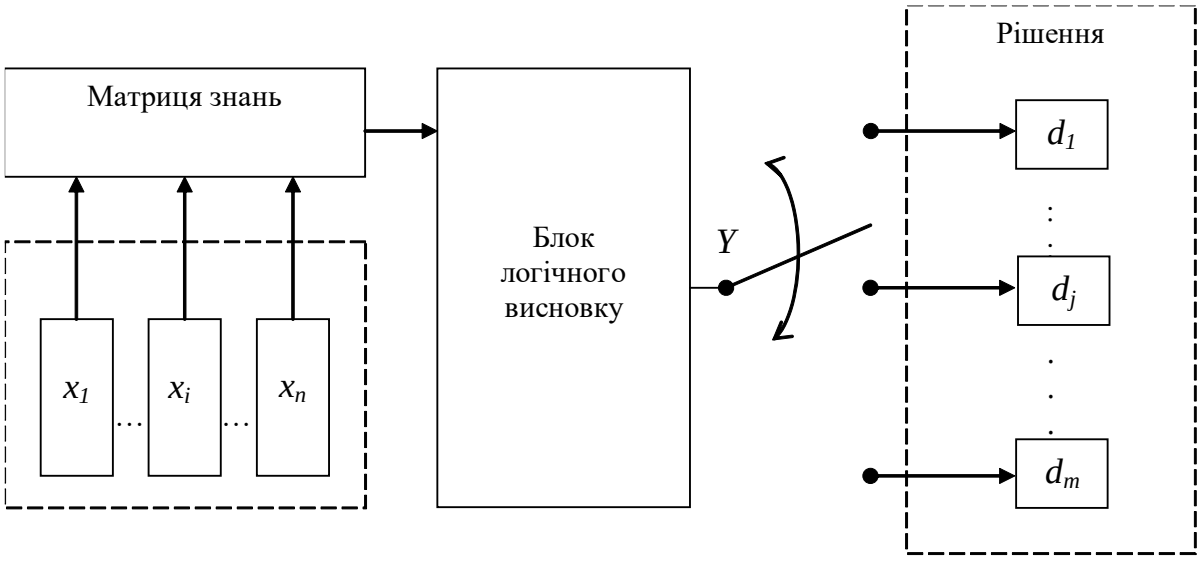

Рис. 4 – Модель процесу оцінювання оперативної обстановки

Розглянемо застосування моделі використання експертної інформації для

синтезу алгоритму оцінювання обробки різнотипних даних в АСУ.

З аналізу функціонування елементів АСУ в різних умовах обстановки визначені напрямки оцінювання: подібність показників обстановки та їх зміни в ході функціонування АСУ.

Опишемо модель оцінювання обробки різнотипних даних в АСУ:

$$D(k) = f\begin{bmatrix} Y_1(k-1), Y_2(k-1), ..., Y_{14}(k-1), Q(k-1), R(k-1), \\ Z_1(k-1), ..., Z_4(k-1) \end{bmatrix}, \quad (17)$$

де $Y_1(k-1)$ – вектор, що характеризує режим функціонування елементу АСУ №1 на k-1 кроці моделювання; $Y_2(k-1)$ – вектор, що характеризує режим функціонування елементу АСУ на k-1 кроці моделювання; $Y_{14}(k-1)$ – вектор, що характеризує режим функціонування елементу АСУ №14 на k-1 кроці моделювання; $Q(k-1)$ – вектор, що характеризує режим функціонування системи управління та зв'язку елементу АСУ №1; $R(k-1)$ – вектор, що характеризує режим функціонування системи управління та зв'язку елементу АСУ; $Z_1(k-1), ..., Z_4(k-1)$ – вектори, що характеризують режими функціонування систем управління та зв'язку групових елементів АСУ.

У свою чергу вектори процесу оцінювання обробки різнотипних даних в АСУ визначаються показниками: $Y_1, ..., Y_{14}, Q, R, Z_1, ..., Z_4 = \{k_{11}(x), ..., k_{145}(x)\}$.

Для показників, що мають кількісний вимір, діапазон зміни розбивається на чотири кванти. Це забезпечить можливість перетворення безперервної універсальної множини $U = [\underline{u}, \bar{u}]$ в дискретну п'ятиелементну множину:

$$U = \{u_1, u_2, ..., u_5\},$$

де $u_1 = \underline{u}$, $u_2 = \underline{u} + \Delta_1$, $u_3 = u_2 + \Delta_2$, $u_4 = u_3 + \Delta_3$, $u_5 = \bar{u}$, причому $\Delta_1 + \Delta_2 + \Delta_3 + \Delta_4 = \bar{u} - \underline{u}$, $\bar{u}(\underline{u})$ – верхня (нижня) границя діапазону зміни показника. Тоді всі матриці парних порівнянь мають розмірність 5×5. Вибір чотирьох квантів визначається можливістю апроксимації нелінійних кривих по п'ятьох точках.

Для оцінки значень лінгвістичних змінних $k_{11},...,k_{21},...,k_{141},...k_{145}$ будемо використовувати наступну шкалу якісних термів.

У загальному випадку вхідні змінні $x_1, x_2,..., x_n$ можуть задаватися числом, лінгвістичним термом або за принципом термометра.

Оцінювання процесу обробки різнотипних даних в АСУ з використанням експертної інформації здійснюється з використанням нечітких логічних рівнянь, які являють собою матрицю знань і систему логічних висловлювань. Ці рівняння дозволяють обчислити значення функцій належності різних результатів ідентифікації при фіксованих значеннях вхідних показників. В якості результату процесу оцінювання процесу обробки різнотипних даних в АСУ, будемо приймати рішення з найбільшим значенням функції належності.

Лінгвістичні оцінки α_i^{jp} змінних $x_1, x_2,..., x_n$, що входять у логічні висловлення щодо рішень d_j, $j = \overline{1,m}$, розглянемо як нечіткі множини, визначені на універсальних множинах $X_i = \left[\underline{x}_i, \overline{x}_i \right], i = \overline{1,n}$.

Нехай $\mu^{a_i^{jp}}(x_i)$ – ФН показника $x_i \in \left[\underline{x}, \overline{x} \right]$ нечіткому терму $\alpha_i^{jp}, i = \overline{1,n},\ j = \overline{1,m},\ p = \overline{1,l_i}$; $\mu^{d_j}(x_1, x_2,..., x_n)$ – ФН вектора вхідних змінних $X = (x_1, x_2,..., x_n)$ значенню вихідної оцінки $y = d_j$, $j = \overline{1,m}$.

Зв'язок між даними функціями визначається нечіткою базою знань і може бути представлений у вигляді наступних логічних рівнянь:

$$\begin{aligned} \mu^{d_j}(x_1, x_2,..., x_n) = \mu^{a_1^{j1}}(x_1) \wedge \mu^{a_2^{j1}}(x_2) \wedge ... \wedge \mu^{a_n^{j1}} \vee \\ \mu^{a_1^{j2}}(x_1) \wedge \mu^{a_2^{j2}}(x_2) \wedge ... \wedge \mu^{a_n^{j2}}(x_n) ... \\ ... \mu^{a_1^{jl_j}}(x_1) \wedge \mu^{a_2^{jl_j}}(x_2) \wedge ... \wedge \mu^{a_n^{jl_j}}(x_n),\ j = \overline{1,m}. \end{aligned} \tag{18}$$

Рівняння отримані з нечіткої бази знань шляхом заміни змінних (лінгвістичних термів) на їх ФН, а операції І та ЧИ – на операції $\wedge$ і $\vee$.

Коротко систему (18) запишемо таким чином:

$$\mu^{d_j}(x_i) = \bigvee_{p=1}^{l_j} \left[\bigwedge_{i=1}^{n} \mu^{a_i^{jp}}(x_i) \right],\ j = \overline{1,m}. \tag{19}$$

Нечіткі логічні рівняння є аналогом введеної Заде процедури нечіткого логічного висновку, що здійснюється за допомогою операції “нечітка (min-max) композиція”, в якій операціям $\wedge$ і $\vee$ відповідають операції min і max, з (19) одержуємо:

$$\mu^{d_j}(x_i) = \max_{p=\overline{1,l_j}} \left\{ \min_{j=\overline{1,n}} \left[\mu^{a_i^{jp}}(x_i) \right] \right\}. \qquad (20)$$

З виразу (20) видно, що для розрахунку необхідно мати лише ФН змінних нечітким термам.

8.5.4 Удосконалена методика настроювання інформаційної системи оцінювання процесу обробки різнотипних даних в АСУ в умовах невизначеності

Сутність методики настроювання інформаційної системи оцінювання процесу обробки різнотипних даних в АСУ в умовах невизначеності полягає у підборі вагових коефіцієнтів продукційних правил при яких помилка між еталонним та експериментальним рішеннями буде мінімальною.

Ідентифікація на основі нечіткого логічного висновку здійснюється відповідно до визначеної бази знань:

ЯКЩО $(x_1 = a_{1,j1})$ I $(x_2 = a_{2,j1})$ I ...I $(x_n = a_{n,j1})$ з вагою w_{j1},

АБО $(x_1 = a_{1,j2})$ I $(x_2 = a_{2,j2})$ I ...I $(x_n = a_{n,j2})$ з вагою w_{j2},

.....

АБО $(x_1 = a_{1,jk_j})$ I $(x_2 = a_{2,jk_j})$ I ...I $(x_n = a_{n,jk_j})$ з вагою w_{jk_j}, (21)

ТО $y = d_j$, $j = \overline{1,m}$,

де $a_{i,jp}$ – нечіткий терм, яким оцінюється змінна x_i в рядку з номером $jp(p = \overline{1,k_j})$, тобто $a_{i,jp} = \int \mu_{jp}(x_i)/x_i$;

k_j – кількість рядків-кон'юнкцій, в яких вихід y оцінюється значенням d_j;

$w_{jp} \in [0,1]$ – ваговий коефіцієнт правила з номером jp.

Функції відповідності процесу обробки різнотипних даних $X^* = (x_1^*, x_2^*, ..., x_n^*)$ класам d_j розраховуються так:

$$\mu_{d_j}(X^*) = \underset{p=\overline{1,k_j}}{\vee} w_{jp} \cdot \underset{i=\overline{1,n}}{\wedge} \left(\mu_{jp}(x_i^*)\right), j = \overline{1,m}, \quad (22)$$

де $\mu_{jp}(x_i^*)$ – функція відповідності входу x_i^* нечіткому терму $a_{i,jp}$; $\wedge(\vee)$ – s-норма (t-норма), якій в задачах класифікації зазвичай відповідають максимум (мінімум).

В якості рішення вибрано клас з максимальною функцією відповідності розрахованого рішення $d_1 ... d_m$:

$$y^* = \underset{\{d_1, d_2, ..., d_m\}}{\arg} \max\left(\mu_{d_1}(X^*), \mu_{d_2}(X^*), ..., \mu_{d_m}(X^*)\right). \quad (23)$$

Таким чином, буде виконана адаптація чи настроювання інформаційної системи оцінювання процесу обробки різнотипних дани в умовах невизначеності.

В роботі застосований спосіб адаптації, заснований на рішенні задачі оптимізації з використанням методу генетичного алгоритму.

Введемо обмеження, що існує еталонна вибірка із M пар експериментальних даних, які зв'язують входи $X = (x_1, x_2, ..., x_n)$ з виходом y залежності, що досліджується:

$$(X_r, y_r), \left(r = \overline{1, M}\right), \quad (24)$$

де $X_r = (x_{r,1}, x_{r,2}, ..., x_{r,n})$ – вхідний вектор в r-ій парі і y – відповідний вихід.

Настроювання моделі представляє собою знаходження таких параметрів функцій відповідності термів вхідних змінних і вагових коефіцієнтів правил, які мінімізують відхилення між очікуваним і отриманим результатом на еталонній вибірці. Критерій близькості можна визначити різними способами.

Перший спосіб полягає у виборі як критерію настройки відсоток помилок класифікації на еталонній вибірці, яка використовується для навчання системи. Введемо наступні позначення:

P – вектор параметрів функцій відповідності термів вхідних і вихідної змінних;

W – вектор вагових коефіцієнтів правил бази знань;

$F(X_r, P, W)$ – результат виводу відповідно базі знань з параметрами (P, W) при значенні входів X_r.

Тоді настроювання нечіткої моделі зводиться до наступної задачі оптимізації: знайти такий вектор (P, W), щоб:

$$\frac{1}{M}\sum_{r=1,M}\Delta_r \to \min, \tag{25}$$

де Δ_r – помилка класифікації стану обробки різнотипних даних X_r:

$$\Delta_r = \begin{cases} 1, якщо \quad y_r \neq F(X_r, P, W) \\ 0, якщо \quad y_r = F(X_r, P, W) \end{cases}. \tag{26}$$

Перевага критерію настроювання полягає в його простоті і ясній змістовній інтерпретації. Відсоток помилок широко використовується як критерій навчання різних систем розпізнавання образів.

Цільова функція задачі оптимізації приймає дискретні значення, що ускладнює використання градієнтних методів оптимізації. Особливо важко підібрати необхідні параметри градієнтних алгоритмів (наприклад, приріст аргументів для розрахунку часткових похідних) при настройці нечіткого класифікатора на невеликій виборці даних.

Другий спосіб передбачає використання як критерію настройки відстань між результатом виводу у вигляді нечіткої множини $\left(\frac{\mu_{d_1}(x)}{d_1}, \frac{\mu_{d_2}(x)}{d_2}, ..., \frac{\mu_{d_m}(x)}{d_m}\right)$ і значенням вихідної змінної в еталонній виборці, яка призначена для навчання системи. Для цього вихідна змінна y еталонної вибірки (23) фазифікується наступним чином:

$$\left.\begin{array}{l}\tilde{y}=(1/d_1,0/d_2,...,0/d_m),\ якщо\quad y=d_1\\ \tilde{y}=(0/d_1,1/d_2,...,0/d_m),\ якщо\quad y=d_2\\ .....\\ \tilde{y}=(0/d_1,0/d_2,...,1/d_m),\ якщо\quad y=d_m\end{array}\right\}. \tag{27}$$

В даному випадку настроювання нечіткого класифікатора зводиться до наступної задачі оптимізації: знайти такий вектор (P,W), щоб*:*

$$\frac{1}{M}\cdot\sum_{r=1}^{M}\sum_{j=1}^{m}\left(\mu_{d_j}(y_r)-\mu_{d_j}(X_r,P,W)\right)^2\to\min, \tag{28}$$

де $\mu_{d_j}(y_r)$ – функція відповідності значення вихідної змінної y в r-ій парі еталонної вибірки до рішення d_j; $\mu_{d_j}(X_r,P,W)$ – функція відповідності виходу нечіткої моделі з параметрами (P,W) до рішення d_j, при значеннях входів із r-ої пари еталонної вибірки (X_r).

Цільова функція в задачі (27) не має протяжних плато, тому вона придатна до оптимізації градієнтними методами. Однак, результати оптимізації не завжди задовільні: нечітка база знань, що забезпечує мінімум критерію (27), не завжди задовольняє також і мінімум помилок класифікації. Це пояснюється тим, що точки, які близькі до максимумів розділів класів, вносять зазвичай однаковий внесок до критерію настройки, як при правильній, так і при помилковій класифікації.

Третій спосіб успадковує переваги попередніх способів. Ідея полягає в тому, що щоб внесок помилково класифікованих об’єктів до критерію настроювання збільшувати, шляхом множення відстані $\sum_{j=1}^{m}\left(\mu_{d_j}(y_r)-\mu_{d_j}(X_r,P,W)\right)^2$ на штрафний коефіцієнт. В результаті задача оптимізації приймає вигляд:

$$\frac{1}{M}\cdot\sum_{r=1}^{M}(\Delta_r\cdot penalty+1)\cdot\sum_{j=1}^{m}\left(\mu_{d_j}(y_r)-\mu_{d_j}(X_r,P,W)\right)^2\to\min, \tag{29}$$

де $penalty>0$ – штрафний коефіцієнт.

Задачі (26), (28) і (29) можуть бути розв'язані різними технологіями оптимізації, серед яких часто застосовується метод найшвидшого спуску, квазіньютонівські методи і генетичні алгоритми.

На змінні, якими управляють, зазвичай, накладають обмеження, що забезпечують лінійну упорядкованість елементів терм-множин. Крім того, ядра нечітких множин не повинні виходити за межі діапазонів зміни відповідних змінних. Це забезпечує прозорість, тобто змістовну інтерперабельність нечіткої бази знань після настроювання. Що стосується вектору W, то його координати повинні знаходитись в діапазоні [305,306]. Якщо до рівняння інтерперабельності бази знань пред'являються високі вимоги, то ваги правил не настроюють, залишаючи їх рівними 1. Можливим є і проміжний варіант, коли вагові коефіцієнти можуть приймати значення 0 і 1. В даному випадку нульове значення вагового коефіцієнту еквівалентно виключенню правила із нечіткої бази знань.

Параметри функцій відповідності і ваги правил можна настроювати одночасно або окремо. При настроюванні тільки ваг правил об'єм обчислень можна значно скоротити, так як функції належності $\mu_{jp}\left(x_i^*\right)$, не залежать від W. Для цього на початку оптимізації необхідно розрахувати ступені виконання правил при одиничних вагових коефіцієнтах $\left(w_{jp}\right)=1$ для кожного об'єкта із еталонної вибірки:

$$g_{jp}\left(X_r\right)=\underset{i=1,n}{\wedge}\mu_{jp}\left(x_{r,i}\right),\ j=\overline{1,m},\ p=\overline{1,k_j},\ r=\overline{1,M}\ .$$

Для нових вагових коефіцієнтів функцій відповідності процесу обробки різнотипних даних в АСУ X_r класам d_j розраховуються так:

$$\mu_{d_j}\left(X_r\right)=\underset{p=1,k_j}{\vee}w_{jp}\cdot g_{jp}\left(X_r\right),\ j=\overline{1,m}\ .$$

Враховуючи специфіку процесу процесу обробки різнотипних даних в АСУ, одним із шляхів її вирішення на основі нечіткої логіки є застосування комбінованих методів оптимізації, які сполучають у собі переваги роботи градієнтного методу та методу випадкового пошуку. Одним з методів такого

типу є метод “генетичного алгоритму”, який дозволяє проводити оптимізацію для полімодальних, негладких та невипуклих функцій зі швидкістю збіжності більшою, ніж в методах випадкового пошуку .

Таким чином, враховуючи визначені особливості процесу обробки різнотипних даних в АСУ, ієрархічність дерева логічного висновку, структурно-семантичну модель обробки різнотипних даних в АСУ, процес настроювання інформаційної системи обробки різнотипних даних доцільно здійснювати диференційовано, тобто шляхом настройки нечіткої бази знань кожного окремого елемента системи АСУ.

Дослідимо можливості застосування та функціонування “генетичного алгоритму” для настроювання інформаційної системи оцінювання оперативної обстановки.

Будемо вважати відомими наступні вихідні дані:

S – вектор структури системи, який визначає параметри системи, що не змінюються під час оптимізації (стосовно інформаційної системи оцінювання оперативної обстановки – набір правил “ЯКЩО-ТО”, представлених за допомогою математичного апарату нечіткої логіки, та коефіцієнтів ваг правил);

B – вектор-еталон, який містить набір зразкових пар стимул-реакцій (вхідні показники - рішення), за якими налаштовується інформаційна система оцінювання оперативної обстановки;

W_{jp} – вектор ваг правил нечіткої бази знань, значення якого оптимізується;

F – функція невідповідності, що визначає якість рішення, яке пропонує інформаційна система оцінювання оперативної обстановки у порівнянні з рішенням у векторі - еталоні.

$F_{ACУ}$ – функція невідповідності, що визначає якість рішення, яке пропонує інформаційна система обробки різнотипних даних щодо окремого елемента АСУ у порівнянні з рішенням у векторі - еталоні.

Під прийняттям рішення інформаційною системою обробки різнотипних даних в АСУ будемо розуміти той вихідний результат, який видає система відповідно введених ознак $x_{11}...x_{nm}$.

Введемо наступні обмеження:

вектор *B* охоплює всю практично значиму область рішень на області застосування інформаційної системи обробки різнотипних даних в АСУ;

вектор *S* (правила "ЯКЩО – ТО") сформований заздалегідь за результатами роботи з експертами та не містить логічних помилок;

за результатами роботи з експертами визначено значення параметрів функцій належності;

функція невідповідності – *F* розраховує помилку прийняття рішення інформаційною системою процесу обробки різнотипних даних в АСУ методом найменших квадратів з використанням порядкової шкали за формулою:

$$e_{АСУ} = \sum_{j=1}^{n} \sum_{i=1}^{k} (d_i^{B_j} - d_i^{O_j})^2, \quad (30)$$

де $e_{ДРВ}$- помилка настроювання окремого елементу АСУ; *n* – розмірність вектору *B*; *к* – максимальна кількість рішень, які видає інформаційна система обробки різнотипних даних в АСУ; $d_i^{B_j}$ - еталонне *i*-те рішення на *j*-ий вхідний елемент вектору *B*; $d_i^{O_j}$ - *i*-те рішення інформаційної системи обробки різнотипних даних в АСУ на *j*-ий вхідний елемент вектору *B* з урахуванням ваг правил W_{jp}; *i* – номер рішення, який видає інформаційна система оцінювання оперативної обстановки, $i \in \overline{1,k}$; *j* – номер набору вхідних показників в еталонному векторі В.

Для настроювання набору баз знань (НБЗ) окремого елементу системи АСУ запропонований критерій:

$$e_{АСУ\_\min} = Min(F_{АСУ}(S,W)), \quad (31)$$

де $e_{ДРВ\min}$ – мінімально допустима кінцева сумарна помилка, як різниця значень ФН рішення щодо стану функціонування окремого елементу системи АСУ та еталонного рішення.

В результаті використання критерію (30) для кожного окремого елементу АСУ логічним є використання критерію (31), що свідчить про настроювання інформаційної системи обробки різнотипних даних в АСУ в цілому.

$$e_{\Sigma \min} = Min(F(S,W)), \quad (32)$$

де $e_{\Sigma \min}$ – мінімально допустима кінцева сумарна помилка як різниця значень ФН рішення щодо стану функціонування системи управління та зв'язку елементу системи АСУ та еталонного рішення.

Настроювання ваг правил (W_{jp}) та параметрів функцій належності – вектору P проведемо за допомогою методу "генетичного алгоритму".

Сукупність показників, які оптимізуються, об'єднуються у вектор параметрів, який має назву хромосома. Показники в хромосомі можуть зберігатись у звичайному та закодованому (перетвореному) вигляді. Окрема ділянка хромосоми, яка відповідає за кодування одного показника, має назву ген; довжина гену залежить від обраного типу кодування.

Кожна хромосома являє собою рішення задачі, яке оптимізується з ефективністю, що виражається певним числом, яке обчислюється за цільовою функцією. Набір хромосом (сукупність рішень) називається популяція. В популяції підтримується постійна кількість хромосом. Основні етапи роботи методу генетичного алгоритму наведено на рис. 5.

Рис. 5 – Схема роботи методу генетичного алгоритму настроювання

Формування початкової популяції залежить від підходу формування хромосом: настроювання вагових коефіцієнтів важливості обстановки; настроювання вагових коефіцієнтів пріоритетності логічних правил БЗ за якими приймаються рішення.

З метою зменшення об’єму масивів даних запропоновано формувати хромосоми де генами є вагові коефіцієнти логічних правил БЗ.

Формування початкової популяції. Вектор-еталон *B* сформований за результатами оцінювання оперативної обстановки в різних умовах функціонування елементів системи АСУ.

Визначимо ФН ознак функціонування елементу системи АСУ та рішень нечітким термам за формулою:

$$\mu^{d_j}(x_i) = \max_{p=\overline{1,l_j}} \left\{ W_{jp} \min_{j=\overline{1,n}} \left[\mu^{a_i^{jp}}(x_i) \right] \right\}. \quad (33)$$

Згенеруємо довільним чином ваги правил W_{jp} НБЗ в інтервалі від 0 до 1.

Визначимо ефективності хромосом. За цільовою функцією визначається: ефективність кожної хромосоми в популяції для кожного набору ваг W_{jp}, після чого проводиться сортування хромосом за зростанням (зменшенням) їх ефективності (формування ряду за величиною похибки).

Відбір. На цьому етапі відкидаються хромосоми з найменшою ефективністю, якщо загальна кількість хромосом в популяції більша за допустиму. Тобто об’єм обчислень, що здійснюється в алгоритмі, зберігається постійним незалежно від номеру ітерації.

Схрещування. З популяції з урахуванням ефективності випадково обираються дві хромосоми та, починаючи з випадкової позиції, обмінюються генами, позицій схрещування може бути й декілька. Якщо у просторі пошуку дві точки, визначені цими хромосомами, знаходяться в околі одного екстремуму, то середнє значення між цими точками, що є результатом схрещування, буде знаходитись ближче до екстремуму, тобто це у певній мірі аналогія градієнтного методу, а якщо дві точки, визначені цими хромосомами, знаходяться в околі різних екстремумів, то середнє значення між цими точками буде випадковим, тобто це аналогія методу випадкового пошуку.

Мутація. Повна аналогія до методу випадкового пошуку – в популяції випадково змінюється значення окремих генів, тобто положення точки пошуку в просторі пошуку.

Закінчення пошуку. Пошук завершується за умови, що протягом L ітерацій ефективність найкращої хромосоми збільшилась менше ніж на λ, інакше

починається нова ітерація. Окрему ітерацію називають поколінням. Наприклад, якщо найефективніша хромосома була знайдена у 30 поколінні, то умова закінчення алгоритму пошуку була виконана на 30 ітерації, а найкращу хромосому у кінцевій виборці вважають оптимальним рішенням.

Таким чином, методику застосування генетичних алгоритмів настроювання інформаційної системи обробки різнотипних даних в АСУ зводиться до наступного алгоритму (рис. 6):

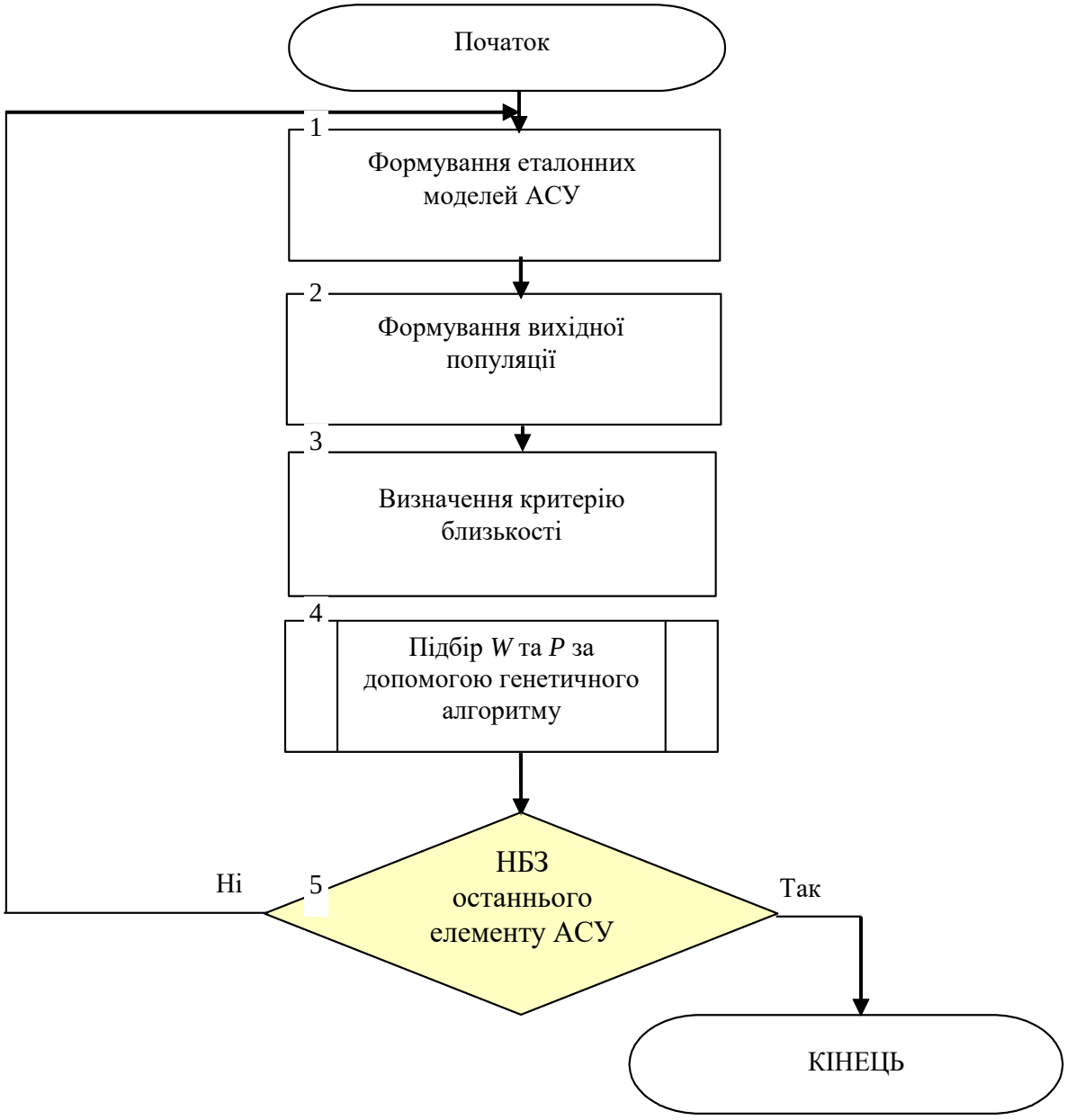

Рис. 6 – Алгоритм застосування ГА для настроювання інформаційної системи обробки різнотипних даних в АСУ

1. Формування еталонних станів процесу обробки різнотипних даних в АСУ.

2. Формування вихідної популяції.

3. Визначення критерію близькості настроювання нечіткого ідентифікатора.

4. Настроювання ваг продукційних правил НБЗ кожного елементу АСУ.

Формування початкової популяції. На

При розрахунку ваги правил НБЗ приймаємо рівними одиниці.

Генеруємо довільним чином ваги правил НБЗ в інтервалі від 0 до 1, тобто $W_{jp} \in [0,1]$. Оскільки правил у НБЗ 43 і кількість хромосом дорівнює 14 то необхідно сформувати двомірний масив 14×43 наборів ваг правил для настроювання інформаційної системи оцінювання оперативної обстановки. Набори ваг правил наведено в таблиці Ж.1.

За цільовою функцією визначається: ефективність кожної хромосоми в популяції для кожного набору ваг W_{jp}, після чого проводиться сортування хромосом за зростанням (зменшенням) їх ефективності.

Відбір. На цьому етапі відкидаються хромосоми з найменшою ефективністю, якщо загальна кількість хромосом в популяції більша за допустиму. Тобто об’єм обчислень, що здійснюється в алгоритмі, зберігається постійним незалежно від номеру ітерації.

Проводимо сортування хромосом з урахуванням їх ефективності:

$\Delta_3, \Delta_9, \Delta_{12}, \Delta_{13}, \Delta_2, \Delta_4, \Delta_1, \Delta_{14}, \Delta_8, \Delta_7, \Delta_5, \Delta_{11}, \Delta_6, \Delta_{10}$

Найбільшу ефективність, тобто найменшу помилку прийняття рішення, мають третя та дев’ята хромосоми, а найменшу ефективність, тобто найбільшу помилку прийняття рішення, мають шоста та десята хромосоми.

Схрещування. В даному випадку для операції схрещування в якості батьківських хромосом обрано хромосому №3 та хромосому №9. В свою чергу результат схрещування – хромосоми-нащадки записуються замість хромосоми №6 та хромосоми №10, відповідно.

Мутація. Над „десятою” хромосомою виконуємо операцію мутації і, якщо сума ваг перевищує одиницю, проводимо нормалізацію *0,15; 0,09; 0,08; 0,02; 0,2; 0,13; 0,07; 0,06; 0,16; 0,04.*

Переходимо до розрахунку її ефективності. Отримуємо результат: Δ=1,1, що значно краще ефективності двох попередніх хромосом, оптимізацію яких проводимо.

Закінчення пошуку.

Висновки

1. За результатами проведеного в дослідженні аналізу встановлено, що застосування нечітких графів та математичного апарату нечіткої логіки в задачах підтримки прийняття рішень розподілу даних та оцінювання процесу обробки різнотипних даних в різних умовах, в тому числі і в умовах невизначеності дозволяє здійснювати розподіл даних між елементами АСУ за важливістю елементів АСУ та кількості ознак в реальному масштабі часу.

2. Удосконалено методику раціонального розподілу даних за важливістю елементів АСУ та кількістю ознак в АСУ в умовах невизначеності, яка відрізняється від відомих комбінуванням математичного апарату теорії інформації, нечіткої логіки та експертного оцінювання, що дозволило виконати формалізацію ознак в єдиному просторі параметрів та за рахунок інтелектуалізації процесів обробки інформації.

Проведено кількісну оцінку ефективності запропонованої методики. Отримані результати цієї оцінки показали, що розподіл даних між елементами АСУ за важливістю та кількістю ознак за допомогою запропонованої методики дозволяє підвищити оперативність обробки даних та прийняти рішення, щодо стану процесу обробки різнотипних даних на 15-17%.

3. Удосконалена методика настроювання інформаційної системи оцінювання процесу обробки різнотипних даних в АСУ в умовах невизначеності на основі використання генетичного алгоритму в умовах невизначеності, де застосування інших методів обмежене у зв’язку з неможливістю варіювання

окремим ознаками при фіксованих значеннях інших показників, дозволило підвищити оперативність розробленої інформаційної системи оцінювання обробки різнотипних даних в АСУ.

Науковим результатом є удосконалення генетичного алгоритму диференційованого настроюванням нечіткої бази знань інформаційної системи оцінювання обробки різнотипних даних в АСУ на основі апостеріорних даних.

Проведено кількісну оцінку ефективності удосконаленої методики. Отримані результати цієї оцінки показали, що використання зазначеної методики дозволяє підвищити оперативності настроювання інформаційної системи обробки різнотипних даних в АСУ в умовах невизначеності.