

С. Д. Штовба, д-р техн. наук, доц.

## МОДЕЛЮВАННЯ ЗАЛЕЖНОСТЕЙ ЗА ДОПОМОГОЮ НЕЧІТКОЇ БАЗИ ЗНАТЬ З НЕЧІТКИМИ РЕГРЕСІЙНИМИ РІВНЯННЯМИ

Запропоновано нову структуру нечіткої бази знань, в якій антецеденти задаються нечіткими термами, а консеквенти — лінійними залежностями «входи—вихід» з нечіткими коефіцієнтами. Запропоновано алгоритм логічного виведення за цією базою знань для випадку чітких та нечітких вхідних даних. На основі теорії нечіткої ідентифікації сформульовано математичну постановку задачі навчання запропонованої бази знань за експериментальними даними.

### Вступ та мета дослідження

Сьогодні все частіше моделювання складних залежностей в техніці, економіці, медицині, соціології, будівництві, сільському господарстві, військовій справі, спорті та в інших областях здійснюють за допомогою нечітких баз знань [1—3]. Нечіткою базою знань називається сукупність нечітких правил «Якщо—Тоді», яка задає взаємозв'язок між входами та виходами досліджуваного об'єкта. Найпоширенішими є нечіткі бази знань Мамдані та Сугено.

Базу знань Мамдані складають правила, в яких антецеденти і консеквенти задано нечіткими множинами [4]. Її можна трактувати як розбиття факторного простору на зони з нечіткими межами, в кожній з яких функція відклику приймає нечітке значення. Кількість таких нечітких зон дорівнює кількості правил, яких для адекватного опису складних залежностей зазвичай, потрібно, достатньо багато.

Базу знань Сугено складають правила, антецеденти яких задано нечіткими множинами, а консеквенти — лінійними залежностями між входами та виходом [5]. Базу знань Сугено можна розглядати як розбиття факторного простору на зони з нечіткими межами, в кожній з яких діє свій закон «входи—вихід». Границі зон розмиті, тому в будь-якій точці факторного простору можуть виконуватися декілька законів «входи—вихід», але з різними ступенями. Основний недолік бази знань Сугено полягає в труднощах, з якими зустрічається експерт під час визначення коефіцієнтів в консеквентах правил.

Метою статті є дослідження нечіткої бази знань нового формату, яка поєднає переваги вищеописаних моделей, а саме компактності бази знань Сугено та зручності експертного формування правил бази знань Мамдані. Ідея, що походить з роботи [6], полягає в описі коефіцієнтів в консеквентах правил бази знань Сугено за допомогою нечітких чисел. Іншими словами, кожен коефіцієнт в лінійних законах «входи—вихід» задається нечітким числом, що можна зіставити з нечітким консеквентом в базі знань Мамдані. З іншого боку консеквенти правил в базі знань нового формату можна ототожнити з нечіткими регресійними рівняннями.

### Нечітка база знань з нечіткими регресійними рівняннями

Нечітку базу знань з нечіткими регресійними рівняннями запишемо у вигляді такої сукупності правил:

$$(x_1 = \tilde{a}_{j1} \ominus_j x_2 = \tilde{a}_{j2} \ominus_j \dots \ominus_j x_n = \tilde{a}_{jn}) \Rightarrow y = \tilde{d}_j, j = \overline{1, m}, \quad (1)$$

де  $\tilde{a}_{ji}$  — нечіткий терм, яким оцінено вхідну змінну  $x_i$  в  $j$ -му правилі  $i = \overline{1, n}$ ,  $j = \overline{1, m}$ ;

$m$  — кількість правил;  $\ominus_j$  — логічна операція (ТА чи АБО), яка пов'язує фрагменти антецедента  $j$ -го правила;  $\Rightarrow$  — нечітка імплікація;  $\tilde{d}_j = \tilde{b}_{j0} + \tilde{b}_{j1}x_1 + \tilde{b}_{j2}x_2 + \dots + \tilde{b}_{jn}x_n$  — консеквент  $j$ -го правила, який описано лінійною функцією з нечіткими коефіцієнтами  $\tilde{b}_{j0}, \tilde{b}_{j1}, \dots, \tilde{b}_{jn}$ .

На відміну від бази знань Сугено в (1) коефіцієнти в консеквентах правил задано нечіткими числами.

ми. Тому експерт може описати ці нечіткі коефіцієнти лінгвістичними оцінками «мало впливає», «помірно впливає», «сильно впливає» тощо, які відображають його знання про ступінь впливу відповідної вхідної змінної на вихідну. За такими лінгвістичними оцінками можна визначати ядро нечіткого коефіцієнта в (1). «Розмазаність» нечіткого коефіцієнта залежить від впевненості експерта в достовірності своїх знань, яку можна виразити термами «абсолютно достовірно», «майже достовірно», «більш-менш достовірно» тощо. Чим достовірніше знання, тим концентрованіший нечіткий коефіцієнт.

### Логічне виведення за чітких початкових даних

Логічне виведення по базі знань (1) здійснюватимемо у такий спосіб. Спочатку для вектора  $X^* = (x_1^*, x_2^*, \dots, x_n^*)$  поточних значень вхідних змінних за правилами нечіткої арифметики розрахуємо нечіткі значення консеквентів:

$$\tilde{d}_j = \tilde{b}_{j0} + \tilde{b}_{j1}x_1^* + \tilde{b}_{j2}x_2^* + \dots + \tilde{b}_{jn}x_n^*, \quad j = \overline{1, m}. \quad (2)$$

Це перетворить (1) в базу знань Мамдані, тому подальші кроки здійснимо за алгоритмом Мамдані [1]. Зауважимо, що для кожного вхідного вектора  $X^*$  створюється база знань Мамдані з унікальним набором нечітких консеквентів.

Відповідно до алгоритму Мамдані ступінь виконання антецедента  $j$ -го правила для поточного вхідного вектора  $X^* = (x_1^*, x_2^*, \dots, x_n^*)$  розрахуємо так:

$$\mu_j(X^*) = \mu_{j1}(x_1^*) \chi_j \mu_{j2}(x_2^*) \chi_j \dots \chi_j \mu_{jn}(x_n^*), \quad j = \overline{1, m}. \quad (3)$$

де  $\mu_{ji}(x_i^*)$  — ступінь належності вхідного значення  $x_i^*$  нечіткому терму  $\tilde{a}_{ij}$ ,  $i = \overline{1, n}$ ;  $\chi_j$  позначає  $t$ -норму, якщо  $\Theta_j = TA$  та  $s$ -норму, якщо  $\Theta_j = ABO$ .

В алгоритмі Мамдані зазвичай використовують норми Заде, тобто мінімум як  $t$ -норма та максимум як  $s$ -норма.

Результатом виведення за  $j$ -м правилом бази знань є така нечітка множина:

$$\tilde{d}_j^* = \text{imp}(\tilde{d}_j, \mu_j(X^*)), \quad j = \overline{1, m}, \quad (4)$$

де  $\text{imp}$  позначає імплікацію, яку реалізують операцією мінімуму.

Геометричною інтерпретацією імплікації є «зрізання» графіка функції належності  $\mu_{d_j}(y)$  нечіткого консеквента (2) по рівню  $\mu_j(X^*)$ :

$$\tilde{d}_j^* = \int_{y \in [\underline{y}, \bar{y}]} \min(\mu_j(X^*), \mu_{d_j}(y)) / y,$$

де  $[\underline{y}, \bar{y}]$  — діапазон зміни вихідної змінної  $y$ .

Результат виведення за усіма правилами знаходимо агрегуванням нечітких множин (4):

$$\tilde{y}^* = \text{agg}(\tilde{d}_1^*, \tilde{d}_2^*, \dots, \tilde{d}_m^*), \quad (5)$$

де  $\text{agg}$  — агрегування, яке реалізують операцією максимуму над функціями належності.

Чітке результуюче значення  $y^*$  визначається через дефазифікацію нечіткої множини  $\tilde{y}^*$  за методом центра тяжіння.

### Приклад

Розглянемо залежність ймовірності ( $p_1$ ) правильного введення символу в процесі набору оператором тексту на комп'ютері від двох факторів:  $x_1$  — резерв часу у разі набору одного символу, який відповідає рівню завантаженості оператора, та  $x_2$  — тривалість роботи [7]. Опишемо цю залежність такою базою знань:

якщо  $x_2 = \text{Початок}$ , тоді  $p_1 = \tilde{b}_{10} + \tilde{b}_{11}x_1 + \tilde{b}_{12}x_2$ ;

якщо  $x_2 = \text{Середина}$ , тоді  $p_1 = \tilde{b}_{20} + \tilde{b}_{21}x_1 + \tilde{b}_{22}x_2$ ;

якщо  $x_2 = \text{Кінець}$ , тоді  $p_1 = \tilde{b}_{30} + \tilde{b}_{31}x_1 + \tilde{b}_{32}x_2$ .

Коефіцієнти в консеквентах правил задамо трикутними нечіткими числами з табл. 1. В ній  $a$  позначає ліву межу носія,  $b$  — ядро та  $c$  — праву межу носія нечіткого числа.

Розрахунок нечітких консеквентів  $\tilde{d}_1$ ,  $\tilde{d}_2$  та  $\tilde{d}_3$  для значень факторів  $x_1^* = 1,5$  с та  $x_2^* = 2$  год. зображено на рис. 1. За функцій належності з рис. 2 графік залежність нечіткої ймовірності правильного введення символу від тривалості роботи віддзеркалює етапи припрацювання, нормальної роботи та втоми оператора (рис. 3).

Таблиця 1

Параметри трикутних функцій належності нечітких коефіцієнтів

| Параметр | $\tilde{b}_{11}$ | $\tilde{b}_{12}$ | $\tilde{b}_{10}$ | $\tilde{b}_{21}$ | $\tilde{b}_{22}$ | $\tilde{b}_{20}$ | $\tilde{b}_{31}$ | $\tilde{b}_{32}$ | $\tilde{b}_{30}$ |
|----------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| $a$      | 0                | 0                | 0,92             | 0                | 0                | 0,95             | 0,01             | -0,01            | 0,92             |
| $b$      | 0,015            | 0,005            | 0,94             | 0,01             | 0,0002           | 0,96             | 0,02             | -0,0015          | 0,935            |
| $c$      | 0,02             | 0,01             | 0,96             | 0,02             | 0,0005           | 0,99             | 0,04             | 0                | 0,95             |

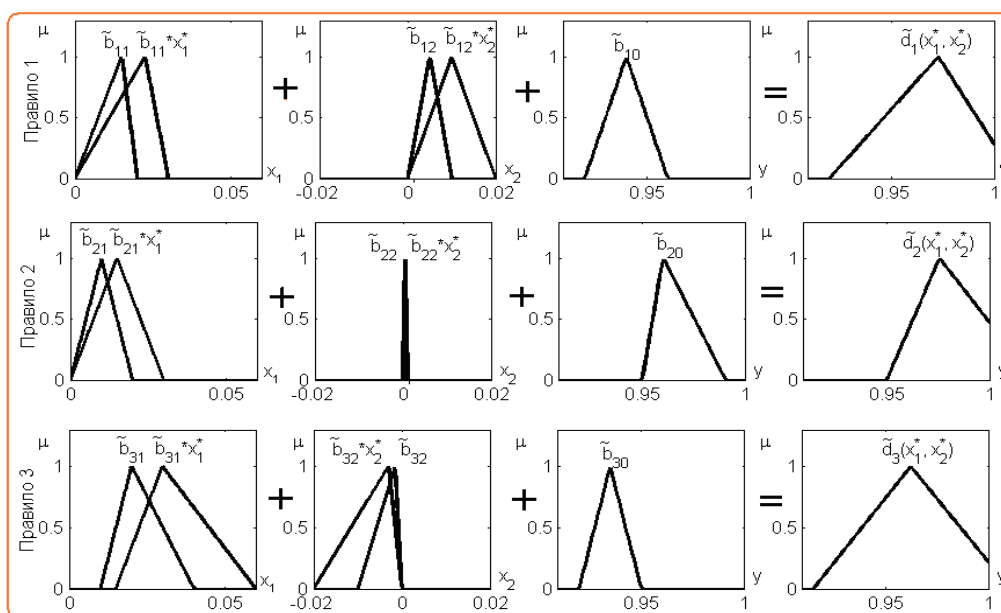


Рис. 1. Розрахунок нечітких консеквентів для  $x_1^* = 1,5$  с та  $x_2^* = 2$  год



Рис. 2. Функції належності термів змінної  $x_2$

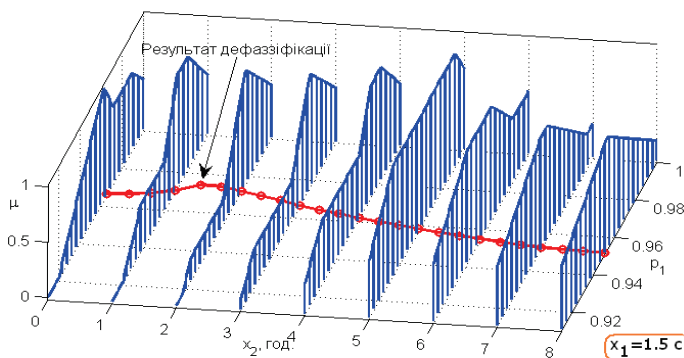


Рис. 3. Залежність нечіткої ймовірності правильного введення символу ( $\tilde{p}_1$ ) від тривалості роботи

**Логічне виведення за нечітких початкових даних**

Позначимо вхідний вектор з нечіткими значеннями вхідних змінних через  $\tilde{X}^* = (\tilde{x}_1^*, \tilde{x}_2^*, \dots, \tilde{x}_n^*)$ . Для цих значень вхідних змінних за правилами нечіткої арифметики розрахуємо нечіткі значення консеквентів  $\tilde{d}_j = \tilde{b}_{j0} + \tilde{b}_{j1}\tilde{x}_1^* + \tilde{b}_{j2}\tilde{x}_2^* + \dots + \tilde{b}_{jn}\tilde{x}_n^*$ ,  $j = \overline{1, m}$ . Подальше виведення здійснимо за формулами (3)—(4), тільки замість  $\mu_{ji}(x_i^*)$  використовуватимемо  $\mu_{ji}(\tilde{x}_i^*)$  — ступінь належності нечіткого вхідного значення  $\tilde{x}_i^*$  нечіткому терму  $\tilde{a}_{ij}$ ,  $j = \overline{1, m}$ ,  $i = \overline{1, n}$ . Ці значення розрахуємо так [1, 8]:

$$\mu_j(\tilde{x}_i^*) = \text{height}(\tilde{x}_i^* \cap \tilde{a}_{ij}), \tag{6}$$

де *height* — висота нечіткої множини.

Якщо функції належності задані в параметричній формі, тоді можна отримати швидкі правила обчислення (6). Як приклад, в табл. 2 пропонуються такі правила для трьох типів функцій належності.

Таблиця 2

**Реалізація формули (6) для параметричних функцій належності**

| Функції належності нечітких множин $\tilde{A}$ та $\tilde{B}$   | Правила швидкого розрахунку <i>height</i> ( $\tilde{A} \cap \tilde{B}$ ) на мові MATLAB  |
|---|--|
| <p>Гаусові:</p> $\mu_{\tilde{A}}(x) = \exp\left(-\frac{(x - b_1)^2}{2c_1^2}\right);$ $\mu_{\tilde{B}}(x) = \exp\left(-\frac{(x - b_2)^2}{2c_2^2}\right)$                | <pre> if b1==b2 height=1; else x1=(b1*c2+b2*c1)/(c2+c1); y1=gaussmf(x1, [b1 c1]); if c1==c2 y2=y1; else x2=(b1*c2-b2*c1)/(c2-c1); y2=gaussmf(x2, [b1 c1]); end height=max(y1, y2); end                     </pre>  |
| <p>Сигмоїдні:</p> $\mu_{\tilde{A}}(x) = \frac{1}{1 + \exp(-a_1(x - c_1))};$ $\mu_{\tilde{B}}(x) = \frac{1}{1 + \exp(-a_2(x - c_2))}$                                    | <pre> if ((a1==a2) &amp; (c1==c2)) height=1; elseif a1==a2 height=0; else x=(a2*c2-a1*c1)/(a2-a1); height=sigmf(x, [a1 c1]); end                     </pre>  |
| <p>Дзвонові:</p> $\mu_{\tilde{A}}(x) = \frac{1}{1 + \left \frac{x - c_1}{a_1}\right ^{2b}};$ $\mu_{\tilde{B}}(x) = \frac{1}{1 + \left \frac{x - c_2}{a_2}\right ^{2b}}$ | <pre> if c1==c2 height =1; else x=(a2*c1+a1*c2)/(a1+a2); y1=gbellmf(x, [a1 b c1]); if a1==a2 y2=0; else x=(a2*c1-a1*c2)/(-a1+a2); y2=gbellmf(x, [a1 b c1]); end height=max(y1, y2); end                     </pre> |

**Навчання нечіткої бази знань**

Навчальну вибірку з  $M$  пар «входи—вихід» запишемо таким чином:

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

де  $X_r$  — вхідний вектор в  $r$ -му рядку вибірки та  $y_r$  — відповідний вихід.

Позначимо через  $y = F(P, X)$  модель на основі нечіткої бази знань (1) з параметрами  $P$ . Грунтуючись на ідеях нечіткої ідентифікації [1, 5, 8], навчання бази знань (1) зведемо до пошуку вектора  $P$ , що забезпечує:

$$RMSE = \sqrt{\frac{1}{M} \sum_{r=1, M} (y_r - F(P, X_r))^2} \rightarrow \min. \quad (7)$$

В (7) керовані змінні  $P$  задають параметри функцій належності нечітких множин з антецедентів та консеквентів правил бази знань (1). Для збереження прозорості на параметри функцій нечітких множин  $\tilde{a}_{ij}$ ,  $i = \overline{1, n}$ ,  $j = \overline{1, m}$  накладемо обмеження згідно з [9]. Задачу (6) можна розв'язати типовими методами математичного програмування, як це відбувається під час ідентифікації залежностей нечіткими базами знань Мамдані та Сугено.

### Висновки

Запропоновано нову структуру нечіткої бази знань, в якій антецеденти задаються нечіткими термами, а консеквенти—лінійними залежностями «входи—вихід» з нечіткими коефіцієнтами. Її перевагою над базою знань Мамдані є компактність, а над базою знань Сугено — можливість задання консеквентів лінгвістичними оцінками, що для експерта зручніше. Запропоновано алгоритм логічного виведення, за яким спочатку для кожного вхідного вектора синтезується унікальна база знань з нечіткими консеквентами, а подальші розрахунки здійснюються за методом Мамдані. Показано, як здійснити логічне виведення за чітких та нечітких початкових даних. На основі теорії нечіткої ідентифікації сформульовано математичні постановки задачі навчання запропонованої бази знань за вибірками експериментальних даних.

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Штовба С. Д. Проектирование нечетких систем средствами MATLAB / С. Д. Штовба. — М. : Горячая линия — Телеком, 2007. — 288 с.
2. Precup R. A Survey on Industrial Applications of Fuzzy Control / R. Precup, H. Hellendoorn // Computers in Industry. — 2011. — Vol. 62, № 3. — P. 213—226.
3. Wong B. A Survey of the Application of Fuzzy Set Theory in Production and Operations Management: 1998—2009 / B. Wong, V. Lai // International Journal of Production Economics. — 2011. — Vol. 129, № 1. — P. 157—168.
4. Заде Л. Основы нового подхода к анализу сложных систем и процессов принятия решений / Л. Заде // Математика сегодня. — М. : Знание, 1974. — С. 5—49.
5. Takagi T. Fuzzy Identification of Systems and Its Applications to Modeling and Control / T. Takagi, M. Sugeno // IEEE Trans. on Systems, Man, and Cybernetics. — 1985. Vol. 15, № 1. — P. 116—132.
6. Штовба С. Д. Ідентифікація залежностей нечіткою базою знань з нечіткими регресійними рівняннями / С. Д. Штовба // Системний аналіз та інформаційні технології : Х міжнар. НТК., 20—24 травня 2008 р. : матер. конф. — К., 2008. — 273 с.
7. Ротштейн А. П. Моделирование надежности человека-оператора с помощью нечеткой базы знаний Сугено / А. П. Ротштейн, С. Д. Штовба // Автоматика и телемеханика. — 2009. — № 1. — С. 180—187.
8. Ротштейн А. П. Идентификация нелинейной зависимости нечеткой базой знаний с нечеткой обучающей выборкой / А. П. Ротштейн, С. Д. Штовба // Кибернетика и системный анализ. — 2006. — № 2. — С. 17—24.
9. Штовба С. Д. Обеспечение точности и прозрачности нечеткой модели Мамдани при обучении по экспериментальным данным / С. Д. Штовба // Проблемы управления и информатики. — 2007. — № 4. — С. 102—114.

Рекомендована кафедрою комп'ютерних систем управління

Стаття надійшла до редакції 26.05.11  
Рекомендована до друку 26.05.11

**Штовба Сергій Дмитрович** — професор кафедри комп'ютерних систем управління.

Вінницький національний технічний університет, Вінниця