

## ДОСЛІДЖЕННЯ БАЗ ЗНАНЬ МАМДАНІ З РІЗНОЮ КІЛЬКІСТЮ НЕЧІТКИХ ПРАВИЛ

Вінницький національний технічний університет

### Анотація

Розроблено систему нечіткого логічного виведення за Мамдані. Здійснено навчання нечіткої бази знань за допомогою середовища програмування MatLab з використанням бібліотеки Fuzzy Logic ToolBox. На основі експериментів знайдено функціональну залежність нев'язки від кількості нечітких правил у базі знань. Визначено оптимальну кількість правил у базі знань.

**Ключові слова:** нечітка логіка, база знань, правила, ідентифікація, алгоритм Мамдані, нев'язка, експерименти, складність.

### Abstract

A system with Mamdani fuzzy logic output has been developed. The fuzzy knowledge base has been trained with the usage MatLab programming environment and Fuzzy Logic ToolBox library. On the basis of experiments, the functional dependence for RMSE from the number of fuzzy rules in the knowledge base has been found. The optimal number of rules in the knowledge base has been determined.

**Keywords:** fuzzy logic, knowledge base, rules, identification, Mamdani algorithm, RMSE, experiments, complexity.

### Вступ

Прийняття правильних рішень в умовах неповної і нечіткої інформації – це найбільш вражаюча властивість людського інтелекту. Побудова моделей наближених до роздумів людини і використання їх у комп'ютерних системах представляє сьогодні одну з найважливіших проблем науки. Оскільки в задачах, що вирішуються інтелектуальними системами, часто доводиться користуватися неточними знаннями, які не можуть бути інтерпретовані як повністю істинні або помилкові (логічні true/false або 1/0), то на практиці широко застосовується використання властивостей нечітких множин [1-3]. На початку 70-х років американський математик Л. Заде запропонував формальний апарат *нечіткої (fuzzy) алгебри і нечіткої логіки*, який дозволяє зберегти властивості розмитості і неточності знань [4].

Популярність теорії нечітких множин в проектуванні пояснюється тим, що нечіткі системи швидше розробляються, вони реалізуються простіше та дешевше, ніж їх чіткі аналоги. Експертні знання легко ввести в нечіткі системи, що дозволяє швидко створювати прототипи зі зрозумілими для людини алгоритмами функціонування. Знання в подібних системах зберігають у особливого роду базах даних (*базах знань*), розроблених для управління знаннями, тобто збиранням, зберіганням, пошуком і видачею знань. Цікаво, що в базах даних немає залежностей, а в базах знань є залежності між даними виходу по вхідним даним. Нечітка база знань може трактуватися як деяке розбиття простору чинників впливу на області з розмитими межами, в кожній з яких функція відклику приймає значення, яке задається відповідною нечіткою множиною. Кількість таких нечітких зон дорівнює кількості *правил*, яких для адекватного опису складних залежностей потрібно досить багато. Правила в базі знань представляють собою «інформаційні згустки» або «гранули знань», тобто вони розглядаються як аналог вербального кодування, яке, як встановили психологи, відбувається в людському мозку під час навчання. Саме тому формування нечіткої бази знань в конкретній предметній області, як правило, не складає труднощів для експерта. Фактично нечіткою базою знань називається сукупність нечітких правил «якщо-то», яка задає взаємозв'язок між входами та виходами досліджуваного об'єкта, описуючи певну предметну область. Перша частина правила («якщо») називається *антецедентом* або *посилкою*, а друга частина («то») – *консеквентом* або *висновком*. Щоб зберегти в нечіткій базі певний ряд правил необхідно ідентифікувати залежність, якою ці знання задаються. *Ідентифікація залежностей*, тобто побудова їх моделей за результатами

спостережень, є важливим завданням в управлінні технологічними процесами, технічній та медичній діагностиці, прогнозуванні часових рядів, розпізнаванні образів тощо [5-7].

Сьогодні найпоширенішими є нечіткі бази знань Мамдані, Сугено, Ларсена, Цукамото, сингтонна. Для них розроблені методи навчання для успішної ідентифікації моделей. У даній роботі досліджується база знань *I. Мамдані*. Базу знань Мамдані складають правила, в яких антецеденти і консеквенти задано нечіткими множинами [8].

Слід зазначити, що проектувальники нечітких систем прагнуть створити адекватну нечітку базу знань найменшої складності – з малою кількістю правил. Таку компактну модель легше верифікувати; вона забезпечує найбільшу швидкість логічного виведення та потребує найменше ресурсів для апаратної реалізації [9].

Саме тому *метою* роботи є зменшення складності бази знань Мамдані шляхом експериментального визначення оптимальної кількості нечітких правил у ній.

### Результати дослідження

*Нечітка модель Мамдані* [8, 9]. У нечіткій базі знань Мамдані антецеденти і консеквенти задано нечіткими множинами. В алгоритмі Мамдані [8] результатом виведення за  $j$ -им правилом бази знань є така нечітка множина

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

де *imp* позначає імплікацію, яку реалізує операція мінімуму;  $\tilde{d}_j$  – нечітке значення, яке вибирається з терм-множини;  $X^*$  – поточний вхідний вектор;  $\mu_j(X^*)$  – ступінь виконання антецедента  $j$ -ого правила.

Результат виведення за усіма правилами знаходять агрегуванням нечітких множин

$$\tilde{y}^* = \text{agg}(\tilde{d}_1^*, \tilde{d}_2^*, \dots, \tilde{d}_m^*), \quad (2)$$

де *agg* – агрегування нечітких множин, яку реалізують операцією максимуму.

Чітке значення виходу  $y^*$ , яке відповідає вхідному вектору  $X^*$ , визначається через *дефаззіфікацію* нечіткої множини  $\tilde{y}^*$ . Найбільш поширені методи дефаззіфікації – це центр тяжіння, медіана, центр максимумів.

Проведемо *навчання нечіткої бази знань Мамдані* [9-12]. Навчанням називається процес підбору параметрів функцій приналежностей, кількості нечітких правил та їх ваг, які забезпечують мінімальне відхилення між експериментальними даними та результатами нечіткого логічного висновку [11].

*Критерієм якості навчання* може виступати *нев'язка RMSE* між модельними та експериментальними даними. Навчальну вибірку з  $M$  пар експериментальних даних  $X = (x_1, x_2, \dots, x_n)$  запишемо таким чином  $(X_r, y_r)$ ,  $r = \overline{1, M}$ , де  $X_r$  – вхідний вектор в  $r$ -ому рядку вибірки, а  $y_r$  – відповідний вихід. Позначимо через  $y = F(H, X)$  модель на основі нечіткої бази знань з параметрами  $H$ , яка пов'язує фактори впливу  $X$  з характеристикою  $y$ . Відповідно до методу найменших квадратів задача ідентифікації зводиться до пошуку вектору керованих змінних  $H$  (задають параметри функцій приналежності), що забезпечує невід'язку

$$RMSE = \sqrt{\frac{1}{M} \sum_{r=1, M} (y_r - F(H, X_r))^2} \rightarrow \min. \quad (3)$$

На *першому етапі* навчання бази знань виконується *структурна ідентифікація*. Вона є формуванням нечіткої бази знань, яка грубо відображає нелінійний взаємозв'язок «входи - вихід» за допомогою *лінгвістичних правил*. Ці правила генеруються експертом або отримуються в результаті екстракції нечітких знань з експериментальних даних [9, 11].

Фактично структурна ідентифікація моделей полягає в знаходженні такої бази знань, яка має стати базу лінгвістичних правил, що має найменшу помилку ідентифікації. Для процесів коригування раніше накопичених знань найбільш інформативними кривими навчання є залежність нев'язки від часу навчання, а для процесів набуття нових знань – залежність нев'язки від кількості правил в базі знань. Максимальна кількість правил в нечіткій базі знань дорівнює  $N_{\max} = l_1 \cdot \dots \cdot l_n$ , де  $l_i$  – кількість термів для оцінки  $i$ -ої вхідної змінної ( $i = 1, \dots, n$ ), а  $n$  – кількість вхідних змінних.

Однією з переваг систем нечіткого логічного висновку є їх задовільна робота при кількості правил в базі знань меншій, ніж  $N_{\max}$ . Звичайно, додавання кожного нового правила покращує якість роботи системи, але в реальних умовах кількість достовірних знань завжди обмежена. Крім того, для нечітких баз знань існує так зване «явище насиченості», суть якого полягає в тому, що, починаючи з певного обсягу бази знань, додавання нового правила практично не покращує характеристики системи. Цьому явищу на графіку залежності нев'язки від обсягу бази знань відповідає «плато насичення» – майже горизонтальний відрізок прямої.

На **другому етапі** відбувається *параметрична ідентифікація* досліджуваної залежності шляхом знаходження таких параметрів нечіткої бази знань, які мінімізують відхилення результатів нечіткого моделювання від експериментальних даних. Параметрами, які налаштовуються, є ваги правил і параметри функцій приналежності нечітких термів [9, 11].

Вибір правил нечіткої бази знань можна звести до оптимізаційної задачі про рюкзак, яка є *NP*-повною [3]. Правилу бази знань відповідає предмет, який може потрапити до рюкзака, точності бази знань – корисність рюкзака, а кількості правил – сумарний обсяг вибраних предметів. Відмінність між задачами полягає в різних типах функції корисності, яка є лінійною в задачі про рюкзак та нелінійною в задачі вибору правил бази знань. Відповідно алгоритм точного розв'язання цієї задачі матиме експоненціальну обчислювальну складність, і тому буде прийнятним лише за невеликої кількості правил-кандидатів.

Розглянемо *приклад експериментального визначення* [11, 12] оптимальної кількості нечітких правил у базі знань Мамдані. Для кожного нечіткого розбиття експерименти проводилися за певною схемою. Робота виконувалася шляхом написання декількох *M*-файлів в середовищі *Matrix Laboratory*. Використовувалась наступна багатопараметрична функціональна залежність

$$z(x, y) = \frac{x^3 \cdot \sin(0,2 \cdot y)}{x + y^2}, \quad \text{де} \quad \begin{matrix} x \in [100, 150], \\ y \in [50, 100], \end{matrix} \quad \text{а кількість термів } N_{\{x, y, z\}} = \{4, 5, 6\}. \quad (4)$$

Виконано перебір баз знань за нечіткими розбиттями  $\{2*2, 2*3, 2*4, 2*5, 3*2, 3*3, 3*4, 3*5, 4*2, 4*3, 4*4, 4*5\}$  для кількості правил  $\{4, 6, 8, 9, 10, 12, 15, 16, 20\}$ . Для кожної множини нечітких баз знань одного розміру знайдено мінімальну, максимальну та середню нев'язки. Експериментальні дані  $RMSE_{\min}$  від кількості правил  $N$  апроксимовано квадратичною функціональною залежністю, яка описує процес навчання нечіткої бази знань Мамдані у формі

$$f(N) = 0,0032111 \cdot N^2 - 0,094706 \cdot N + 1,1021. \quad (5)$$

Експерименти показали, що після навчання нев'язка  $RMSE$  спадає зі збільшенням кількості нечітких правил, але її мінімум в багатьох випадках досягається за неповної бази знань. Бази знань, які містять біля 75-80% від максимальної кількості правил, мають найвищу точність, тобто мають найменшу похибку. У даному випадку (формула 5) мінімум досягається при 15 правилах (розбиття  $3*5$ , максимальна кількість комбінацій правил  $2^{15} = 32768$ ). Для зниження обчислювальної складності повну нечітку базу знань Мамдані можна скоротити в 2 рази до 10 правил (розбиття  $2*5$ , максимальна кількість комбінацій правил  $2^{10} = 1024$ ) без великих втрат точності. Експериментально встановлено, що базу знань Мамдані з достатньо малою похибкою ідентифікації можна отримати при наповненості бази знань на 30-50% (6-10 правил), після чого збільшення кількості правил істотно не зменшує нев'язку. Такі компактні бази знань є прозорішими та легше навчаються через меншу складність відповідної задачі оптимізації.

## Висновки

У даній роботі розроблено систему нечіткого логічного виведення за Мамдані та здійснено навчання бази знань з використанням програмного забезпечення. Подібна система може бути використана у зв'язці з нейронними мережами, як інструмент для аналізу даних фахівцями у різних предметних областях. Виконані комп'ютерні експерименти показують, що повна нечітка база знань Мамдані не дає найменшу похибку ідентифікації. Це обумовлено ускладненням задачі оптимізації через збільшення кількості керованих параметрів. Експериментально визначено оптимальну кількість правил у нечіткій базі знань, що дозволяє зменшити її складність та підвищити інтерпретованість.

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Штовба С.Д. Проектирование нечетких систем средствами MATLAB / С.Д. Штовба. – М.: Горячая линия-Телеком, 2007. – 288 с.
2. Ротштейн А.П. Интеллектуальные технологии идентификации: нечеткая логика, генетические алгоритмы, нейронные сети / А.П. Ротштейн. – Винница: УНІВЕРСУМ, 1999. – 320 с.
3. Иванов Ю.Ю. Вступ до Computer Science. Дискретна математика: цікава та не дуже: лекції, алгоритми та задачі / Ю.Ю. Иванов. – 2018. – 89 с. – Режим доступу: [https://iq.vntu.edu.ua/method/read\\_url.php?tbl\\_num=2&url=/fdb/1166/Discrete\\_Math\\_by\\_IVANOV.djvu](https://iq.vntu.edu.ua/method/read_url.php?tbl_num=2&url=/fdb/1166/Discrete_Math_by_IVANOV.djvu).
4. Заде Л.А. Понятие лингвистической переменной и его применение к принятию приближенных решений / Л.А. Заде. – М.: Мир, 1976. – 165 с.
5. Штовба С.Д. Проектирование систем управления в Fuzzy Logic Toolbox [Электронный ресурс] / С.Д. Штовба. – Режим доступа: <http://matlab.exponenta.ru/fuzzylogic/book2/index.php>.
6. Штовба С.Д. Идентификация нелинейных зависимостей с помощью нечеткого логического вывода в системе MATLAB / С.Д. Штовба // Математика в приложениях. – 2003. – №2 (2). – С. 9-15.
7. 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. – V. 15. – № 1. – P. 116-132.
8. Mamdani E.H. An Experiment in Linguistic Synthesis with Fuzzy Logic Controller / E.H. Mamdani, S. Assilian // Int. J. Man-Machine Studies. – 1975. – V. 7. – № 1. – P. 1-13.
9. Штовба С.Д. Интеллектуальні технології ідентифікації залежностей. Лабораторний практикум: електронний навчальний посібник / С.Д. Штовба, В.В. Мазуренко. – Вінниця: ВНТУ, 2014. – 113 с.
10. Штовба С.Д. Дослідження навчання компактних нечітких сингтонних баз знань / С.Д. Штовба, В.В. Мазуренко // Вимірювальна та обчислювальна техніка в технологічних процесах. – 2011. – № 1. – С. 133-139.
11. Штовба С.Д. Дослідження навчання компактних нечітких баз знань типу Мамдані / С.Д. Штовба, В.В. Мазуренко // Штучний інтелект. – 2011. – № 4. – С. 521-529.
12. Штовба С.Д. Вплив кількості нечітких правил на точність бази знань Мамдані / С.Д. Штовба, В.В. Мазуренко, О.Д. Панкевич // Вісник Хмельницького національного університету. Технічні науки. – 2011. – № 2. – С. 182-185.

**Олійник Артем Костянтинович** — магістрант групи ІАКІТ-17м, факультет комп'ютерних систем і автоматики, Вінницький національний технічний університет, м. Вінниця.

**Іванов Юрій Юрійович** — канд. техн. наук, асистент кафедри автоматики та інформаційно-вимірювальної техніки, Вінницький національний технічний університет, м. Вінниця, e-mail: Yura881990@i.ua.

Науковий керівник: **Іванов Юрій Юрійович** — канд. техн. наук, асистент кафедри автоматики та інформаційно-вимірювальної техніки, Вінницький національний технічний університет, м. Вінниця.

**Oliynik Artem K.** — graduate student, Faculty of Computer Systems and Automation, Vinnytsia National Technical University, Vinnytsia.

**Ivanov Yuriy Yu.** — Cand. Sc. (Eng), Assistant Professor, Faculty of Computer Systems and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: Yura881990@i.ua.

Supervisor: **Ivanov Yuriy Yu.** — Cand. Sc. (Eng), Assistant Professor, Faculty of Computer Systems and Automation, Vinnytsia National Technical University, Vinnytsia.