

## РОЗРОБКА МЕТОДИКИ ОЦІНЮВАННЯ ЗНАЧЕНЬ ФУНКЦІЇ НАЛЕЖНОСТІ НА ОСНОВІ ГРУПОВОЇ ЕКСПЕРТИЗИ У МЕТОДІ НЕЧІТКОГО ДЕРЕВА РІШЕНЬ

Швед А. В. – д-р техн. наук, професор, професор кафедри інженерії програмного забезпечення Чорноморського національного університету імені Петра Могили, Миколаїв, Україна.

### АНОТАЦІЯ

**Актуальність.** Останнім часом нечіткі дерева рішень набули широкого застосування при вирішенні задач класифікації та регресії. У випадку відсутності об'єктивної інформації для побудови функції належності елементів вузлам дерева, єдиним шляхом отримання інформації є залучення експертів. У випадку залучення групи фахівців виникає задача агрегування думок експертів з метою синтезу групового рішення. Об'єктом дослідження є групові експертні оцінки ступеню належності елемента заданому класу, атрибуту, що потребують структуризації та агрегування при побудові та аналізі нечіткого дерева рішень.

**Метою роботи** є розробка методики визначення значень функції належності елемента заданому класу (атрибуту) за результатами опитування групи експертів при побудові та аналізі нечітких дерев рішень.

**Метод.** Методика дослідження ґрунтується на комплексному застосуванні математичного апарату теорії правдоподібних та парадоксальних міркувань та методів нечіткої логіки для вирішення задачі агрегування нечітких експертних оцінок значень атрибутів (ознак) класифікації при побудові та аналізі нечіткого дерева рішень. Запропонований підхід використовує механізм комбінування експертних свідочств, сформованих в рамках гібридної моделі Дезера-Смарандаке, на основі правила перерозподілу конфліктів PCR5 для побудови групового рішення.

**Результати.** Розглянуті питання структуризації нечіткої експертної інформації та запропонована методика синтезу групової експертної оцінки відносно значень атрибутів (ознак) класифікації при побудові та аналізі нечіткого дерева рішень.

**Висновки.** Дістали подальшого розвитку моделі та методи структуризації та синтезу групових рішень в умовах нечіткої експертної інформації. На відміну від існуючих експертних методів визначення значень функції належності в умовах групового вибору, запропонований підхід дозволяє синтезувати групове рішення з різним значенням конфліктної маси при комбінуванні вихідних експертних свідочств. Такий підхід дозволяє коректно агрегувати як узгоджені, так і суперечливі (конфліктні) експертні свідочства.

**КЛЮЧОВІ СЛОВА:** теорія правдоподібних та парадоксальних міркувань, нечіткі дерева рішень, правило перерозподілу конфліктів.

### АБРЕВІАТУРИ

ДР – дерева рішень;

ЕП – експертні переваги;

ЕС – експертні свідочства;

ТДС – теорія Дезера-Смарандаке, теорія правдоподібних та парадоксальних міркувань;

ТДШ – теорія Демпстера-Шейфера, теорія свідочств;

ІІІ – штучний інтелект;

### НОМЕНКЛАТУРА

$A$  – множина атрибутів (ознак) класифікації,  $a_i \in A$ ;

$B_i$  – компонента профілю ЕП, кожен елемент якої відображає вибір експерта  $E_i$  відносно значень термів лінгвістичної змінної  $\beta$  для заданої сукупності об'єктів  $O$ ;

$P^{gr}$  – груповий профіль ЕП;

$D^\Omega$  – множина всіх можливих підмножин, які можуть бути сформовані на множині  $\Omega$ , на основі операцій об'єднання та перетину, включаючи порожню множину;

$E$  – група експертів;

$Gr$  – синтаксичне правило, що породжує нові терми;

$H_i$  – компонента профілю ЕП, кожен елемент якої відображає оцінку упевненості експерта  $E_i$  в тому, що

для заданої сукупності об'єктів  $O$  лінгвістична змінна  $\beta$  набуває визначені в  $B_i$  значення;

$M_i$  – компонента профілю ЕП, кожен елемент якої містить числові значення основної маси ймовірності (функції належності) для визначених в  $B_i$  термів лінгвістичної змінної  $\beta$ ;

$Mp$  – множина належностей;

$Mr$  – семантичне правило генерації нечітких множин для кожного терму лінгвістичної змінної;

$O$  – множина аналізованих об'єктів класифікації;

$P$  – множина індивідуальних профілів ЕП;

$T(\beta)$  – терм-множина лінгвістичної змінної  $\beta$ ;

$V_{a_i}$  – множина значень атрибуту  $a_i$ ;

$U_k$  – область визначення змінної  $k$  (універсум);

$\tilde{X}$  – нечітка множина;

$\alpha$  – найменування нечіткої змінної;

$\beta$  – найменування лінгвістичної змінної;

$\mu_{\tilde{X}}(u)$  – характеристична функція належності

елемента  $u$  нечіткій множині  $\tilde{X}$ ;

$\Omega$  – основа задачі;

$d_c$  – атрибут-рішення;

$d_{metric}$  – метрика відстані між групами ЕС;

$m_k(Y_k)$  – основна маса ймовірності, призначена множині  $Y_k$ ;

$m_{d+l}(\Omega)$  – основна маса ймовірності, призначена множині  $\Omega$ ;

$w_i$  – коефіцієнт компетентності експерта  $E_i$ ;

$\oplus$  – правило комбінування ЕС.

## ВСТУП

Сучасні тенденції розвитку інформаційних технологій сприяють активній інтеграції методів ШІ у всі сфери людської діяльності. Використання технологій ШІ дозволяє швидше та ефективніше створювати більш точні моделі процесів, що протікають у складних соціальних, економічних, технічних, організаційних та інших системах під впливом сукупності різного роду факторів зовнішнього та внутрішнього оточення. Ключовою особливістю технологій ШІ є здатність навчатися, накопичувати знання та застосовувати їх. Саме машинне навчання, як один із ключових напрямків розвитку ШІ, дозволяє створювати на основі отриманих знань, досвіду та нових даних моделі здатні до навчання та адаптації для вирішення поставленої задачі. Однією з основних задач машинного навчання є задача класифікації, суть якої полягає у формуванні розбиття вихідної сукупності об'єктів на визначені класи відповідно до заданого набору ознак. Даний тип завдань у машинному навчанні відноситься до розділу контрольованого машинного навчання (навчання з учителем), і передбачає наявність навчальної вибірки для виявлення закономірностей у вихідному наборі даних та побудови вирішальних правил класифікації на їх основі.

Серед методів класифікації досить широкого поширення набув метод ДР [1–3]. Дерева рішень є графічним методом побудови класифікаційної моделі, що легко інтерпретується та дозволяє обробляти як категоріальні, так і інтервальні дані. Даний метод не вимагає спеціальної підготовки вихідних даних і дозволяє оперувати великими обсягами даних (*Big data*), їх можна ефективно застосовувати до даних з пропущеними значеннями атрибутів. У той же час дерева чутливі до зміни в навчальній вибірці, що може призвести до коригування побудованої моделі (класифікаційних правил). У процесі побудови та навчання дерева необхідно дотримуватись балансу між точністю та складністю одержуваної ієрархічної структури (можуть утворюватись занадто складні конструкції, які недостатньо повно відображають наявні дані, внаслідок чого може виникати проблема перенавчання дерева) [4].

Метод ДР набув широкого застосування в галузі машинного навчання, прогнозного моделювання та прийняття рішень у різних галузях: медичні дослідження [5–6], фінансовий сектор, маркетинг [7], діагностика несправностей у технічних системах [8], та ін.

Останнім часом активного поширення набув метод побудови ДР на основі нечіткого підходу [9–10].

**Об'єктом дослідження** є групові експертні оцінки ступеню належності елемента заданому класу, атрибуту,

що потребують структуризації та агрегування при побудові та аналізі нечіткого ДР.

**Предметом дослідження** є моделі та методи структуризації групових експертних оцінок на основі математичного апарату теорії правдоподібних та парадоксальних міркувань та нечіткої логіки.

**Метою роботи** є розробка методики визначення значень функції належності елемента заданому класу (атрибуту) за результатами опитування групи експертів при побудові та аналізі нечітких ДР.

## 1 ПОСТАНОВКА ЗАДАЧІ

Для побудови ДР вихідна сукупність даних може бути сформована на основі різних джерел, і являти собою статистичну, аналітичну, експериментальну / емпіричну інформацію, що отримана на основі методів спостереження (реєстрації, моніторингу), вимірювань (експериментів, тестів), тощо. У випадку повної відсутності об'єктивної вихідної інформації єдиним шляхом отримання даних є залучення сторонніх осіб (інтерв'ю, опитування, фокус-групи, методи експертних оцінок, тощо). Експертні знання можуть бути отримані у формі індивідуальних та групових експертних оцінок. Зазвичай для аналізу проблемної ситуації залучають групу експертів (фахівців певної предметної області), оскільки групова експертиза дає можливість отримати більш об'єктивну оцінку на основі аналізу певної сукупності індивідуальних думок експертів.

При аналізі та побудові ДР (при вирішенні задач класифікації та регресії) для графічного подання вихідної сукупності даних, відображення їх семантики використовується реляційна система – таблиця, рядки якої відповідають аналізованим об'єктам, а стовпчики – ознакам (критеріям, атрибутам) цих елементів. В комірку на перетині  $j$ -го рядка та  $l$ -го стовпчика відображається значення  $l$ -ї ознаки для  $j$ -го елемента, таким чином кожен рядок таблиці відображає один об'єкт (приклад) і відповідні йому значення атрибутів. Останній рядок зазвичай являє собою атрибут-рішення.

Таким чином, сукупність вихідних даних, яку необхідно класифікувати на основі визначеної множини критеріїв (атрибутів), можна подати у формі  $DT = (O, A, V, d_c, C, f)$ , де  $O = \{o_j \mid j = \overline{1, z}\}$  – не порожня скінчена множина аналізованих об'єктів;  $A = \{a_l \mid l = \overline{1, m}\}$  – не порожня скінчена множина примітивних атрибутів (ознак);  $V = \bigcup_{a_l \in A} V_{a_l}$ ,  $V_{a_l}$  – множина значень атрибуту  $a_l$  (область атрибуту  $a_l$ );  $d_c$  ( $|d_c| = 1$ ) являє собою атрибут-рішення, який характеризує можливі класи, до яких може бути віднесено об'єкт  $o_j \in O$ ;  $C$  – множина значень атрибуту  $d_c$ ;  $f$  – інформаційна функція, така, що  $\forall a_l \in A, o \in O, f(o, a_l) \in V_{a_l}$ .

Припустимо, що  $\exists a_i \in A$ , для яких значення множини  $V_{a_i}$  формується на основі групової експертної оцінки. Задача полягає в агрегуванні відповідних значень релевантних атрибутів  $a_i^i(o_j)$ , що формуються на основі оцінок експертів  $E_i$ ,  $i = \overline{1, n}$ , і синтезу групової оцінки:  $\forall t = \overline{1, k} : agr(a_i^i(o_j)) \rightarrow a_i^{gr}(o_j)$ ,  $o_j \in O$ ,  $1 \leq k \leq m$ . Агрегування групових експертних оцінок здійснюється для кожного атрибуту окремо.

## 2 ОГЛЯД ЛІТЕРАТУРИ

В даний час запропоновано значну кількість алгоритмів побудови дерев рішень [1–3, 11–14], серед яких найбільшого поширення набули методи ID3 [11, 12], C4.5 [13] та CART [14]. Більшість відомих алгоритмів, при пошуку оптимальної структури, використовують стратегію жадібного пошуку. У [8] для синтезу дерева запропоновано використовувати стохастичні, евристичні методи багатовимірної безградієнтної оптимізації, що дозволяє отримувати структуру дерева з кращими апроксимаційними властивостями.

Класичний метод побудови ДР передбачає, що кожен аналізований об'єкт належить конкретному задалегідь визначеному класу (вузлу). Однак у реальній практиці можуть виникати ситуації, при яких не завжди вдається досягти однозначної належності аналізованого об'єкта до певної категорії (класу). У цьому випадку виникає невизначеність при спробі віднести об'єкт до одного із заданих класів, або встановити чітке та однозначне значення класифікаційного атрибуту (ознаки). Математичний апарат нечіткої логіки дозволяє ефективно моделювати ситуації, за яких об'єкт може належати як одному, так і декільком класам одночасно, але з різним ступенем, який можна охарактеризувати значенням функції належності об'єкту до заданого класу [15–16]. Зазвичай нечіткість проявляється у ситуаціях, коли на основі якісних оцінок оцінюються значення кількісних параметрів досліджуваного об'єкту, процесу або явища.

Застосування нечітких ДР (*fuzzy decision trees*) дозволяє для кожного класифікаційного атрибуту (ознаки) виділити декілька лінгвістичних значень і встановити відповідні ступені належності об'єктів (прикладів) до них [9–10]. Таким чином, кожний об'єкт може мати властивості більше одного лінгвістичного значення, що характеризує певний заданий атрибут.

## 3 МАТЕРІАЛИ І МЕТОДИ

Дерева рішень складаються із двох основних компонентів: процедури побудови символічного дерева та процедури виведення для прийняття рішення. У загальному випадку процес побудови ДР складається з наступних послідовних етапів: вибір атрибуту розбиття; вибір критерію зупинення навчання; вибір методу відсікання гілок; оцінка точності побудованої моделі.

У нечіткому ДР кожний атрибут розглядається як лінгвістична змінна, яка може приймати деяке задалегідь зумовлене вербальне значення, одне чи декілька. Кожному значенню лінгвістичної змінної відповідає певна нечітка множина зі своєю функцією належності.

Введемо деякі позначення. Припустимо задана деяка універсальна множина  $U$ , тоді нечітка множина  $\tilde{X} \subseteq U$  є множиною всіх впорядкованих пар виду [15–16]:

$$\tilde{X} = \{(U, \mu_{\tilde{X}}(u))\}, \quad u \in U, \quad \mu_{\tilde{X}}(u) \in Mp. \quad (1)$$

Нечітки та лінгвістичні змінні використовуються для опису нечітких множин на основі слів та / або словосполучень природної людської мови, що є більш природним для людини, ніж оперувати кількісними значеннями.

Нечітка змінна задається трійкою [15–16]:

$$\langle \alpha, U_\alpha, \tilde{X} \rangle. \quad (2)$$

В якості значень лінгвістичної змінної використовуються словесний (вербальний) опис різних об'єктів, процесів та явищ.

В теорії нечітких множин лінгвістична змінна визначається наступним кортежем [15–16]:

$$\langle \beta, T(\beta), U_\beta, Gr, Mr \rangle. \quad (3)$$

Функція належності будується, зазвичай, на основі статистичних даних (частотний метод), або на основі оцінок експертів (одного чи групи). У разі відсутності достатньої кількості статистичної інформації єдиним можливим способом побудови функції належності є проведення експертного опитування.

Умовно методи побудови функцій належності можна поділити на дві групи: прямі та непрямі [17–18]. У прямих методах ступінь належності задається безпосередньо експертами (метод відносних частот, та ін.). У непрямих методах значення функції належності визначаються виходячи із задалегідь сформульованих умов відповідно до заданого алгоритму (метод парних порівнянь, та ін.).

Ступінь суб'єктивності одержаних ЕП можна зменшити в умовах групової експертизи. У цьому випадку виникають дві основні задачі: оцінка узгодженості ЕП і побудова узагальненої оцінки.

Розглянемо основні положення підходу, при якому значення лінгвістичної змінної, що ставиться у відповідність до деякого атрибуту ДР, визначається на основі групової експертизи.

Припустимо, група експертів  $E = \{E_i \mid i = \overline{1, n}\}$  оцінюючи значення лінгвістичної змінної  $\beta$  для заданої сукупності об'єктів  $O = \{o_j \mid j = \overline{1, z}\}$  сформувала множину профілів ЕП виду  $P = \langle B, L \rangle$ . За умови, що

сукупність значень лінгвістичної змінної  $\beta$  утворює терм-множину  $T(\beta) = \{t_l \mid l = \overline{1, k}\}$ , перша компонента кортежу являє собою сукупність  $B = \{B_i \mid i = \overline{1, n}\}$ , кожен елемент якої  $B_i = \{b_j^i \mid j = \overline{1, z}\}$  може приймати значення відповідного терму або декількох термів, що набуває лінгвістична змінна  $\beta$  для об'єкту  $o_j$ .

Друга компонента  $L$  містить оцінки експертів, на основі яких можуть бути обраховані числові значення функцій належності для визначених в  $b_j^i$  термів лінгвістичної змінної  $\beta$  відносно об'єкта  $o_j$ , що встановлені експертом  $E_i$ .

Задача полягає у синтезі групового профілю  $P^{gr} = \langle B^{gr}, L^{gr} \rangle$ . Кожен елемент  $b_j^{gr} \in B^{gr}$ ,  $j = \overline{1, z}$  відображає групове рішення відносно значення що набуває лінгвістична змінна  $\beta$  для об'єкту  $o_j$ , і формується на основі агрегування оцінок  $B_i = \{b_j^i \mid j = \overline{1, z}\}$ ,  $\forall i = \overline{1, n}$ . Елементи  $L_j^{gr} \in L^{gr}$ ,  $j = \overline{1, z}$  містять значення функції належності визначених в  $b_j^{gr}$  термів лінгвістичної змінної  $\beta$ , отримані за результатами агрегування індивідуальних ЕП. Якщо задано декілька лінгвістичних змінних  $\{\beta_1, \dots, \beta_q\}$ , значення яких встановлюються шляхом експертного опитування, то буде сформовано відповідна кількість групових профілів  $\{P_1^{gr}, \dots, P_q^{gr}\}$ .

Для синтезу групової оцінки в роботі використано математичний апарат ТДС в рамках якого множина  $T(\beta)$  розглядається як основа задачі  $\Omega$  [19]. На відміну від моделі Шейфера (ТДШ), основа задачі  $\Omega$  розглядається виключно як множина вичерпних елементів. Такі елементи можуть формально описувати неточні, нечіткі (розмиті) поняття та знання про оточуючий світ, наприклад, відтінки кольору, градації віку, температури і т.п. Внаслідок чого деякі елементи можуть перекривати один одного, і відповідно неможливо досягти їх взаємної виключності.

В рамках нотації ТДС на множині вихідних даних  $\Omega = T(\beta)$  може бути сформовано  $|D^\Omega|$  підмножин на основі операцій  $\cup$  та  $\cap$ , включаючи порожню множину  $\emptyset$  [19].

Таким чином експертом можуть бути виділені підмножини  $X_i \subseteq D^\Omega$ ,  $i = \overline{1, |D^\Omega|}$ , що задовольняють умовам [19]:

1.  $X_i = \{\emptyset\}$ ;
2.  $X_i = \{t_l\}$  – експертом обраний (оцінений) один елемент  $t_l \in \Omega$ .
3.  $X_i = \{t_l \mid l = \overline{1, p}\}$ ,  $p < k$  – експертом виділено  $p$  елементів  $t_l \in \Omega$ .

$$4. X_i = \Omega = \{t_l \mid l = \overline{1, k}\};$$

5. якщо  $X_i, X_j \subset D^\Omega$ , тоді  $X_i \cap X_j \in D^\Omega$  та  $X_i \cup X_j \in D^\Omega$ .

ТДС оперує двома видами моделей: вільна та гібридна модель [19]. Вільна модель містить всі можливі підмножини основи задачі. Гібридна модель визначається із вільної моделі шляхом введення обмежень на деякі підмножини елементів  $X_i$  із множини  $D^\Omega$ , за умови, що  $X_i \neq \emptyset$ . Це пояснюється тим, що в реальних задачах немає необхідності визначати основні маси ймовірності всім можливим підмножинам  $D^\Omega$ , тому, що завжди можливо існування елементів, які є взаємовиключними.

Будемо вважати, що в процесі експертного опитування на основі множини  $T(\beta)$  експертом  $E_i$  можуть бути сформовані підмножини, що відображають його судження, або одноелементні,  $|b_j^i| = 1$ , або використовуючи операцію  $\cup$ . Таким чином кількість можливих виділених підмножин становить  $2^\Omega$ , враховуючи порожній вибір ( $b_j^i = \{\emptyset\}$ ).

Обмеження, що накладаються, і умови проведення процедури експертного опитування можуть привести до наступних ситуацій.

*Випадок 1.* Експерт може встановити відповідність об'єкта до декількох термів лінгвістичної змінної, але з різним ступенем належності (безпосередньо задати значення функцій належності).

За результатами експертного опитування формуються профілі ЕП виду  $P = \langle B, M \rangle$ . Перша компонента кортежу являє собою сукупність  $B = \{B_i \mid i = \overline{1, n}\}$ , кожен елемент якої  $B_i = \{b_j^i \mid j = \overline{1, z}\}$  відображає вибір експерта  $E_i$  відносно значень, що може набувати лінгвістична змінна  $\beta$  для об'єкту  $o_j$ . При цьому  $b_j^i = \{Y_k \mid k = \overline{1, d}\}$ ,  $|Y_k| = 1$ ,  $d < 2^{|\Omega|}$  являє собою більш ніж одне значення (декілька термів лінгвістичної змінної). Друга компонента кортежу являє собою сукупність  $M = \{M_i \mid i = \overline{1, n}\}$ , кожен елемент якої  $M_i = \{m_j^i \mid j = \overline{1, z}\}$  містить основну масу ймовірності  $m_j^i$  призначену для  $b_j^i$ , що відповідає числовим значенням функцій належності для визначених в  $b_j^i$  термів лінгвістичної змінної  $\beta$  відносно об'єкта  $o_j$ , які встановлено експертом  $E_i$ . При цьому  $m_j^i = \{m_k \mid k = \overline{1, d}\}$ ,  $d < 2^{|\Omega|}$ ,  $\forall i, j: |m_j^i| = |b_j^i|$ ,  $i = \overline{1, n}$ ,  $j = \overline{1, z}$ . Таким чином, елемент  $m_k = m_j^i$  містить значення функції належності терму  $Y_k \subseteq b_j^i$  лінгвістичної змінної  $\beta$  для об'єкту  $o_j$  задане експертом  $E_i$ .

Виходячи з нотації ТДС, кожен елемент  $Y_k \subseteq b_j^i$  повинен підпорядковуватися системі правил (4); значення, що містить  $m_k = m_j^i$  відповідає системі умов:

$$0 \leq m_k(Y_k) \leq 1, \forall (Y_k \in b_j^i), m_k(\emptyset) = 0, \sum_{Y_k \in b_j^i} m_k(Y_k) = 1. \quad (5)$$

Агрегування суджень експертів виконується у відповідності до запропонованої процедури:

1.1 Структуризація задачі. Для кожного  $o_j \in O$  на основі сформованого набору профілів ЕП Р формується кортеж  $\langle B_j^*, M_j^*, T_j^* \rangle$ , де  $B_j^* = \{b_j^i\}$ ,  $i = \overline{1, n}$  – сукупність тверджень групи експертів;  $M_j^* = \{m_j^i\}$ ,  $i = \overline{1, n}$  – сукупність відповідних значень функції належності, сформованих на основі суджень групи експертів; множину  $T_j^* = \{t_f^{*j} | f = \overline{1, v}\}$ ,  $v \leq |T(\beta)|$ , утворюють елементи множини  $T(\beta)$  на основі яких утворені елементи  $b_j^i \in B_j^*$ ,  $i = \overline{1, n}$ .

1.2 Визначення порядку агрегування (комбінування) ЕС. Для комбінування обирається пара ЕС  $b_j^i, b_j^h \in B_j^*$ , таких, що при  $i \neq h$   $\min d_{metric}(m_j^i, m_j^h) \in [0; 1]$ , де  $d_{metric}$  – деяка метрика відстані між групами ЕС [20–23].

1.3 Агрегування ЕС здійснюється шляхом комбінування отриманих основних мас ймовірності  $M_j^* = \{m_j^i | i = \overline{1, n}\}$  та  $B_j^* = \{b_j^i\}$ , за всіма експертами  $E_i$ ,  $i = \overline{1, n}$ , для кожного  $o_j \in O$  окремо:  $M_j^{comb} = m_j^1 \oplus m_j^2 \oplus \dots \oplus m_j^n$ ,  $B_j^{comb} = b_j^1 \oplus b_j^2 \oplus \dots \oplus b_j^n$ , де  $m_j^i \in M_j^*$ ,  $b_j^i \in B_j^*$ . В якості правила комбінування запропоновано використовувати правило PCR5 [19].

Комбінована маса ймовірності  $m_{PCR5}(C)$  згідно з правилом перерозподілу конфліктів PCR5 ( $\forall C \subset D^\Omega \setminus \{\emptyset\}$ ) визначається відповідно до

$$m_{PCR5}(C) = m_{12}(C) + \sum_{\substack{Y \in D^\Omega \setminus \{X\} \\ Y \cap X = \emptyset}} \left[ \frac{m_1(X)^2 \cdot m_2(Y) + m_2(X)^2 \cdot m_1(Y)}{m_1(X) + m_2(Y) + m_2(X) + m_1(Y)} \right]. \quad (6)$$

Правило комбінування PCR5 перерозподіляє основну масу ймовірності, віднесена до порожньої множини, на підмножини, залучені до локальних конфліктів, пропорційно основним масам ймовірності цих

підмножин. Ця властивість дозволяє коректно поводитися з узагальненою масою ймовірності, що віднесена до порожніх перетинів. Правило комбінування PCR5 дозволяє обробляти ЕП у ситуації, коли конфліктна маса впевненості досягає максимально можливого значення (набуває 1), при цьому будуть розраховані комбіновані основні маси ймовірності для всіх виділених експертами підмножин, включаючи одноелементні.

*Випадок 2.* Експерт може встановити відповідність об'єкта до декількох термів лінгвістичної змінної із різним ступенем упевненості у своєму виборі.

За результатами експертного опитування формуються профілі ЕП виду  $P = \langle B, N \rangle$ .

Перша компонента кортежу являє собою сукупність  $B = \{B_i | i = \overline{1, n}\}$ , кожен елемент якої  $B_i = \{b_j^i | j = \overline{1, z}\}$  відображає вибір експерта  $E_i$  відносно значень, що може набувати лінгвістична змінна  $\beta$  для об'єкта  $o_j$ . При цьому  $b_j^i = \{Y_k | k = \overline{1, d}\}$ ,  $|Y_k| > 1$ ,  $d < 2^{|\Omega|}$  ( $|Y_k| > 1$ , то всі елементи в групі  $Y_k$  є рівнозначними, може бути заданий тільки один терм із групи, наприклад, якщо  $Y_k = \{t_1, t_2\}$ , то мається на увазі, що  $\beta$  набуває тільки одне із можливих значень, або  $t_1$ , або  $t_2$ ).

Друга компонента кортежу являє собою сукупність  $N = \{H_i | i = \overline{1, n}\}$ ,  $H_i = \{h_j^i | j = \overline{1, z}\}$ . Кожен елемент  $X_k \in h_j^i$ ,  $k = \overline{1, d}$  відображає оцінку упевненості експерта  $E_i$  в тому, що для об'єкта  $o_j$  лінгвістична змінна  $\beta$  набуває значення терму  $Y_k \in b_j^i$ . Виходячи з нотації ТДС, кожен елемент  $Y_k \subseteq b_j^i$  повинен підпорядковуватися системі правил (4); кожен елемент  $X_k \in h_j^i$  може бути виражений в рамках певної заданої шкали, використовуючи діапазон чисел від 0 до деякого заданого  $N$  ( $N > 0$ ).

Агрегування суджень експертів виконується у відповідності до запропонованої процедури:

2.1 Структуризація задачі. Для кожного  $o_j \in O$  на основі сформованого набору профілів ЕП Р формується кортеж  $\langle B_j^*, H_j^*, T_j^* \rangle$ , де  $B_j^* = \{b_j^i\}$ ,  $i = \overline{1, n}$  – сукупність тверджень групи експертів;  $H_j^* = \{h_j^i\}$ ,  $i = \overline{1, n}$  – сукупність відповідних оцінок впевненості експертів в своєму виборі; множину  $T_j^* = \{t_f^{*j} | f = \overline{1, v}\}$ ,  $v \leq |T(\beta)|$ , утворюють елементи множини  $T(\beta)$  на основі яких сформовані  $b_j^i \in B_j^*$ ,  $i = \overline{1, n}$ .

2.2 Визначення основних мас ймовірності, що відповідають виділеним підмножинам  $Y_k \subseteq b_j^i$ ,  $\forall b_j^i \in B_j^*$ .

Для кожної сформованої системи підмножин  $b_j^i = \{Y_k | k = \overline{1, d}\}$  буде отримано вектор  $m_j^i = \{m_k | k = \overline{1, d+1}\}$ , елементи якого відповідають умові (3) і визначаються відповідно до виразу [24]:

$$m_k(Y_k) = \frac{R_1}{R_2 + \sqrt{d}}, \quad m_{d+1}(\Omega) = \frac{\sqrt{d}}{R_2 + \sqrt{d}}, \quad (7)$$

де  $R_1$  відповідає значенню  $X_k \in h_j^i$ ;  $R_2$  відповідає значенню  $\sum_{k=1}^d X_k$ ;  $d$  – загальна кількість сформованих  $E_i$  підмножин  $Y_k \in b_j^i$ .

Значення  $m_{d+1}(\Omega)$  є мірою невизначеності вибору експерта  $E_i$  відносно значень, що може набувати лінгвістична змінна  $\beta$  для об'єкту  $o_j$ .

Якщо задано вектор коефіцієнтів компетентності експертів  $W = \{w_i | i = \overline{1, n}\}$ , то

$$R_1 = X_k \cdot w_i; \quad R_2 = \sum_{k=1}^d X_k \cdot w_i. \quad (8)$$

На основі отриманих оцінок  $m_j^i$  для кожного аналізованого  $o_j \in O$  формується вектор  $M_j^* = \{m_j^i | i = \overline{1, n}\}$ .

2.3 Визначення порядку агрегування (комбінування) ЕП: обирається така пара  $b_j^i, b_j^h \in B_j^*$ , що при  $i \neq h \min d_{metric}(m_j^i, m_j^h) \in [0; 1]$ .

2.4 Агрегування ЕП здійснюється шляхом комбінування отриманих основних мас ймовірності  $M_j^* = \{m_j^i | i = \overline{1, n}\}$  та  $B_j^* = \{b_j^i\}$ , за всіма експертами  $E_i$ ,  $i = \overline{1, n}$ , для кожного  $o_j \in O$  окремо. В якості правила комбінування запропоновано використовувати правило PCR5.

Результатом комбінування є вектор  $B_j^{comb} = \{Y_k^{comb} | k = \overline{1, v}\}$ ,  $v \leq 2^{|\Omega|}$  і вектор  $M_j^{comb} = \{m(Y_k^{comb}) | k = \overline{1, v}\}$ , відповідно.

2.5 Розрахунок верхньої і нижньої межі ймовірності для кожного  $t_f^{*j} \in T_j^*$  у відповідності до виразів [19]:

– функція впевненості (довіри)  $Bel: D^\Omega \rightarrow [0; 1]$ :

$$Bel(\theta) = \sum_{b_j^{*i} \subseteq \theta, b_j^{*i} \in B_j^*} m(b_j^{*i}); \quad (9)$$

– функція правдоподібності  $Pl: D^\Omega \rightarrow [0; 1]$ :

$$Pl(\theta) = \sum_{b_j^{*i} \cap \theta \neq \emptyset, b_j^{*i} \in B_j^*} m(b_j^{*i}). \quad (10)$$

Значення функції  $Bel(\theta)$  відображають ступінь підтримки, що надається підмножині  $\theta$ .

Значення функції  $Pl(\theta)$  відображають ступінь потенційної підтримки, яка може бути надана підмножині  $\theta$ .

Формування інтервалів  $[Bel(\{t_f^{*j}\}), Pl(\{t_f^{*j}\})]$  для кожного  $t_f^{*j} \in T_j^*$ ,  $f = \overline{1, v}$ .

#### 4 ЕКСПЕРИМЕНТИ

Продемонструємо запропоновані вище підходи на прикладі вирішення задачі синтезу групових рішень щодо значень лінгвістичної змінної  $\beta = \langle \text{Якість} \rangle$ , яка може набувати наступні значення  $T(\beta) = \langle \text{«низька»}, \text{«середня»}, \text{«висока»}, \text{«дуже висока»} \rangle$ .

*Приклад 1.* Припустимо, група експертів  $E = \{E_i | i = \overline{1, n}\}$ ,  $n = 3$ , оцінюючи значення заданої лінгвістичної змінної  $\beta$  для заданої сукупності аналізованих об'єктів  $O = \{o_j | j = \overline{1, z}\}$ ,  $z = 3$ , сформувала профілі ЕП  $P = \langle B, M \rangle$ , де  $B = \{B_i | i = \overline{1, n}\}$ ,  $B_i = \{b_j^i | j = \overline{1, z}\}$ . Результати експертного опитування наведені у табл. 1.

Таблиця 1 – Профілі ЕП (приклад 1)

p	$E_1$				$E_2$			
	$t_1$	$t_2$	$t_3$	$t_4$	$t_1$	$t_2$	$t_3$	$t_4$
$o_1$		0,2	0,8				0,8	0,2
$o_2$	0,15	0,85			0,6	0,4		
$o_3$			0,3	0,7		0,6	0,4	
p	$E_3$							
	$t_1$	$t_2$	$t_3$	$t_4$				
$o_1$		0,7	0,3					
$o_2$	0,7	0,3						
$o_3$			0,6	0,4				

Таблиця 1 відображає лише суб'єктивні судження 3 експертів відносно значення лінгвістичної змінної  $\beta$  для заданої множини аналізованих об'єктів  $O = \{o_j | j = \overline{1, 3}\}$ .

*Приклад 2.* Припустимо, група експертів  $E = \{E_i | i = \overline{1, n}\}$ ,  $n = 3$ , оцінюючи значення заданої лінгвістичної змінної  $\beta$  для аналізованого об'єкту  $o_1$ , сформувала профілі експертних переваг виду  $P = \langle B, H \rangle$ .

Результати експертного опитування наведені у табл. 2.

Таблиця 2 – Профілі ЕП (приклад 2)

P	E <sub>1</sub>		E <sub>2</sub>		E <sub>3</sub>	
	$Y_{k \in b_j^1}$	$X_{k \in h_j^1}$	$Y_{k \in b_j^2}$	$X_{k \in h_j^2}$	$Y_{k \in b_j^3}$	$X_{k \in h_j^3}$
o <sub>1</sub>	{t <sub>2</sub> }	4	{t <sub>3</sub> }	8	{t <sub>2</sub> }	7
	{t <sub>3</sub> }	8	{t <sub>4</sub> }	2	{t <sub>3</sub> , t <sub>4</sub> }	2

Значення  $X_k \in h_j^i$  визначались за 10-ти бальною шкалою (0 – відповідає найнижчому ступеню упевненості; 10 – експерт абсолютно впевнений в своєму виборі).

**Приклад 3.** Розглянемо приклад бінарної класифікації за методом нечіткого ДР при вирішенні наступної задачі: необхідно оцінити можливість надання кредиту потенційному позичальнику – фізичній особі. В якості атрибутів класифікації розглядаються:  $a_1$  – кредитний рейтинг, який визначається на основі скорингової (бальної) оцінки кредитоспроможності потенційного позичальника;  $a_2$  – платоспроможність пози-

чальника (рівень середньомісячного доходу за останні шість місяців).

Атрибут-рішення  $d_c$  має два значення: «Видати кредит» – 1 / «Відмовити у видачі кредиту» – 0.

Припустимо, атрибут  $a_1$  = «Кредитний рейтинг» може приймати значення  $t_1$  = «низький»,  $t_2$  = «середній»,  $t_3$  = «високий»; атрибут  $a_2$  = «Платоспроможність» може приймати значення  $t_1$  = «низька»,  $t_2$  = «середня»,  $t_3$  = «висока».

За результатами обробки персональних даних п'яти потенційних позичальників було побудовано скорингову карту та отримано чисельні оцінки атрибутів  $a_1$  та  $a_2$ . На основі проведеного експертного опитування групою із трьох експертів за отриманими чисельними оцінками класифікаційних атрибутів було визначено ступінь належності кожного об'єкту (потенційного позичальника) до відповідних значень атрибутів.

Результати експертної оцінки наведені в табл. 3.

Таблиця 3 – Профілі ЕП за атрибутами (лінгвістичними змінними)  $a_1$  та  $a_2$

P	Атрибут $a_1$									Атрибут $a_2$									$d_c$
	E <sub>1</sub>			E <sub>2</sub>			E <sub>3</sub>			E <sub>1</sub>			E <sub>2</sub>			E <sub>3</sub>			
	$t_1$	$t_2$	$t_3$	$t_1$	$t_2$	$t_3$	$t_1$	$t_2$	$t_3$	$t_1$	$t_2$	$t_3$	$t_1$	$t_2$	$t_3$	$t_1$	$t_2$	$t_3$	
$a_1$	0,2	0,8		0,45	0,55		0,6	0,4		0,6	0,4		0,4	0,6		0,7	0,3		0
$a_2$	0,15	0,85		0,6	0,4			0,3	0,7		0,35	0,65	0,4	0,6			0,3	0,7	1
$a_3$		0,3	0,7		0,4	0,6	0,2	0,8			0,2	0,8		0,6	0,4		0,7	0,3	1
$a_4$	0,8	0,2		0,35	0,65		0,3	0,7			0,35	0,65		0,8	0,2		0,4	0,6	1
$a_5$	0,3	0,7		0,55	0,45		0,65	0,35		0,3	0,7		0,55	0,45		0,65	0,35		0

## 5 РЕЗУЛЬТАТИ

Розглянемо реалізацію запропонованих алгоритмів для синтезу групового рішення стосовно значення лінгвістичної змінної  $\beta$  для об'єкту  $o_1$ .

Аналізуючи дані таблиці 1 можна бачити, що для  $o_1 \in O$  на основі множини значень  $T = \{t_1, t_2, t_3, t_4\}$ , що може приймати лінгвістична змінна  $\beta$  = «Якість», групою експертів була сформована сукупність  $B_1^* = \{b_j^i\}$  і сукупність оцінок  $M_1^* = \{m_j^i\}$ ,  $i = \overline{1, n}$ , де

$$b_1^1 = \{\{t_2\}, \{t_3\}\}; \quad m_1^1 = \{0,2, 0,8\};$$

$$b_1^2 = \{\{t_3\}, \{t_4\}\}; \quad m_1^2 = \{0,8, 0,2\}$$

$$b_1^3 = \{\{t_2\}, \{t_3\}\}; \quad m_1^3 = \{0,7, 0,3\}.$$

На основі значень множини  $B_1^*$  сформуємо множини  $T_1^* = \{t_2, t_3, t_4\}$ .

Таблиця 4 – Ступінь перетину виділених експертами підмножин (приклад 1)

В	Експерт $E_2$		
	$b_j^i$	$t_3$	$t_4$
Експерт $E_1$	$t_2$	$E_1(t_2) \cap E_2(t_2) = \emptyset$	$E_1(t_2) \cap E_2(t_4) = \emptyset$
	$t_3$	$E_1(t_3) \cap E_2(t_3) = t_3$	$E_1(t_3) \cap E_2(t_4) = \emptyset$

Виходячи із даних табл. 4 маємо три локальні конфлікти:  $E_1(t_2) \cap E_2(t_3)$ ,  $E_1(t_2) \cap E_2(t_4)$  та  $E_1(t_3) \cap E_2(t_4)$ .

Розрахуємо комбіновані значення основної маси ймовірності виділених підмножин за привалом (6):

$$m_{123}\{t_2\} = 0,17; \quad m_{123}\{t_3\} = 0,77; \quad m_{123}\{t_4\} = 0,06.$$

Таким чином, маємо  $B^{gr} = \{t_2, t_3, t_4\}$ ,  $M^{gr} = \{0,17; 0,77; 0,06\}$  відповідно.

За даними таблиці 2 експертами була сформована сукупність  $B_1^* = \{b_j^i\}$  і сукупність оцінок  $O_1^* = \{o_j^i\}$ ,  $i = \overline{1, n}$ , де

$$b_1^1 = \{\{t_2\}, \{t_3\}\}; \quad o_1^1 = \{4; 8\};$$

$$b_1^2 = \{\{t_3\}, \{t_4\}\}; \quad o_1^2 = \{8; 2\};$$

$$b_1^3 = \{\{t_2\}, \{t_3, t_4\}\}; \quad o_1^3 = \{7; 2\}.$$

На основі значень множини  $B_1^*$  сформуємо множини  $T_1^* = \{t_2, t_3, t_4\}$ .

Розрахована основна маса ймовірності виділених фокальних елементів відповідно до виразу (7):

$$E_1: m(t_2) = 0,3; \quad m(t_3) = 0,6; \quad m(\Omega) = 0,1;$$

$$E_2: m(t_3) = 0,7; \quad m(t_4) = 0,18; \quad m(\Omega) = 0,12;$$

$$E_3: m(t_2) = 0,67; \quad m(t_3, t_4) = 0,19; \quad m(\Omega) = 0,14.$$

Комбінована основної маси ймовірності виділених підмножин визначена у відповідності до (6):

$$\begin{aligned} m_{123}\{t_2\} &= 0,316; & m_{123}\{t_3\} &= 0,63; \\ m_{123}\{t_4\} &= 0,05; & m_{123}\{t_3, t_4\} &= 0,0023; \\ m_{123}\{\Omega\} &= 0,0017. \end{aligned}$$

Значення функцій (9) та (10) для кожного елемента множини  $T_1^*$ :

$$\begin{aligned} t_2: \begin{cases} Bel(\{t_2\}) = 0,316; \\ Pl(\{t_2\}) = 0,3177; \end{cases} & t_3: \begin{cases} Bel(\{t_3\}) = 0,63; \\ Pl(\{t_3\}) = 0,634; \end{cases} \\ t_4: \begin{cases} Bel(\{t_4\}) = 0,05; \\ Pl(\{t_4\}) = 0,054. \end{cases} \end{aligned}$$

Із наведених результатів видно, що найбільше значення функції довіри та правдоподібності отримав терм  $t_2$ ; найменше – терм  $t_4$ .

Для переходу від інтервальних до точкових оцінок введемо коефіцієнт песимізму  $\gamma \in [0, 1]$ :

$$\gamma \cdot Bel(\theta) + (1 - \gamma) \cdot Pl(\theta), \quad (11)$$

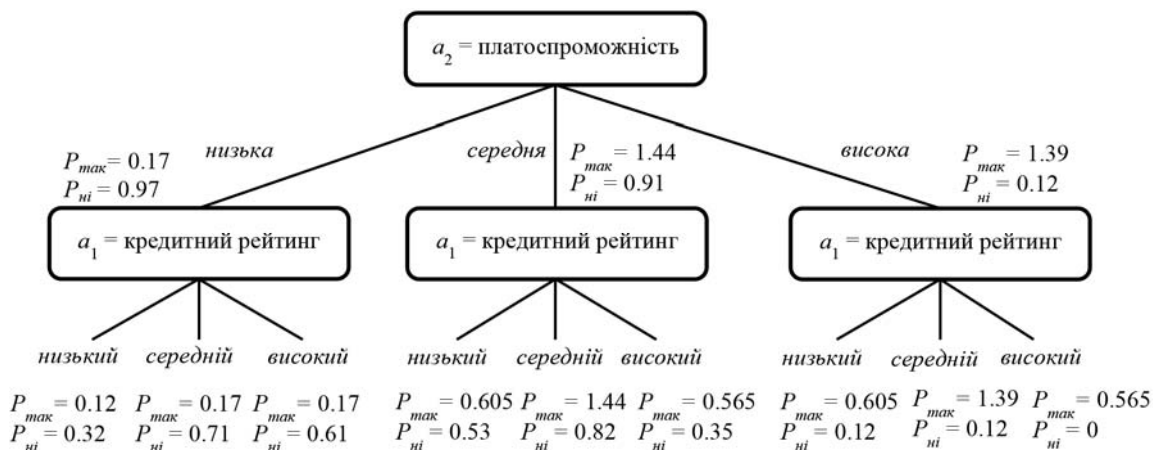


Рисунок 1 – Графічне подання побудованого нечіткого ДР

## 6 ОБГОВОРЕННЯ

Запропонована в роботі процедура визначення значень функції належності на основі групової експертної оцінки ґрунтується на математичному апараті нечіткої логіки та теорії правдоподібних та парадоксальних міркувань. Агрегування ЕП (свідoctв) відбувається шляхом їх комбінування на основі обраного правила [19, 25–27]. Основною проблемою, що виникає в процесі комбінування ЕС, отриманих на основі двох або більше незалежних груп свідoctв (експертів), є поводження з конфліктами. Причиною виникнення конфліктів є неузгодженість окремих груп ЕС, коли окремі вихідні ЕС (фокальні елементи) не перетинаються.

Виділяють локальний конфлікт, який виникає в результаті порожнього перетину двох вихідних фокальних елементів, і глобальний (сума всіх локальних конфліктів). Якщо вихідні фокальні елементи перети-

наються частково, то комбінована маса ймовірності віддається підмножині, яка є результатом такого перетину. Таким чином, результуючий фокальний елемент безпосередньо залежить від ступеня перетину вихідних фокальних елементів – чим вищий ступінь перетину вихідних фокальних елементів, тим менше втрачається вихідної інформації (менше виникає локальних конфліктів) і тим достовірнішими будуть результати комбінування.

$$\mu(t_2) = 0,32; \quad \mu(t_3) = 0,63; \quad \mu(t_4) = 0,05.$$

На основі запропонованого в роботі підходу щодо визначення агрегованих експертних оцінок значень функції належності, за результатами обробки даних табл. 3 було побудовано агреговані експертні оцінки, табл. 5.

Таблиця 5 – Агреговані експертні оцінки значень класифікаційних атрибутів

$p^{gr}$	$a_1$			$a_2$			$d_c$
	$t_1$	$t_2$	$t_3$	$t_1$	$t_2$	$t_3$	
$a_1$	0,39	0,61		0,65	0,35		0
$a_2$	0,12	0,7	0,18	0,17	0,39	0,44	1
$a_3$	0,055	0,56	0,385		0,54	0,46	1
$a_4$	0,43	0,57			0,51	0,49	1
$a_5$	0,53	0,47		0,32	0,56	0,12	0

Побудоване за вихідними даними табл. 5 ДР на основі алгоритму [10] зображено на рис. 1.

наються частково, то комбінована маса ймовірності віддається підмножині, яка є результатом такого перетину. Таким чином, результуючий фокальний елемент безпосередньо залежить від ступеня перетину вихідних фокальних елементів – чим вищий ступінь перетину вихідних фокальних елементів, тим менше втрачається вихідної інформації (менше виникає локальних конфліктів) і тим достовірнішими будуть результати комбінування.

Для коректного поводження з конфліктними масами ймовірності у роботі запропоновано використовувати правила комбінування, що дозволяють врахувати ступінь перетину вихідних фокальних елементів і коректно перерозподіляти конфліктні маси ймовірності на підмножини, що залучені до локальних конфліктів [19]. Правила перерозподілу конфліктів PCR1-PCR5 можуть бути використані як на моделі Шейфера [25–27], так і на гібридній моделі Дезера-



Смарандаке [19]. Правило PCR5 є єдиним правилом, застосування якого дозволяє розділити кожен локальну конфліктну масу ймовірності на частки, які пропорційно перерозподіляються на підмножини, залучені до конфлікту, у відповідності до основних мас ймовірності вихідних фокальних елементів. В цьому випадку не відбувається втрати інформації (маси ймовірності, що відповідає порожнім перетинам фокальних елементів), також не відбувається «розмиття» комбінованих мас ймовірності на результуючих фокальних елементах (утворених шляхом об'єднання вихідних підмножин), як при використанні правил на основі диз'юнктивного консенсусу [25]. Більшість правил комбінування, що базуються на принципах кон'юнктивного консенсусу [25], також мають ряд недоліків, до яких можна віднести втрату частини інформації пов'язаної з вихідними фокальними елементами, що не перетинаються, (така маса ймовірності може відноситись до порожньої множини; до основи задачі або використовуватися для нормування отриманих результатів); так зване «стиснення» вихідних фокальних елементів (крім ситуації, коли вони ідентичні один одному) при утворенні результуючих підмножин, отриманих шляхом перетину вихідних фокальних елементів; не береться до уваги ступінь перетину вихідних фокальних елементів (за винятком правила Жанга).

Втрату вихідної інформації у процесі комбінування можна зменшити за рахунок вибору оптимального порядку комбінування ЕС з огляду на міру близькості між ними [20–23].

Ще одним способом отримання агрегованої оцінки у разі відсутності прийнятного рівня узгодженості є виявлення та виключення конфліктних ЕС, або розбиття вихідної сукупності ЕС на кілька підгруп, усередині яких свідчення характеризуються прийнятним рівнем узгодженості [28].

## ВИСНОВКИ

У роботі розглянуто задачу визначення значень функції належності аналізованій сукупності об'єктів до визначених термів лінгвістичних змінних, що ставляться у відповідність класифікаційним ознакам у методі нечіткого ДР в умовах групового вибору. Розглянуто два випадки: задані безпосередньо числові оцінки функції належності групою експертів; отримано індивідуальні ЕП щодо ступеня належності об'єкта деякому вербальному значенню лінгвістичної змінної асоційованої з класифікаційним атрибутом. Синтез групового рішення виконано на основі математичного апарату ТДС.

**Наукова новизна** отриманих результатів полягає в тому, що дістали подальшого розвитку моделі та методи структуризації та синтезу групових рішень в умовах нечіткої експертної інформації.

**Практична цінність** полягає в тому, що запропонований підхід, при синтезі групового рішення, дозволяє отримувати достовірніші результати комбінування ЕС різної структури за рахунок застосування

© Швед А. В., 2024  
DOI 10.15588/1607-3274-2024-2-11

правила PCR5, розподіляючи конфліктну масу впевненості на підмножини, залучені до локальних конфліктів. Цей підхід можна застосувати для агрегування як узгоджених, так і суперечливих ЕС.

**Перспективи подальших досліджень** полягають у дослідженні можливості застосування правил комбінування ЕС при побудові ДР в умовах інтервальної невизначеності.

## ПОДЯКИ

Робота виконана за підтримки іменної стипендії Верховної Ради України для молодих учених – докторів наук за 2023 рік.

## ЛІТЕРАТУРА

1. Decision trees as a predictive model in digital marketing / [C. Pérez-Quinde, W. Llerena-Llerena, F. Zúñiga-Vásquez, M. P. Silva], Ranganathan G., Allioui Y., Piramuthu S. (eds) // *Soft Computing for Security Applications. ICSCS 2023. Advances in Intelligent Systems and Computing.* – Springer, Singapore. – 2023. – Vol. 1449. – P. 403–414. DOI: 10.1007/978-981-99-3608-3\_28
2. Overview of use of decision tree algorithms in machine learning / [A. Navada, A. N. Ansari, S. Patil, B. A. Sonkamble] // *Control and System Graduate Research Colloquium (ICSGRC 2011): Graduate Research Colloquium, Shah Alam, Malaysia, 27–28 June 2011: proceedings.* – Shah Alam : IEEE, 2011. – P. 37–42. DOI: 10.1109/ICSGRC.2011.5991826.
3. Rokach L. *Data mining with decision trees. Theory and Applications* / L. Rokach, O. Maimon. – London : World Scientific Publishing Co, 2008. – 264 p. DOI: 10.1142/9097.9
4. Bramer M. *Avoiding overfitting of decision trees / M. Bramer.* // *Principles of data mining. Undergraduate topics in computer science.* – London : Springer, 2013. – P. 121–136. DOI: 10.1007/978-1-4471-4884-5\_9
5. Fuzzy decision trees in medical decision making support system / [V. Levashenko, P. Hrkut, A. Kovalenko, L. Kurmasheva] // *Modern technologies of biomedical engineering: the 1st International scientific and technical conference, Odesa, Ukraine, 25–27 May 2022: proceedings.* Odesa, 2022. – P. 190–198.
6. Classification of cancer data: analyzing gene expression data using a fuzzy decision tree algorithm / [S. A. Ludwig, S. Picek, D. Jakobovic], Kahraman C., Topcu Y. I. (eds) // *Operations research applications in health care management.* – Cham : Springer. – 2018. – Vol. 262. – P. 327–347. DOI: 10.1007/978-3-319-65455-3\_13
7. Zhang Z. *Applications of the decision tree in business field / Z. Zhang* // *Economic Management and Cultural Industry (ICEMCI 2021): the 3rd International Conference, Guangzhou, China, 22–24 October 2021: proceedings.* Atlantis Press International B. V., 2011. – P. 926–929. DOI: 10.2991/assehr.k.211209.151
8. Гофман С. О. Еволюційний метод синтезу дерев рішень / С. О. Гофман, А. О. Олійник, С. О. Субботін // *Штучний інтелект.* – 2011. – №2. – С. 6–14.
9. Fuzzy decision trees / [A. Altay, D. Cinar, Kahraman C., Kabak Ö. (eds)] // *Fuzzy statistical decision-making: theory and applications.* – 2016. – Vol. 343. – P. 221–261. DOI: 10.1007/978-3-319-39014-7\_13
10. Janikow C. Z. *Fuzzy decision trees: issues and methods / C. Z. Janikow* // *IEEE Transactions on Systems, Man and*

- Cybernetics. – 1998. – Vol. 28(1). – P. 1–14. DOI: 10.1109/3477.658573
11. Quinlan J. R. Induction on decision tree / J. R. Quinlan // *Machine Learning*. – 1986. – Vol. 1. – P. 81–106. DOI: 10.1007/BF00116251
12. Kantarci S. A fuzzy ID3 induction for linguistic data sets / S. Kantarci, E. Nasibov // *International Journal of Intelligent Systems*. – 2018. – Vol. 33. – P. 858–878. DOI: 10.1002/int.21971.
13. Quinlan J. R. C4.5: Programs for machine learning / J. R. Quinlan. – San Mateo: Morgan Kaufmann Publishers, 1993. – 312 p.
14. Classification and regression trees / [L. Breiman, J. H. Friedman, R. A. Olsen, C. J. Stone]. – California : Wadsworth & Brooks, 1984. – 368 p.
15. Klir G. J. Fuzzy sets and fuzzy logic: theory and applications / G. J. Klir, B. Yuan. – NJ, Prentice Hall; Upper Saddle River, 1995. – 592 p.
16. Peckol J. K. Introduction to fuzzy logic / J. K. Peckol. – Hoboken, NJ : Wiley, 2021. – 287 p.
17. Kondratenko Yu. Hesitant fuzzy information processing based on the generalized aggregation of resulting trapezoidal linguistic terms / Yu. Kondratenko, G. Kondratenko, I. Sidenko // *ICT in Education, Research and Industrial Applications (ICTERI-2019): the 15th International Conference, Kherson, Ukraine, 12–15 June 2019: CEUR workshop proceedings. Integration, Harmonization and Knowledge Transfer*. – 2019. – Vol. I. – P. 479–484.
18. Sancho-Royo A. Methods for the construction of membership functions / A. Sancho-Royo, J. Verdegay // *International Journal of Intelligent Systems*. – 1999. – Vol. 14. – P. 1213–1230. DOI: 10.1002/(SICI)1098-111X(199912)14:123.0.CO;2-5.
19. Smarandache F. Advances and applications of DSMT for information fusion / F. Smarandache, J. Dezert. – Vol. 1. – Rehoboth : American Research Press, 2004. – 760 p.
20. Bhattacharyya A. On a measure of divergence between two statistical populations defined by their probability distribution / A. Bhattacharyya // *Bulletin of the Calcutta Mathematical Society*. – 1943. – Vol. 35. – P. 99–110.
21. Cuzzolin F. A geometric approach to the theory of evidence / F. Cuzzolin // *Transactions on Systems, Man, and Cybernetics (Part C: Applications and Reviews)*. – 2007. – Vol. 38(4). – P. 522–534. DOI: 10.1109/TSMCC.2008.919174
22. Jousselme A. L. A new distance between two bodies of evidence / A. L. Jousselme, D. Grenier, E. Bossé // *Information Fusion*. – 2001. – Vol. 2. – P. 91–101. DOI: 10.1016/S1566-2535(01)00026-4
23. Tessem B. Approximations for efficient computation in the theory of evidence / B. Tessem // *Artificial Intelligence*. – 1993. – Vol. 61. – P. 315–329. DOI: 10.1016/0004-3702(93)90072-J
24. Beynon M. J. The Dempster–Shafer theory of evidence: an alternative approach to multicriteria decision modeling / M. J. Beynon, B. Curry, P. Morgan // *Omega*. – 2000. – Vol. 28(1). – P. 37–50.
25. Sentz K. Combination of evidence in Dempster-Shafer theory. Technical report SAND 2002-0835 / K. Sentz, S. Ferson. – Albuquerque : Sandia National Laboratories, 2002. – 94 p.
26. Dempster A. P. Upper and lower probabilities induced by a multi-valued mapping / A. P. Dempster // *Annals of Mathematical Statistics*. – 1967. – Vol. 38(2). – P. 325–339. DOI: 10.1214/aoms/1177698950
27. Shafer G. A mathematical theory of evidence / G. Shafer. – Princeton : Princeton University Press, 1976. – 297 p.
28. Davydenko Ye. O. Development of technique for structuring of group expert assessments under uncertainty and inconsistency / Ye. O. Davydenko, A. V. Shved, N. V. Honcharova // *Radio Electronics, Computer Science, Control*. – 2023. – Vol. 30(4). – P. 30–38. DOI: 10.15588/1607-3274-2023-4-3

Received 19.02.2024.  
Accepted 24.04.2024.

UDC 004.827:519.816

## DEVELOPMENT OF TECHNIQUE FOR DETERMINING THE MEMBERSHIP FUNCTION VALUES ON THE BASIS OF GROUP EXPERT ASSESSMENT IN FUZZY DECISION TREE METHOD

**Shved A. V.** – Dr. Sc., Professor, Professor of Department of Software Engineering, Petro Mohyla Black Sea National University, Mykolayiv, Ukraine.

### ABSTRACT

**Context.** Recently, fuzzy decision trees have become widely used in solving the classification problem. In the absence of objective information to construct the membership function that shows the degrees of belongingness of elements to tree nodes, the only way to obtain information is to involve experts. In the case of group decision making, the task of aggregation of experts' preferences in order to synthesize a group decision arises.

**The object of the study** is group expert preferences of the degree of belonging (membership function) of an element to a given class, attribute, which require structuring and aggregation in the process of construction and analysis of a fuzzy decision tree.

**Objective.** The purpose of the article is to develop a methodology for determining the membership degree of elements to a given class (attribute) based on the group expert assessment in the process of construction and analysis of fuzzy decision trees.

**Method.** The research methodology is based on the complex application of the mathematical apparatus of the theory of plausible and paradoxical reasoning and methods of fuzzy logic to solve the problem of aggregating fuzzy judgments of the classification attribute values in the process of construction and analysis of a fuzzy decision tree. The proposed approach uses the mechanism of combination of expert evidences (judgments), formed within the framework of the Dezert-Smarandache hybrid model, based on the PCR5 proportional conflict redistribution rule to construct a group solution.

**Results.** The issues of structuring fuzzy expert judgments are considered and the method of synthesis of group expert judgments regarding the values of membership degree of elements to classification attributes in the process of construction and analysis of fuzzy decision trees has been proposed.

**Conclusions.** The models and methods of structuring and synthesis of group decisions based on fuzzy expert information were further developed. In contrast to the existing expert methods for the construction of membership function in context of group decision making, the proposed approach allows synthesizing a group decision taking into account the varying degree of conflict mass in the process of combination of original expert evidenced. This approach allows to correctly aggregate both agreed and contradictory (conflicting) expert judgments.

**KEYWORDS:** the theory of plausible and paradoxical reasoning, fuzzy decision trees, proportional conflict redistribution rule.

## REFERENCES

1. Páez-Quinde C., Llerena-Llerena W., Zúñiga-Vásquez F., Silva M. P., Ranganathan G., EL Alloui Y., Piramuthu S. (eds), Decision trees as a predictive model in digital marketing, *Soft Computing for Security Applications. ICSCS 2023. Advances in Intelligent Systems and Computing*. Springer, Singapore, 2023, Vol. 1449, pp 403–414. DOI: 10.1007/978-981-99-3608-3\_28
2. Navada A., Ansari A. N., Patil S., Sonkamble B. A. Overview of use of decision tree algorithms in machine learning, *Control and System Graduate Research Colloquium (ICSGRC 2011): Graduate Research Colloquium, Shah Alam, Malaysia, 27–28 June 2011: proceedings*. Shah Alam, IEEE, 2011, pp. 37–42. DOI: 10.1109/ICSGRC.2011.5991826.
3. Rokach L., Maimon O. Data mining with decision trees. Theory and Applications. London, World Scientific Publishing Co, 2008, 264 p. DOI: 10.1142/9097.9
4. Bramer M. Avoiding overfitting of decision trees. In: Principles of data mining, *Undergraduate topics in computer science*. London, Springer, 2013, pp. 121–136. DOI: 10.1007/978-1-4471-4884-5\_9
5. Levashenko V., Hrkut P., Kovalenko A., Kurmasheva L. Fuzzy decision trees in medical decision making support system, *Modern technologies of biomedical engineering: the 1st International scientific and technical conference, Odesa, Ukraine, 25–27 May 2022, proceedings*. Odesa, 2022, pp. 190–198.
6. Ludwig S. A., Picek S., Jakobovic D., Kahraman C., Topcu Y. I. (eds) Classification of cancer data: analyzing gene expression data using a fuzzy decision tree algorithm. Operations research applications in health care management. Cham, Springer, 2018, Vol. 262, pp. 327–347. DOI: 10.1007/978-3-319-65455-3\_13
7. Zhang Z. Applications of the decision tree in business field, *Economic Management and Cultural Industry (ICEMCI 2021): the 3rd International Conference, Guangzhou, China, 22–24 October 2021: proceedings*. Atlantis Press International B. V., 2011. pp. 926–929. DOI: 10.2991/assehr.k.211209.151
8. Gofman E. A., Oliinyk A. A., Subbotin S. A. Evolutionary method of decision trees synthesis, *Artificial Intelligence*, 2011, №2, pp. 6–14.
9. Altay A., Cinar D. Fuzzy decision trees. In Kahraman C., Kabak Ö. (eds) Fuzzy statistical decision-making: theory and applications, 2016, Vol. 343, pp. 221–261. DOI: 10.1007/978-3-319-39014-7\_13
10. Janikow C. Z. Fuzzy decision trees: issues and methods, *IEEE Transactions on Systems, Man and Cybernetics*, 1998, Vol. 28(1), pp. 1–14. DOI: 10.1109/3477.658573
11. Quinlan J. R. Induction on decision tree, *Machine Learning*, 1986, Vol. 1, pp. 81–106. DOI: 10.1007/BF00116251
12. Kantarci S., Nasibov E. A fuzzy ID3 induction for linguistic data sets, *International Journal of Intelligent Systems*, 2018, Vol. 33, pp. 858–878. DOI: 10.1002/int.21971.
13. Quinlan J. R. C4.5: Programs for machine learning. San Mateo, Morgan Kaufmann Publishers, 1993, 312 p.
14. Breiman L., Friedman J. H., Olsen R. A., Stone C. J. Classification and regression trees. California, Wadsworth & Brooks, 1984. 368 p.
15. Klir G. J., Yuan B. Fuzzy sets and fuzzy logic: theory and applications. NJ, Prentice Hall; Upper Saddle River, 1995, 592 p.
16. Peckol J. K. Introduction to fuzzy logic. Hoboken, NJ, Wiley, 2021, 287 p.
17. Kondratenko Yu., Kondratenko G., Sidenko I. Hesitant fuzzy information processing based on the generalized aggregation of resulting trapezoidal linguistic terms, *ICT in Education, Research and Industrial Applications (ICTERI-2019): the 15th International Conference, Kherson, Ukraine, 12–15 June 2019: CEUR workshop proceedings. Integration, Harmonization and Knowledge Transfer, 2019, Vol. I. P. 479–484.*
18. Sancho-Royo A., Verdegay J. Methods for the construction of membership functions, *International Journal of Intelligent Systems*, 1999, Vol. 14, pp. 1213–1230. DOI: 10.1002/(SICI)1098-111X(199912)14:123.0.CO;2-5.
19. Smarandache F., Dezert J. Advances and applications of DSMT for information fusion. Rehoboth, American Research Press, 2004, Vol. 1, 760 p.
20. Bhattacharyya A. On a measure of divergence between two statistical populations defined by their probability distribution, *Bulletin of the Calcutta Mathematical Society*, 1943, Vol. 35, pp. 99–110.
21. Cuzzolin F. A geometric approach to the theory of evidence, *Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 2007, Vol. 38(4), pp. 522–534. DOI: 10.1109/TSMCC.2008.919174
22. Jousselme A. L., Grenier D., Boss'e E. A new distance between two bodies of evidence, *Information Fusion*, 2001, Vol. 2, pp. 91–101. DOI: 10.1016/S1566-2535(01)00026-4
23. Tessem B. Approximations for efficient computation in the theory of evidence, *Artificial Intelligence*, 1993, Vol. 61, pp. 315–329. DOI: 10.1016/0004-3702(93)90072-J
24. Beynon M. J., Curry B., Morgan P. The Dempster–Shafer theory of evidence: an alternative approach to multicriteria decision modeling, *Omega*, 2000, Vol. 28(1), pp. 37–50.
25. Sentz K., Ferson S. Combination of evidence in Dempster–Shafer theory. Technical report SAND 2002-0835. Albuquerque, Sandia National Laboratories, 2002, 94 p.
26. Dempster A. P. Upper and lower probabilities induced by a multi-valued mapping, *Annals of Mathematical Statistics*, 1967, Vol. 38(2), pp. 325–339. DOI: 10.1214/aoms/1177698950
27. Shafer G. A mathematical theory of evidence. Princeton: Princeton University Press, 1976, 297 p.
28. Davydenko Ye. O., Shved A. V., Honcharova N. V. Development of technique for structuring of group expert assessments under uncertainty and inconcistency, *Radio Electronics, Computer Science, Control*, 2023, Vol. 30(4), pp. 30–38. DOI: 10.15588/1607-3274-2023-4-3