

ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ ДІЯЛЬНОСТІ НАУКОВИХ УСТАНОВ МЕТОДОМ ОБРОБКИ РІЗНОРІДНОЇ СЛАБОСТРУКТУРОВАНОЇ ІНФОРМАЦІЇ

Глоба Л.С., Задосенко Б.О.

Інститут телекомунікаційних систем КПІ ім. Ігоря Сікорського, Україна

E-mail: lgloba@its.kpi.ua, zadoenko.bogdan@gmail.com

Evaluation of the effectiveness of the activities of the scientific institutions by the method of processing of different slow structured information

An approach to information processing is considered, which includes information processing method based on fuzzy tree-based knowledge bases, advanced fuzzy logic model and appropriate methods, approaches and algorithms that allow structuring, formation and tuning of fuzzy knowledge base, which is the main part of the model of scientific activity evaluation institutions and agencies.

Наукова діяльність є одним із найбільш впливових чинників розвитку економіки кожної країни. Об'єктивний аналіз ефективності діяльності наукових установ та значення їх діяльності з точки зору світової та вітчизняної науки, освіти, економіки, культури, підвищення якості життя людини є вкрай актуальним. Важливим є отримання та аналіз даних відкритим та прозорим способом із застосуванням об'єктивних методів отримання показників діяльності.

Метою досліджень, яким присвячена дана доповідь, є отримання об'єктивної оцінки ефективності діяльності наукових установ НАН України з урахуванням наявних умов їхнього кадрового, інформаційного, фінансового та матеріально-технічного забезпечення; стимулювання науково-дослідної та інноваційної діяльності; прийняття обґрунтованих управлінських, організаційних та фінансових рішень; наближення до європейської практики оцінювання ефективності діяльності наукових установ [1].

Державні установи такі як міністерства та відомства відносяться до складних організаційних систем. Вони мають розвинену ієрархічну структуру, що складається з багатьох підрозділів та підлеглих установ. Інформаційні технології складають основу інформаційно-аналітичної діяльності для задач оцінювання таких об'єктів, а обробка інформації є однією з найбільш складних частин процесу оцінювання. Це обумовлено великою кількістю складних інформаційних потоків, що циркулюють в таких об'єктах, високою динамікою зміни їх структури, яка обумовлена законодавчими змінами та наявністю на кожному рівні суб'єктивних факторів, необхідністю врахування при обробці як кількісної так і якісної інформації.

Обробка інформації в таких установах виконується експертами, які, базуючись на власному досвіді та за допомогою спрощених математичних методів, проводять оцінку загального стану результатів діяльності. При цьому необхідно визначити ключові моменти такої діяльності [2]:

- оцінювана інформація часто є неповною, частина даних подається в кількісному вигляді, частина – в якісному;
- змінність вимог до проведення обробки інформації і до моделі системи;
- робота з оцінки ситуації виконується людиною на всіх рівнях ієрархії об'єкту, тому навіть в однотипних підрозділах при однакових вхідних даних можуть бути отримані різні оцінки;
- внаслідок того, що обробка інформації здійснюється людиною така система втрачає оперативність;
- виходячи з кінцевих звітів, які подаються керівнику установи, важко простежити та

визначити ті параметри, які найбільше впливають на такий стан справ.

Внаслідок цього є актуальною наукова-технічна задача розробки підходу до обробки інформації в складних організаційних системах, який би включав метод обробки інформації на основі деревоподібних баз знань, який використовує змішану, а саме чітку і нечітку, схему виведення.

Інформація може мати не тільки кількісний вигляд, але і надаватись у вигляді знань [3], а саме: правил, евристик, наборів ранжованих альтернатив, семантичних оцінок кількісних параметрів та інших характеристик, які описуються за допомогою значень лінгвістичних змінних (ЛЗ) і висловлювань. Тому необхідно використовувати такі математичні методи і підходи, які включають засоби для представлення і оперування такою інформацією, а саме методи, що базуються на теорії нечітких множин та нечіткій логіці.

В даному дослідженні запропоновано розробку підходу до обробки інформації, до складу якого повинні входити методи та алгоритми, що дозволять впорядковувати бази знань (БЗ) введенням ієрархії правил, проводити настроювання БЗ у відповідності до вхідних даних, метод обробки

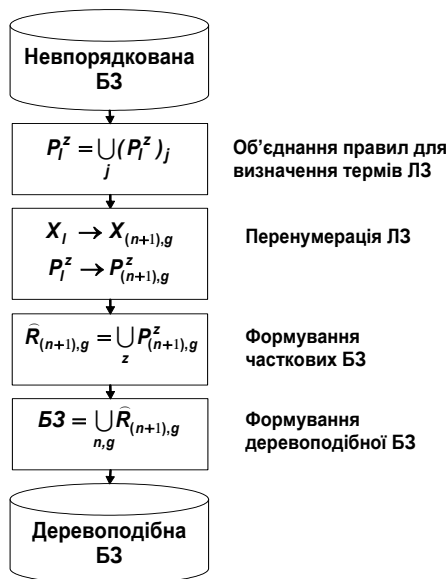


Рис.1. Кроки вирішення задачі зведення нечіткої БЗ до деревоподібної структури.

Задля зменшення кількості перебору та усунення недоліків, що пов'язані з неупорядкованістю нечіткої БЗ, сформулюємо задачу зведення неупорядкованої нечіткої БЗ до деревоподібної структури.

Постановка задачі: Дано:

1) $BZ = \bigcup_{l,z,j} (P_l^z)_j$ – неупорядкована нечітка БЗ, де

$(P_l^z)_j$ - j -е правило ($j = \overline{1, M}$, $M = M_{l_z}$) для визначення z -го терму ЛЗ з ідентифікатором l

2) $I = \{I_i | i = \overline{1, N}\}$ - множина з уведеною на ній ієрархією, що відповідає ієрархії оцінювання ситуації

3) Відповідність елементів ієрархії оцінювання ситуації лінгвістичним змінним X_i , $i = \overline{1, N}$

Отримати:

$$BZ = \bigcup_{n,g} BZ_{(n+1),g} = \bigcup_{n,g} \widehat{R}_{(n+1),g} \quad (1)$$

інформації на основі деревоподібних нечітких баз знань, який дозволяє простежувати шлях виведення в зворотному напрямку з метою визначення всієї послідовності вершин, які вплинули на остаточне значення результуючої змінної.

Для рішення цієї задачі в роботі [4] запропоновано підхід до зведення нечіткої БЗ до деревоподібної структури за рахунок використання природної ієрархічності, а саме ієрархії оцінювання ситуації, та притаманних їм евристик. Він полягає у співставленні з кожною ЛЗ, окрім тих, що відповідають вхідним змінним, частини нечіткої БЗ, яка визначатиме залежність цієї ЛЗ від ЛЗ попереднього рівня. Таку частину нечіткої БЗ будемо називати частковою нечіткою БЗ. Тоді для кожного набору правил, що відповідають одному терму ЛЗ буде проводитись композиція, яка буде полягати в знаходженні максимуму серед всіх отриманих значень функції належності (ФН).

де BZ – деревоподібна нечітка БЗ,

$$\begin{aligned}
 & \text{ЯКЩО } (X_{n1} = X_{n1}^{k_1^{z1}}) TA (X_{n2} = X_{n2}^{k_2^{z1}}) TA \dots TA (X_{nS} = X_{nS}^{k_S^{z1}}) \text{ АБО} \\
 & (X_{n1} = X_{n1}^{k_1^{z2}}) TA (X_{n2} = X_{n2}^{k_2^{z2}}) TA \dots TA (X_{nS} = X_{nS}^{k_S^{z2}}) \text{ АБО } \dots \\
 \hat{R}_{(n+1),g} &= \text{АБО } (X_{n1} = X_{n1}^{k_1^{zM^z}}) TA (X_{n2} = X_{n2}^{k_2^{zM^z}}) TA \dots TA (X_{nS} = X_{nS}^{k_S^{zM^z}}) \quad (1) \\
 & \text{ТО } X_{(n+1),g} = X_{(n+1),g}^z \\
 & z = \overline{1, Z_{(n+1),g}}
 \end{aligned}$$

де $\hat{R}_{(n+1),g}$ – часткова нечітка БЗ для визначення залежності ЛЗ $X_{(n+1),g}$ від ЛЗ $\{X_{ns} = X_{nq_s} \mid s = \overline{1, S}\}$, n – номер рівня, q_s – номер ЛЗ на рівні n , s – номер ЛЗ у рядку-кон'юнкції, $S = S_{(n+1),g}^z$ – кількість ЛЗ у рядку-кон'юнкції, $M^z = M_{(n+1),g}^z$ – кількість рядків кон'юнкцій для визначення z -го терму ЛЗ $X_{(n+1),g}$, $X_{ns}^{k_s^{zr}}$ – терм ЛЗ X_{ns} , $k_s^{zr} = k_{ns}^{g_zr}$ – номер терму.

Сформована таким чином БЗ буде мати деревоподібну структуру (рис.2). В ній можна виділити дві частини: нижню і верхню. Зауважимо, що виходячи з особливостей предметної області застосовувати автоматичні механізми формування та настроювання можливо лише для нижньої частини деревоподібної нечіткої БЗ. Це обумовлено тим, що адекватну вибірку значень вхідних та відповідних результуючих змінних можна забезпечити лише для частин нижнього рівня.

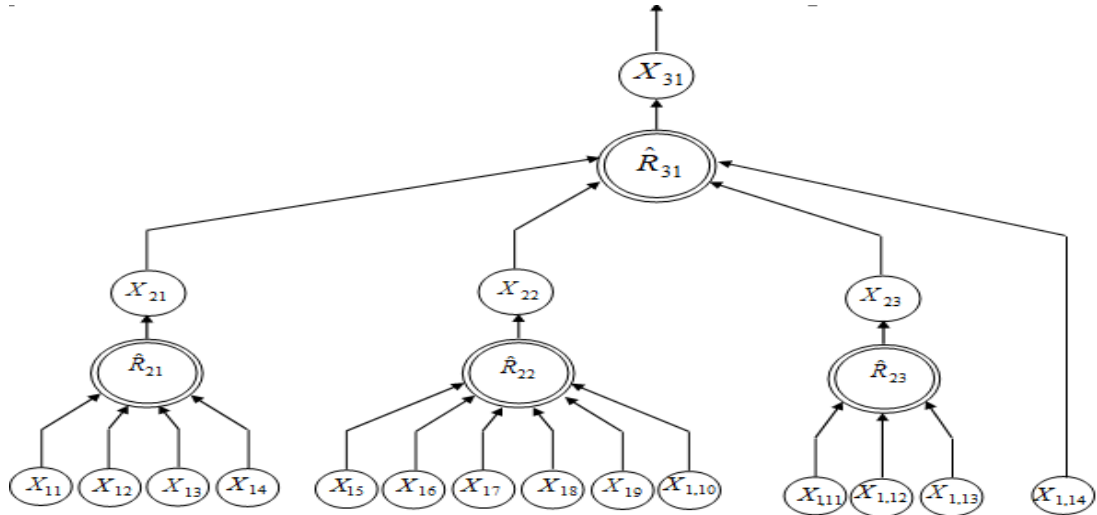


Рис. 2. Деревоподібна нечітка БЗ.

Для вирішення задач представлення знань в умовах невизначеності в працях [5,6] пропонується будувати системи нечіткого виведення на основі нечітких БЗ. Нечіткі БЗ в таких системах представляють собою сукупність лінгвістичних висловлювань типу ЯКЩО <входи>, ТО <виходи> та є частиною підсистем нечіткого виведення (ПНВ).

Рішення задач класифікації, коли результуюча змінна є якісною за своєю природою, застосовується нечітке виведення Мамдані, яке здійснюється в чотири етапи, що наведені на рис. 3:

1. Введення нечіткості.

Знаходяться ступені істинності для передумов кожного правила.

2. Нечітке виведення.

Знаходяться рівні «відсікання» для передумов кожного з правил (з використанням операції мінімум), та знаходяться «усічені» функції належності для вихідних змінних.

3. Композиція.

Проводиться об'єднання усічених функцій з використанням операції максимум, що призводить до одержання підсумкової нечіткої підмножини для змінної виходу з відповідною функцією належності.

4. Приведення до чіткості.

Додатковий етап, який використовується при необхідності перетворення нечіткої множини виведення до чіткого числа.

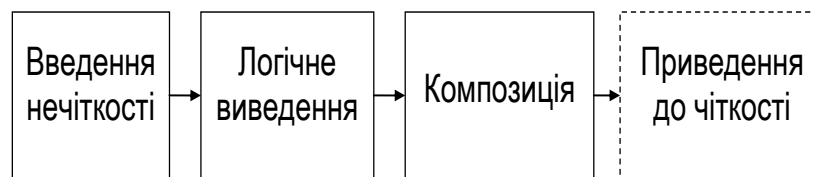


Рис. 3. Загальна схема нечіткого виведення.

Запропонований підхід у роботі [2] дозволяє проводити аналіз якості наукових досліджень у наукових установах та разом із застосуванням онтологій, які описуватимуть набори параметрів, за якими оцінюється наукова діяльність установ, дозволить розробити методіку, що отримуватиме більш достовірний інтегральний узагальнюючий показник – рейтинг.

Висновки. В роботі визначено підхід до обробки інформації, який включає метод обробки інформації на основі нечітких деревоподібних баз знань, удосконалену нечітку логічну модель та відповідні способи, підходи та алгоритми, що дозволяють провести структурування, формування та настроювання нечіткої БЗ, яка є головною частиною моделі оцінки наукової діяльності установ та відомств. Також визначено підхід до зведення невпорядкованої нечіткої бази знань до деревоподібної структури, за рахунок використання ієрархічності оцінювання інформації.

Література

1. Методика оцінювання ефективності діяльності наукових установ Національної академії наук України ЗАТВЕРДЖЕНО постановою Президії НАН України від 15.03.2017 № 75 зі змінами, внесеними постановою Президії НАН України від 11.07.2018 № 241.
2. Терновой М.Ю. -ПІДХІД ДО ОБРОБКИ ІНФОРМАЦІЇ НА ОСНОВІ ДЕРЕВОПОДІБНИХ БАЗ ЗНАНЬ ЗІ ЗМІШАНОЮ СХЕМОЮ ВИВЕДЕННЯ.
3. Згуровский М.З., Доброногов А.В., Померанцева Т.Н. Исследование социальных процессов на основе методологии системного анализа. – К.: Наукова думка, 1997. – 224 с.
4. Глоба Л.С., Терновой М.Ю. Оптимизация использования базы знаний экспертной системы введением иерархии правил. // Труды седьмой международной научно-практической конференции "Современные информационные и электронные технологии" (СИЭТ-2006). – Одеса: ОНПУ. – 2006. – С.166.
5. Герасимов Б.М., Грабовский Г.Г., Рюмшин Н.А. Нечеткие множества в задачах проектирования, управления и обработки информации. –К.: Техніка, 2002. –140 с.
6. Ротштейн А.П. Интеллектуальные технологии идентификации: нечеткие множества, генетические алгоритмы, нейронные сети. – Винница: УНИВЕРСУМ, 1999. –320 с.