

УДК 519.254

Л. М. Божуха

Дніпропетровський національний університет імені Олеся Гончара

ПРО ПОБУДОВУ КОМПЛЕКСНОЇ ОЦІНКИ У СТРУКТУРІ ІЄРАРХІЇ ФАКТОРІВ

Розроблено обчислювальну технологію побудови вузлових точок функції належності для подальшого використання трапецієподібного числа лінгвістичного класифікатора при аналізі кількісних характеристик без допомоги експерта, для побудови комплексної оцінки на відношеннях між факторами у структурі ієрархії цих факторів. Розроблено програмне забезпечення та виконано практичну апробацію методики з підтвердженням її ефективності.

Ключові слова: *лінгвістичний критерій; унімодальна гистограма; стандартний 01-класифікатор; функція належності.*

Разработана вычислительная технология построения узловых точек функции принадлежности для дальнейшего использования трапециевидного числа лингвистического классификатора при анализе количественных характеристик без помощи эксперта, для построения комплексной оценки на отношениях между факторами в структуре иерархии этих факторов. Разработано программное обеспечение и выполнена практическая апробация методики с подтверждением ее эффективности.

Ключевые слова: *лингвистический критерий; унимодальная гистограмма; стандартный 01-классификатор; функция принадлежности.*

A computational technology was developed for constructing node points of the membership function for the further use of the trapezoidal number of a linguistic classifier in the analysis of quantitative characteristics without the assistance of an expert, to construct a comprehensive assessment of the relationships between factors in the structure of the hierarchy of these factors. The software has been developed and a practical testing of the methodology has been performed to confirm its effectiveness.

Keywords: *linguistic criterion; unimodal histogram; standard 01-classifier; membership function.*

Вступ. При аналізі параметрів різної природи виникає проблема якісної інтерпретації чисельних характеристик. Задача може бути

ускладнена наявністю даних різних знаків та великою кількістю показників, які змінюються незалежно один від одного, і виписати якусь залежність не стає можливим. Серед багатьох методів з'єднання усіх чисельних характеристик в одну характеристику можна виділити лінгвістичний критерій [1].

При створенні лінгвістичного критерію користуються поняттями нечіткої множини, які характеризується функцією належності. Основною проблемою при побудові функції належності можуть стати висновки експерта або групи експертів, які належать до непрямих методів побудови критерію. Експертні дані можуть бути лише вихідною інформацією для подальшої обробки інформації і можна вважати їх більше теоретичними. Функція належності, яка може бути побудована на основі цих експертних висновків, не може бути перевірена на адекватність безпосередньо засобами теорії. Складністю при побудові функції належності є наявність експертної оцінки для існуючих чисельних даних. Тому на етапі обрання відповідної функції належності при побудові класифікатора важливим є обрання вузлових точок [2].

Результати побудови найпростішої комплексної оцінки на основі трапецієподібного числа лінгвістичного класифікатора з використанням показника Херста для виявлення трендовості показників представлено у роботі [3].

Постановка проблеми. З метою забезпечення збору, обробки, збереження та аналізу інформації довільних даних та розробки рекомендацій для прийняття ефективних управлінських рішень поставлено задачу розробити інформаційно-аналітичну систему. Для розробки рекомендацій щодо прийняття рішень використати методику створення функції належності при відсутності експертної оцінки та виконати побудову найпростішої комплексної оцінки на основі трапецієподібного числа лінгвістичного класифікатора. Використання запропонованої комбінації методів може розширити область якісних критеріїв оцінки з лінгвістичним класифікатором при аналізі кількісних характеристик.

Основна частина. Ступінь належності функції лінгвістичному критерію розглядається на заданій реальній множині. Тому і інтенсивність належності можна визначити з гістограми факторів, аналізуючи кути нахилу відрізків, які з'єднують інтервали розподілу. Для побудови класифікатора використовується метод знаходження вузлових точок для функції належності при відсутності експертної оцінки та формування зон впевненості та невпевненості лінгвістичного

критерію за пропорцією $\alpha : u : \beta$ залежно від обраних вузлових точок [2].

Гістограму $G = \{X, N, \Delta, n\}$ визначимо як математичний об'єкт, де $X = [x_{\min}, x_{\max}]$ – інтервал аналізу параметра, N – кількість даних обраного параметра або фактора, $\Delta = (x_{\max} - x_{\min}) / (1 + \log_2 N)$ – крок гістограми, $n = (n_1, n_2, \dots, n_M)$ – вектор кількості входження даних у відповідну комірку гістограми розмірності M .

По вигляду гістограми, яка приведена до вигляду щільності $F(x_i) = \frac{n_i}{(N \Delta)}$, $i = \overline{1, M}$ при $x_j = x_{\min} + j\Delta$, $j = \overline{0, M}$ ідентифікується ймовірнісний закон розподілу. Розглядається найпростіший випадок, коли для фактора існує унімодальна гістограма і за цією гістограмою стоїть нормальний розподіл. У якості функції належності можна обрати стандартний 01-класифікатор з трьома рівнями «Низький (Н) – Середній (С) – Високий (В)».

На кожному інтервалі аналізу параметра гістограми $[x_{i-1}; x_i]$ позначимо точки $y_i = (x_i - x_{i-1})/2$, $i = \overline{1, M}$ та значення $F(y_i)$, $i = \overline{1, M}$. По центральних точках $(y_i, F(y_i))$, $i = \overline{1, M}$ наведеної гістограми побудуємо ламану, відрізки якої мають різні кути нахилу до осі абсцис. Кути відрізків ламаної φ_i , $i = \overline{1, M}$, приймають від'ємні $(\varphi_{\max}^-, \varphi_{\min}^-)$ або додатні значення $(\varphi_{\max}^+, \varphi_{\min}^+)$ залежно від знаходження цих відрізків відносно центрального піка гістограми. Чисельні значення $K1 = (\varphi_{\max}^+ + \varphi_{\min}^+) / 2$, $K2 = (\varphi_{\max}^- + \varphi_{\min}^-) / 2$ надають можливість проаналізувати інформацію про суттєві зміни кута нахилу відрізків ламаної. Для гістограми з невеликою швидкістю зростання і спадання даних доцільним є введення додаткових змінних $K1$ і $K2$, які є важливими характеристиками для визначення вузлів m_1, m_2, m_3 . Чисельні значення вузлів m_1, m_2, m_3 можуть бути прийняті як початковий крок при формуванні лінгвістичного критерію при відсутності експертної оцінки. При діленні відрізків

$[m_i, m_{i+1}]$, $i=1, 2$ на зони абсолютної впевненості, зниженої впевненості та абсолютної невпевненості є можливість побудови лінгвістичного критерію на функціях належності μ_i ($i = \overline{1, 3}$), операції з якими можна звести до операцій з їх вершинами.

Змінюючи положення та кількість вузлових точок класифікатора, можна отримувати нові функції належності з нестандартною пентшкалою на 01-носії. Так, наприклад, при побудові критеріїв з трьома зонами у пропорції $\alpha : u : \beta$, де параметр $u \geq 0$ виражає глибину невпевненості, параметр $k = (m_3 - m_1) / (\alpha + u + \beta)$, трапецієподібне число $(a_{1i}, a_{2i}, a_{3i}, a_{4i})$ для $i = \overline{1, 3}$ може бути обчислено так:

$$((2m_i - uk - 2\alpha k)/2, (2m_i - uk)/2, (2m_i + uk)/2, (2m_i + uk + 2\beta k)/2) .$$

Побудова класифікатора за наведеною схемою може надати можливість побудувати зони після аналізу кількісних характеристик та сформулювати правило обрання довжин зон за пропорцією $\alpha : u : \beta$, а не використовувати експертну оцінку для формування лінгвістичної змінної «рівень фактора» для вимірних змінних. У більш складному випадку, коли симетрії не існує у функції належності, можна проаналізувати характер даних гістограми: параметри пропорції α, u, β при побудові критеріїв можна обрати з урахуванням отриманих вузлів m_1, m_2, m_3 гістограми та середньоквадратичного відхилення від середнього σ .

Для побудови комплексної оцінки проведено етап аналізу кількісних характеристик факторів. На наступному етапі необхідно проаналізувати відношення між факторами у структурі ієрархії цих факторів.

Розглянемо модифікацію схеми агрегування даних на одному рівні ієрархії при аналізі довільної системи із зафіксованим набором її показників [3].

Нехай математична модель: $FSM = \langle G, L, S \rangle$, де G – деревоподібна ієрархія факторів.

Набір якісних оцінок рівнів кожного фактора в ієрархії G має вигляд:

$$L = \{ \text{низький рівень } (H), \text{середній рівень } (C), \text{високий рівень } (B) \}.$$

Система переваг одних факторів іншим для одного рівня ієрархії факторів:

$$S = \left\{ F_{k,i} \Re F_{k,j} \mid \Re = \left\{ \begin{array}{l} > \text{відношення переваги,} \\ \approx \text{відношення рівноваги} \end{array} \right\} \right\}.$$

Деревоподібна ієрархія факторів G може бути описана орієнтованим графом без циклів, петель, горизонтальних ребер у межах одного рівня ранжування, який містить одну кореневу вершину:

$G = \langle F, V \rangle$, де $F = \{ F_i \}$ – множина вершин факторів, $V = \{ V_i \}$ – множина дуг.

Вибрано ряд окремих показників, про які можна сказати, що вони якнайкраще характеризують окремі сторони і при цьому утворюють якусь закінчену сукупність, що дає вичерпне уявлення в цілому. Вибір системи показників для аналізу може бути індивідуальним для кожної системи. Значимість тих чи інших показників для оцінки системи різна і тому перед експертом постає важке завдання відбору і ранжування чинників аналізу. Показники, які класифіковані за групами, можуть утворювати ієрархію, але в найпростішому випадку вони просто складають невпорядкований набір.

Розглянемо деревовидну ієрархію F . Дерево можна розширити, якщо додавати до графа G нові вузли. На ієрархії критеріїв можна обрати систему відношень переваг на основі експертної оцінки груп факторів. Всі фактори рівнів ієрархії знаходяться у відношенні рівноваги або переваги S .

Формується лінгвістична змінна «Рівень фактора» з термножиною значень L . Лінгвістичні оцінки рівнів факторів відповідають набору якісних оцінок L .

Для запису якісних характеристик параметра рівня надалі використовується стандартний кількісний вигляд у формі відповідної функції належності. Функція належності $\mu^*(x)$ представляється стандартним п'ятирівневим 01-класифікатором (з трьома зонами у пропорції $\alpha : u : \beta$).

За кожним показником $F^* = (F_1^*, F_2^*, \dots, F_i^*)$ на обраному підрівні графа G відомі лінгвістичні оцінки $L^* = (L_1^*, L_2^*, \dots, L_i^*)$ та визначена вагова система Фішберна $P^* = (p_1^*, p_2^*, \dots, p_i^*)$. Показник підрівня F^* характеризується своєю лінгвістичною оцінкою, яка визначається

функцією належності на 01-носії $x \in [0;1]$ за допомогою OWA-оператора Ягера.

Обчислення проводяться послідовно знизу вгору по усіх рівнях ієрархії G та отримуємо функцію належності фактору F та його інтерпретацію: Н – Низький; С – Середній; В – Високий.

Для того щоб визначити рівень кожного з отриманих факторів, необхідно визначити ступінь схожості трапецієподібного числа

$\left(a_1^*, a_2^*, a_3^*, a_4^* \right)$ відповідного рівня і стандартного

трапецієподібного числа $\left(b_1^*, b_2^*, b_3^*, b_4^* \right)$ за допомогою міри

розпізнавання рівня (різновидність міри Хемінга):

$$0 \leq \nu = 1 - \max_i \left| a_i^* - b_i^* \right| \leq 1.$$

Висновки. Запропоновано інформаційну технологію аналізу кількісних характеристик факторів, відношення між факторами у структурі ієрархії цих факторів та побудови комплексної оцінки. Технологія передбачає можливість створення функції належності при відсутності експертної оцінки та виконання побудови найпростішої комплексної оцінки.

Засобами розробленої системи проведено аналіз даних. Дослідження кількісних характеристик надає можливість заміни експертної оцінки на правило обрання довжин зон за пропорцією $\alpha : u : \beta$ лінгвістичної змінної «рівень фактора». Алгоритм побудови змінної агрегування даних проходить за напрямом дуг графа ієрархії при застосуванні OWA-оператора Ягера з ваговими коефіцієнтами у згортці у вигляді коефіцієнтів Фішберна і кожному рівню відповідає функція належності.

Використання інформаційно-аналітичної системи надає можливість подальшого аналізу даних для коригування області існування критеріїв комплексної оцінки з лінгвістичним класифікатором при використанні кількісних характеристик показників.

Бібліографічні посилання

1. Недосекин А. О. Комплексная оценка риска банкротства корпорации на основе нечетких описаний. URL: http://sedok.narod.ru/sc_group.html

2. Божуха Л. М., Подберезный Д. В. Про особливості процесу аналізу гістограм при побудові трьохрівневого критерію на функціях

належності // Збірник наукових праць. IV Всеукраїнська науково-практ. конф. молодих учених та студентів «Інтелектуальні технології в системному програмуванні», Секція Штучний інтелект: методи та засоби. Хмельницький: ХНУ, 2015. С. 96–98.

3. Божуха Л. М. Про систему переваг лінгвістичного класифікатора. // Системні технології: зб. наук. праць НМЕТАУ. – Дніпропетровськ, 2017. № 2(109). С. 35–42.

Надійшла до редколегії 01.07.2017