

ВИКОРИСТАННЯ МЕТОДІВ НЕЧІТКОГО ВИВЕДЕННЯ В СИСТЕМАХ ПРОГНОЗУВАННЯ СТАНІВ ПРИРОДНИЧИХ ПРОЦЕСІВ

© Ковальчук А., Суха О., Фабрі Л., 2009

Запропоновано методику використання можливостей нечіткого виведення в системах прогнозування природничих і соціальних процесів.

It is offered the theory of operation of capabilities of an indistinct conclusion in systems of the forecast of condition of natural and social processes.

Вступ

На сучасному етапі розвитку людства вже практично неможливо уявити собі жодну галузь побуту та промисловості, де б не використовувались комп'ютери та різноманітне програмне забезпечення. Цінність оперативної, достовірної і максимально повної інформації неминуче зростає. Із розвитком комп'ютерної техніки та програмних засобів стало можливо легко та зручно прогнозувати дані за допомогою різних алгоритмів [1].

Проблема прогнозування стану будь-якого процесу є складною та багатогранною, як і складність вироблення єдиної універсальної технології оцінки та контролю прогнозування, оскільки, як відомо, точність оцінок є обернено пропорційною складності об'єктів, характеристики яких оцінюються [1]. Отже, ця проблема вимагає вирішення.

Постановка задачі

Метою є розроблення та впровадження різноманітних методів прогнозування, зокрема можливість їх порівняння та вибору найкращого. Сьогодні існує багато різних технологій та мов програмування для розроблення програмного продукту з метою розв'язання поставленої задачі. Також є дуже багато операційних систем, але програмний продукт має працювати з найпоширенішою з них, а саме – Windows.

На вибір методу оцінювання варіантів впливає низка чинників. Оскільки на процес оцінювання та контролю діяльності значний вплив, окрім об'єктивних факторів, мають ще й суб'єктивні та зважаючи на складність поставленої задачі, часто пропонуються методики оцінювання на основі нечітких множин [1, 2]. Такий підхід, звісно, є доволі ефективним внаслідок простоти обчислень та можливості урахування якісних показників оцінювання, проте має і деякі обмеження, пов'язані зі значним підпорядкуванням суб'єктивному впливу, що за певних умов може призвести до зменшення ролі об'єктивного фактора.

Методи нечіткого виведення

У нечітких продукційних моделях використовуються переважно два способи нечіткого виведення : прямий та обернений.

Прямий спосіб нечіткого виведення, або прямий нечіткий ланцюг міркувань (fuzzy forward-chaining reasoning) ґрунтується на правилі виведення “нечіткий модус поненс” (fuzzy modus ponens), узагальнена схема якого містить такі етапи.

Етап 1. Задання нечіткої імплікації $R: A \dot{\Rightarrow} B$, що визначає нечітке причинно-наслідкове відношення між передпосиланням (антецедентом) A та висновком (консеквентом) B , яке представляється у вигляді нечіткої імплікації:

$$\text{якщо } x \in A, \text{ то } y \in B, \quad (1)$$

де x – вхідна змінна, $x \in X$, де X – область визначення посилення нечіткого продукційного правила; A – нечітка множина, що визначена на X з функцією належності $\mu_A(x) \in [0,1]$ y – вихідна змінна, $y \in Y$, де Y – область визначення виведення; B – нечітка множина, визначена на Y , з функцією належності $\mu_B(y) \in [0,1]$.

Операція нечіткої імплікації займає центральне місце в нечітких продукційних моделях, визначаючи причинно-наслідкове відношення між посиленнями та виведеннями правил. Сьогодні існує декілька десятків різноманітних варіантів цієї операції. Всі вони поділяються на три основні класи:

- S -імплікації, що визначаються як

$$\mu_R(x, y) = S(\text{not}(\mu_A(x), \mu_B(y))), \quad (2)$$

де S – оператор S -норми; **not** – оператор нечіткого доповнення на інтервалі $[0,1]$. Основною операцією імплікації цього класу є булева імплікація. Типовими прикладами S -імплікації можуть бути імплікації

Кліні–Денса і Лукашевича.

- R -імплікації, створені на основі порівняння з T -нормою:

$$\mu_R(x, y) = \sup\{z \in [0,1] \mid T(\mu_A(x), z) \leq \mu_B(y)\}. \quad (3)$$

- T -імплікації, визначені як

$$\mu_R(x, y) = T(\mu_A(x), \mu_B(y)), \quad (4)$$

де T – оператор T -норми.

У роботі запропоновано оцінювати операції нечіткої імплікації за двома критеріями:

- максимальна помилка, обумовлена операцією імплікації, не повинна перевищувати максимальної помилки лінгвістичного представлення даних;
- нечітка модель, побудована з використанням вибраної операції нечіткої імплікації, має відповідати виду функції належності.

У результаті аналізу за вищевказаними критеріями у роботах [2, 3] зроблено висновок про найефективніше використання таких операцій нечіткої імплікації:

- класичної нечіткої імплікації, або імплікації Kleene-Dienes

$$\mu_R(x, y) = \max\{1 - \mu_A(x), \mu_B(y)\}, \text{ при } \mu_A(x) \geq \mu_B(y); \quad (5)$$

- нечіткої імплікації Zadeh

$$\mu_R(x, y) = \max\{\min\{\mu_A(x), \mu_B(y)\}, 1 - \mu_A(x)\}; \quad (6)$$

- нечіткої імплікації Mamdani

$$\mu_R(x, y) = \min\{\mu_A(x), \mu_B(y)\}; \quad (7)$$

- нечіткої імплікації Larsen

$$\mu_R(x, y) = \mu_A(x) \cdot \mu_B(y); \quad (8)$$

- нечіткої імплікації Lukasiewicz

$$\mu_R(x, y) = \min\{1, 1 - \mu_A(x) + \mu_B(y)\},$$

або

$$\mu_R(x, y) = \max\{0, \mu_A(x) + \mu_B(y) - 1\}; \quad (9)$$

- стандартної чіткої (standard strict) імплікації

$$\mu_R(x, y) = \begin{cases} 1, & \text{якщо } \mu_A(x) \leq \mu_B(y), \\ 0, & \text{якщо } \mu_A(x) > \mu_B(y); \end{cases} \quad (10)$$

- нечіткої імплікації Геделя (Gödel)

$$\mu_R(x, y) = \begin{cases} 1, & \text{якщо } \mu_A(x) \leq \mu_B(y), \\ \mu_B(y), & \text{якщо } \mu_A(x) > \mu_B(y); \end{cases} \quad (11)$$

- нечіткої імплікації Гейнса (Gaines)

$$\mu_R(x, y) = \begin{cases} 1, & \text{якщо } \mu_A(x) \leq \mu_B(y), \\ \mu_B(y) / \mu_A(x), & \text{якщо } \mu_A(x) > \mu_B(y); \end{cases} \quad (12)$$

- нечіткої імплікації Гогена (Goguen)

$$\mu_R(x, y) = \min\{1, \mu_B(y) / \mu_A(x)\}, \text{ при } \mu_A(x) > 0 \quad (13)$$

- нечіткої імплікації Кліні–Даенса–Лукашевича (Kline–Dienes–Lukasiewicz)

$$\mu_R(x, y) = 1 - \mu_A(x) + \mu_A(x) \mu_B(y); \quad (14)$$

- нечіткої ймовірнісної імплікації

$$\mu_R(x, y) = \min\{1 - \mu_B(y) + \mu_A(x) \mu_B(y)\}; \quad (15)$$

- нечіткої імплікації з обмеженою сумою

$$\mu_R(x, y) = \min\{1, \mu_A(x) + \mu_B(y)\}; \quad (16)$$

- нечіткої імплікації Н. Ваді

$$\mu_R(x, y) = \max\{\mu_A(x) \mu_B(y), 1 - \mu_B(y)\}; \quad (17)$$

- нечіткої імплікації Ягера (Yager)

$$\mu_R(x, y) = \mu_A(x)^{\mu_B(y)}; \quad (18)$$

Окрім вказаних вище критеріїв, ту чи іншу нечітку імплікацію вибирають залежно від базису нечітких операцій, які використовуються, та її найефективнішої обчислювальної реалізації [4, 5].

Етап 2. Задання нечіткої умови (факту):

$$\langle\langle x' \in A' \rangle\rangle, \quad (19)$$

де x' – фактичне значення змінної x , A' – нечітка множина, що відображає значення x' , що визначене на X , з функцією приналежності $\mu_{A'}(x) \in [0, 1]$.

Етап 3. Формування виведення:

$$\langle\langle y' \in B' \rangle\rangle, \quad (20)$$

де y' – отримане значення змінної y ; B' – нечітка множина, що відображає значення y' , яке визначене на Y , з функцією приналежності $\mu_{B'}(y) \in [0, 1]$.

Оскільки нечітка умова (19) містить нечітку множину A' , що є «близькою» в певному значенні до множини A , то в результаті формується виведення у вигляді функції приналежності нечіткої множини B' , що характеризує її «близькість» до нечіткої множини B .

Отже, процес отримання результату прямого нечіткого виведення B' з використанням нечіткої імплікації $A \dot{\rightarrow} B$ та нечіткої умови $\langle\langle x' \in A' \rangle\rangle$ можна подати у вигляді:

$$B' = A' \bullet R = A' \bullet (A \dot{\rightarrow} B), \quad (21)$$

де « \bullet » – операція згортки (композиційне правило нечіткого логічного виведення).

Нечітка імплікація $A \dot{\rightarrow} B$ відповідає нечіткому відношенню R , яке можна розглядати як нечітку підмножину декартового (прямого) добутку $X \cdot Y$ повної множини посилок X та виведень Y з функцією приналежності $\mu_R(x, y)$. Тому функцію належності нечіткої множини B' можна подати так:

$$\mu_{B'}(y) = \sup_{x \in X} \{ \mu_{A'}(x) T \mu_R(x, y) \}, \quad (22)$$

де $\sup\{\dots\}$ означає верхню границю множини елементів $\{\dots\}$, де x «пробігає» усі значення із X ; T – операція T -норми.

Якщо X являє собою множину зі скінченним числом елементів, то формула (22) набуває вигляду:

$$\mu_{B'}(y) = \max_{x \in X} \{ \mu_{A'}(x) T \mu_R(x, y) \}, \quad (23)$$

Те саме композиційне правило (23) називається (T -**max**) композицією, або макстриангулярною композицією.

Серед правил композиції найпоширенішими є такі операції:

•(**sup-min**) – композиція

$$\mu_{B'}(y) = \sup_{x \in X} \{ \min[\mu_{A'}(x), \mu_R(x, y)] \}; \quad (24)$$

•(**sup-prod**) – композиція

$$\mu_{B'}(y) = \sup_{x \in X} \{ \mu_{A'}(x) \cdot \mu_R(x, y) \}, \quad (25)$$

а також їх різновиди, правила макстриангулярної композиції, для випадків, якщо X є множиною з кінцевим числом елементів:

•(**max-min**) – композиція

$$\mu_{B'}(y) = \max_{x \in X} \{ \min[\mu_{A'}(x), \mu_R(x, y)] \}; \quad (26)$$

•(**max-prod**) – композиція

$$\mu_{B'}(y) = \max_{x \in X} \{ \mu_{A'}(x) \cdot \mu_R(x, y) \}, \quad (27)$$

Широке вживання правила (**max-min**) – композиції обумовлено добрими властивостями алгебри цієї операції (асоціативністю і дистрибутивністю відносно **max**), необхідними в заданнях нечіткого моделювання, а правила (**max-prod**) – композиції – простотою реалізації і більшою чутливістю до змін вхідних змінних в передумовах нечітких продукційних правил [3].

Обернений спосіб нечіткого виводу, або зворотний нечіткий ланцюжок міркувань (fuzzy backward-chaining reasoning) ґрунтується на використанні правила виведення нечіткий модус толленс (fuzzy modus tollens). Його схема містить такі етапи [4].

Етап 1. Задання нечіткої імплікації $R: A \dot{\Rightarrow} B$, що визначає, як і в разі прямого способу виведення, нечітке причинно-наслідкове відношення між передумовою і висновком, який представляється у вигляді нечіткої продукції (1).

Етап 2. Задання нечіткої умови « $y' \in B'$ ».

Етап 3. Формування виведення « $x' \in A'$ ». Метою способу зворотного нечіткого виведення є встановлення істинності передумови нечіткої продукції (1). А сам процес здобуття результату зворотного нечіткого виводу A' з використанням нечіткої імплікації $A \dot{\Rightarrow} B$ і нечіткої умови « $y' \in B'$ » можна представити у вигляді:

$$A' = R \bullet B' = (A \dot{\Rightarrow} B) \bullet B'. \quad (37)$$

Причому виведення за цим способом виходить у вигляді функції приналежності нечіткої множини передумови A' на основі функції приналежності нечіткої імплікації $A \dot{\Rightarrow} B$ і функції приналежності нечіткої множині виведення B' :

$$\mu_{A'}(x) = \sup_{x \in X} \{ \mu_R(x, y) T \mu_{B'}(y) \}, \quad (38)$$


Надалі якщо не вказане інше, розглядатимемо моделі, що реалізують прямий спосіб нечіткого виведення.

Результати

Програмний продукт призначений для прогнозування даних на майбутній рік, на основі бази даних за попередні роки. Початковою сторінкою (рис.1) є сторінка введення даних про систему користувача, а саме: початковий рік, назва прогнозованого параметра, розмірність прогнозованих даних та назва системи.

The screenshot shows a window titled "Система прогнозування" with a menu bar containing "Файл", "Мова", and "Про програму". The main area is divided into two sections. The top section, "Введіть дані про систему :", contains four input fields: "початковий рік для прогнозування" (with value '9'), "назва прогнозованого параметра" (with value "мій параметр"), "розмірність прогнозованих даних" (with value '25'), and "назва прогнозованої системи" (with value "Моя система"). A red error message "Рік повинен бути числом" is displayed next to the first field. The bottom section, "Інформація :", contains text about fuzzy implication operations and a list of classes: S-implication, R-implication, and T-implication. At the bottom right, there is a "Далі >>" button and a copyright notice "Оксана Суха (с)".

Рис. 1. Початкова сторінка

У випадку помилкових даних система повідомить користувача, що дані введені неправильно, і чекатиме виправлення помилок. Також користувач може дізнатися про існуючі класи нечітких імплікацій, з якими у цій програмі він матиме безпосередній зв'язок. У меню користувач може такі пункти, як **Файл**  відкрити з файла; за допомогою цього меню можна відкрити зі спеціального файлу дані, щоб не вводити їх постійно до системи.

The screenshot shows the same window as Figure 1, but with the "Оберіть метод нечіткої імплікації :" section expanded into a list box. The list contains the following items: "метод KleeneDienes", "метод Zadeh", "метод Mamdani" (highlighted), "метод Larsen", "метод Lukasiewicz", "метод Standard Strict", "метод Godel", "метод Gaines", "метод Goguen", "метод Kleene-Dienes-Lukasiewicz", "метод Probabilistic", "метод Limited Summa", "метод Vadi", "метод Yager", and "усі методи". The copyright notice "Оксана Суха (с)" is visible at the bottom right.

Рис. 2. Сторінка вибору методу прогнозування

Програма дає змогу легко і швидко змінювати мову інтерфейсу, що полегшує розуміння роботи продукту для широкого кола споживачів. Після переходу на наступну сторінку (рис.2) користувач побачить список можливих методів, а також зможе вибрати будь-який з них.

Рис.3.Сторінка введення даних для прогнозування

Наступною сторінкою для користувача є сторінка введення даних для прогнозування (рис.3). Користувачу запропоновано ввести у вектори дані для прогнозування. На цій сторінці використані валідатори, тому користувач побачить помилку, коли введе неправильне значення.

Рис. 4. Сторінка результатів прогнозування за вибраним методом

Наступною користувач побачить сторінку результатів прогнозування за обраним методом (рис.4), а якщо ж він обрав усі методи, тоді ця сторінка не відобразиться, але відобразиться наступна (рис.5). На цій сторінці користувач зможе побачити та оцінити числові спрогнозовані дані за певним методом і порівняти їх з реальними даними, а також визначити абсолютну та відносну похибки прогнозування.

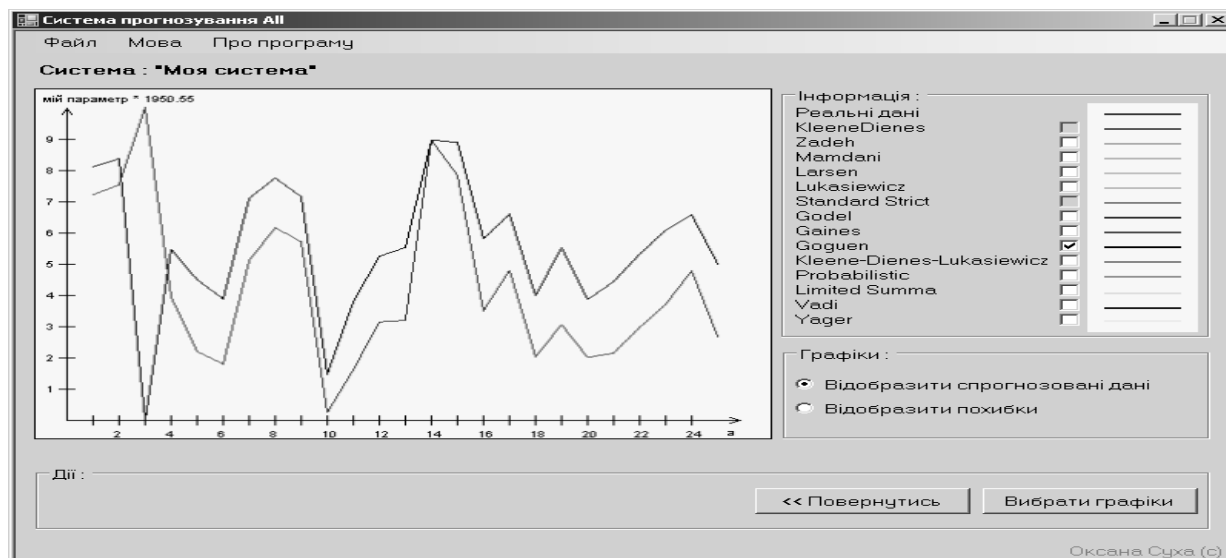


Рис. 5. Сторінка графіків результатів для усіх прогнозованих методів

На сторінці графічних результатів (рис.5) відображені графіки результатів прогнозування, а також графік реальних даних. Користувач на лівій панелі може легко вибрати ті графіки, які його цікавлять і подивитись, наскільки він відхиляється від реальних тестових даних. Отже, користувач може візуально обрати для своєї системи найкращий метод прогнозування.

Для тестування програмного продукту використано дані за декілька років, що характеризують стік вуглецю у фітомасу лісів. Ці дані характеризують фізичні явища, що не мають дуже значних стрибків.

Висновок

Розроблена програма була ретельно відлагоджена та відтестована. В процесі експериментування прогнозуючої програми довелось записувати і аналізувати можливі варіанти поведінки програми. Було досліджено, який з методів нечіткого виводу є найкращим, що підтверджено графічно. Графічне відображення дає змогу зробити висновок про те, який метод є найкращим для вибраної користувачем системи.

1. Борисов В.В., Круглов В.В., Федулов А.С. *Нечеткие модели и сети*. – М.: Горячая линия-Телеком, 2007. – 284 с. 2. Мацяшек, Лешек А. *Анализ и проектирование информационных систем с помощью UML*, 3-е изд.: Пер. с англ. – М.: ООО «И.Д.Вильямс», 2008. – 816с.: ил. – Парал. тит. англ. 3. *Прикладные нечеткие системы: Перевод с япон./ К. Асаи, Д. Вапада, С. Иваи и др.; под ред. Т. Тэрано, К. Асаи, М. Сугено*. – М.: Мир, 1993. 4. <http://www.intuit.ru/department/ds/fuzzysets/> – Яхьяева Г.Э. *Основы теории нечетких множеств, Интернет-университет информационных технологий*. – ИНТУИТ.ру, БИНОМ.Лаборатория знаний, 2008. –, 320 с. 5.://offline.computerra.ru/2001/415/1. – Сергей Гриняев *Нечеткая логика в системах управления, журнал "Компьютерра" №38 от 08 октября 2001 года*.