© І. М. Цідило, 2011
ISSN 2076-8184. Інформаційні технології і засоби навчання. 2011. №5 (25).
Режим доступу до журналу: http://www.journal.iitta.gov.ua

УДК 62:50

Цідило Іван Миколайович, кандидат педагогічних наук, доцент кафедри комп'ютерних технологій Тернопільського національного педагогічного університету імені Володимира Гнатюка, м.Тернопіль, e-mail: tsidylo@mail.ru

НЕЧІТКЕ МОДЕЛЮВАННЯ РЕЙТИНГУ НАУКОВИХ ПУБЛІКАЦІЙ

Анотація

Розглядається проблема прийняття рішень в умовах невизначеності на основі застосування нечітких множин і нечіткої логіки. Побудовано нечітку експертну систему, що дозволяє провести оцінку якості публікацій, журналів, науковців і наукових організацій, формалізація якої складна й неможлива в умовах неповної або неточної інформації. У процесі дослідження показано, що нечіткі системи Сугено, налаштовані за допомогою алгоритмів навчання, мають найвищу точність завдяки відносно простим програмам навчання і практичній реалізації. Описано структуру і функції нечіткої експертної системи для оцінки якості наукових публікацій. Проведено тестування контролера на тестовій вибірці, що наведена у вихідних даних.

Ключові слова: нечітка логіка, експертна оцінка, нечітка система Сугено, нечіткий контролер.

Постановка проблеми в загальному вигляді. Нечітка логіка, на якій засновується нечітке керування, ближча за суттю до мислення людини і природних мов, ніж традиційні логічні системи. Вона, в основному, забезпечує ефективні засоби відображення невизначеностей і неточностей. Методи нечітких множин особливо корисні за відсутності точної математичної моделі функціонування системи. Теорія нечітких множин дає можливість застосувати для прийняття рішень неточні і суб'єктивні експертні знання про предметну галузь без формалізації їх у вигляді традиційних математичних моделей.

З використанням теорії нечітких множин вирішуються питання узгодження суперечливих критеріїв прийняття рішень, створення логічних регуляторів систем.

Нечіткі множини дають змогу застосовувати лінгвістичний опис складних процесів, встановлювати нечіткі відношення між поняттями, прогнозувати поведінку системи, формувати множину альтернативних дій, виконувати формальний опис нечітких правил прийняття рішень. Вихідною інформацією під час розробки алгоритму служать дані спеціально організованого опитування досвідчених операторів або спеціалістів, тобто, на базі експертних оцінок предметної галузі, за результатами яких після спеціальної обробки формується алгоритм.

Методи експертних оцінок набули бурхливого розвитку лише в останні десятиріччя, що зумовлено, насамперед, посиленням інноваційного розвитку і зростанням у результаті цього ступеня невизначеності через відсутність необхідної інформації. Експертні оцінки в теперішній час є найбільш розповсюдженим способом отримання й аналізу якісної інформації.

«Проведення наукових досліджень завжди супроводжується пошуком наукової інформації з певної тематики. При чому існує необхідність впорядкування наукових публікацій за їх цінністю для досліджень. Ці задачі вирішують бібліометричні і вебометричні методи. У світовій практиці існує велика кількість індексів формування рейтингів, призначених для ранжирувань публікацій, авторів та журналів. Індекс цитування, індекс Хірша, еіджен-фактор та інші індекси в своїй основі містять показники кількості цитувань іншими публікаціями. Але подібні розрахунки важко привести для українських видань, адже норми оформлення посилань часто змінюють, крім того не всі редакції пильно перевіряють їх дотримання. Іншою проблемою використання таких індексів є відсутність всіх електронних версій видань, що знижує цитування, тобто перевіряти дієвість подібних рейтингів буде можливо, лише через тривалий час після створення електронних аналогів більшості паперових видань» [1, с. 10]. Тому, за основу нашого дослідження пропонується використовувати метод обчислення рейтингів наукових публікацій, журналів, науковців і наукових організацій, що запропонований у роботі [1, с. 10] і поєднує в собі експертне оцінювання, з наданням ваг експертам за рахунок власних рейтингів і кількості переглядів публікацій користувачами.

Під час розв'язування складних задач управління з високою точністю при погано визначених умовах або станах нелінійних об'єктів керування опціонально достатнім є програмне середовище MATLAB зі спеціальними програмними пакетами

Fuzzy Logic Toolbox, Neural Networks Toolbox і Genetic Algorithm and Direct Search Toolbox [3–6].

Аналіз останніх досліджень та публікацій. Слід зазначити, що значний вклад у розвиток теорії і практики експертних оцінок внесли українські вчені: В. М. Глушков, Г. М. Гнатієнко, Н. З. Шор та ін. Саме бази знань, що сформовані експертами відповідно з особливостями предметної галузі, на нашу думку, забезпечують успіх у побудові ефективної нечіткої експертної системи.

Зокрема, дослідження застосування інтелектуальних технологій у практичних задачах одержали досить глибоке відображення в науковій літературі вітчизняних і зарубіжних учених: В. В. Борисова, Р. Ю. Голунов, М. І. Длі, В. В. Круглова, Ф. Уосермен, С. Д. Штовба; інтелектуальні інформаційні системи – А. І. Змитрович; теорії прийняття рішень – Г. А. Черноморова; прикладні нечіткі системи – К. Асаі, Т. Терано, М. Сугено; основні концепції нейронних мереж – Є. В. Бодянського, А. І. Галушкина, Р. Каллана, С. Осовського, О. Г. Руденко та ін.

Проте, дослідження, що поєднують застосування бази знань експертних груп під час оцінювання якості наукових публікацій, журналів, науковців і наукових організацій для побудови нечітких моделей управління прийняття рішень не проводилось.

Метою статті є розробка нечітких систем моделювання рейтингу наукових публікацій, журналів, науковців і наукових організацій.

Виклад основного матеріалу. Важливим застосуванням теорії нечітких множин є контролери нечіткої логіки, які використовуються в різноманітних системах управління прийняття рішень. Замість математичної моделі для опису системи такі контролери використовують інтегровані знання експертів, які за структурою подання наближаються до розмовної мови й описуються за допомогою лінгвістичних змінних і нечітких множин [7; 8].

Загальна структура fuzzy-контролера містить такі складові: блок фазифікації, база знань, блок рішень, блок дефазифікації.

Блок фазифікації перетворює чіткі величини, виміряні на виході об’єкта управління, на нечіткі величини, описані лінгвістичними змінними у базі знань.

Блок рішень використовує нечіткі умовні (if – then) правила, закладені у базі знань, для перетворення нечітких вхідних даних на необхідні керуючі впливи, що мають також нечіткий характер.

Блок дефазифікації перетворює нечіткі дані з виходу блоку рішень на чітку величину, яка подається на виконавчий пристрій для управління об’єктом.

Проілюструємо методику створення моделі на основі нечітких систем на прикладі. Нехай потрібно створити нечітку систему для експертної оцінки якості наукових публікацій. Відповідно до запропонованого в роботі [1] методу обчислення рейтингу наукових публікацій, журналів, науковців і наукових організацій якість публікації оцінюється за такими критеріями: експертна оцінка якості і кількість переглядів.

Критеріями оцінки експертів виступають вимоги ВАК України до фахових видань. Офіційно в документі «Про підвищення вимог до фахових видань, внесених до переліку ВАК України» зазначається обов’язкова наявність деяких елементів у публікації [2].

Кількість переглядів публікацій задається відповідно до умови, що вони знаходяться у вільному доступі й ураховується величина від 0 до 1000 переглядів. Вихід системи буде бальна оцінка якості статті в діапазоні від 0 до 1. Вихідні дані подані на рис. 1. Поверхня залежності якості статті від експертної оцінки й кількості переглядів наведено на рис. 2.

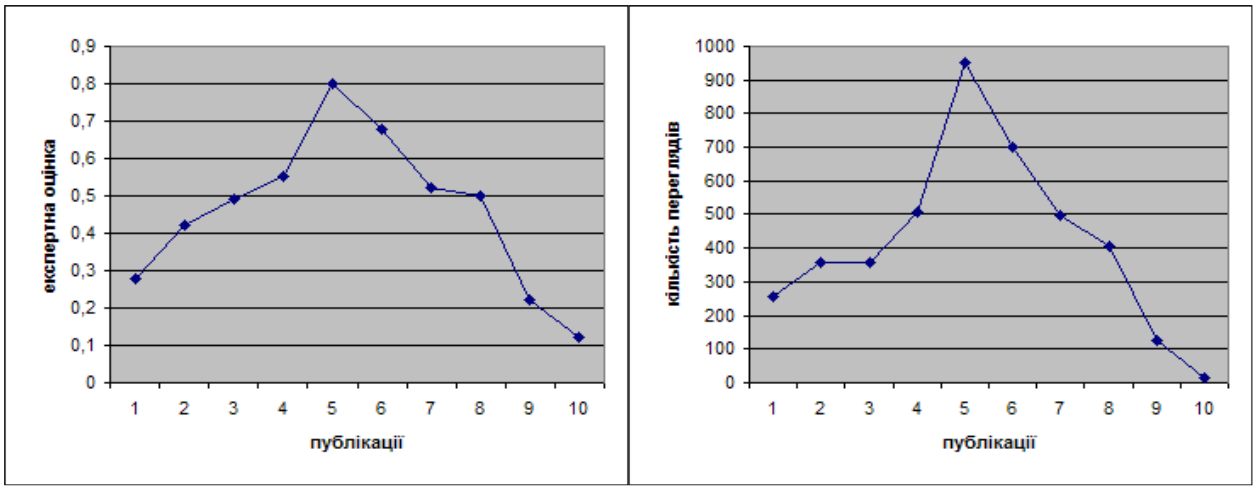

Рис. 1. Вихідні дані

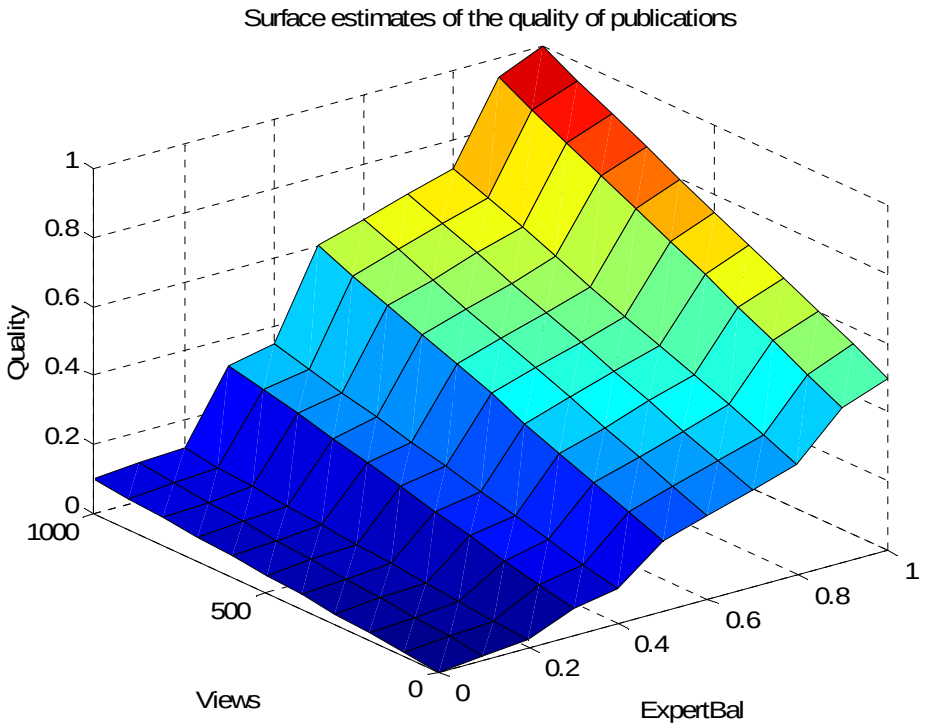

Рис. 2. Поверхня залежності якості статті від експертної оцінки і кількості переглядів

Названі вище критерії і вихід – якість публікації приймаються як лінгвістичні змінні. Структура нечіткої системи оцінки якості публікацій показано на рис. 3. Як видно з рис. 3 входами в нечітку систему є дві нечіткі лінгвістичні змінні «ExpertBal» – оцінка експертів, що задаємо трьома терм-множинами: «low», «midle», «high» і «Views» – кількість переглядів, що задаємо чотирма терм-множинами: «verylow», «low», «midle», «high». Виходом системи є змінна «Quality» – якість публікації, що задаємо трьома терм-множинами: «low», «midle» і «high».

Побудувати нечітку систему можна програмно або за допомогою редактора нечіткого висновку FIS [6].

Найбільш поширеними алгоритмами нечіткого логічного висновку є алгоритми Мамдані і Сугено. У середовищі MATLAB 7 за допомогою FIS-редактора створимо дані нечіткі системи.

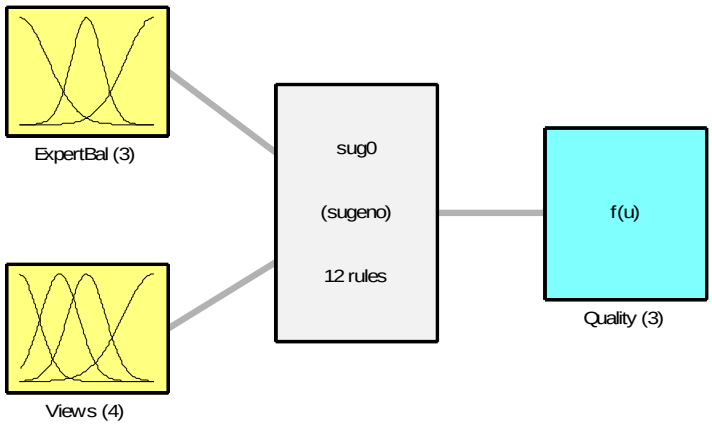

Рис. 3. Структура нечіткої системи оцінки якості публікацій

База нечітких знань і правил формується за типом: ЯКЩО… ТОДІ, і є сукупністю поєднаних нечітких значень вхідних лінгвістичних змінних, яке відповідає конкретній терм-множині вихідної лінгвістичної змінної. Формуємо таку базу правил для нечіткої системи Сугено:

П1: If (ExpertBal is low) and (Views is verylow) then (Quality is low) (1)

П2: If (ExpertBal is low) and (Views is low) then (Quality is low) (1)

П3: If (ExpertBal is low) and (Views is midle) then (Quality is low) (1)

…

…

П11: If (ExpertBal is high) and (Views is midle) then (Quality is high) (1)

П12: If (ExpertBal is high) and (Views is high) then (Quality is high) (1),

де в дужках після кожного правила записана його вага (для побудови експериментальної експертної системи вагу кожного правила задаємо 1).

За допомогою команди **View/Surface** FIS-редактора можна переглянути поверхню відгуку створеної системи (рис. 4) і поверхню похибок нечіткої системи (див. рис. 5).

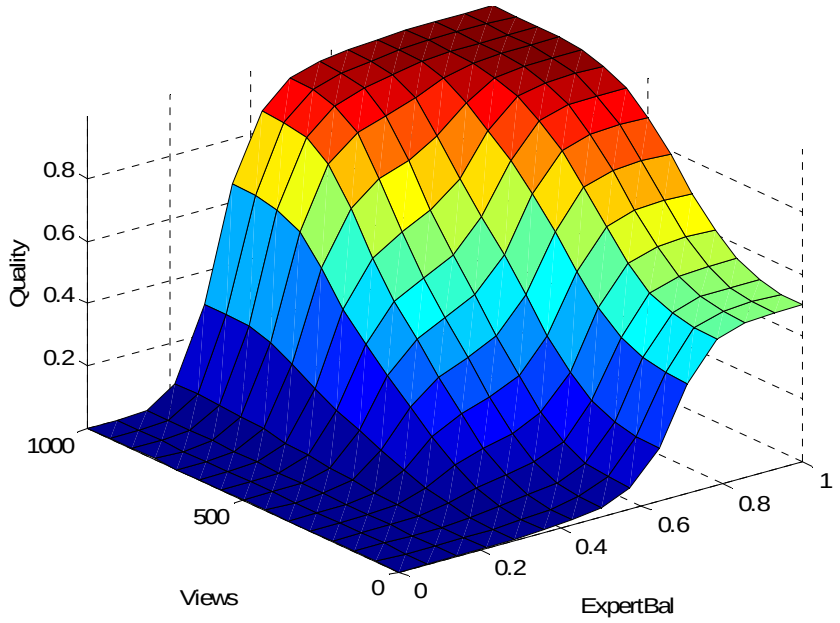

Рис. 4. Поверхня відгуку нечіткої системи Сугено нульового порядку

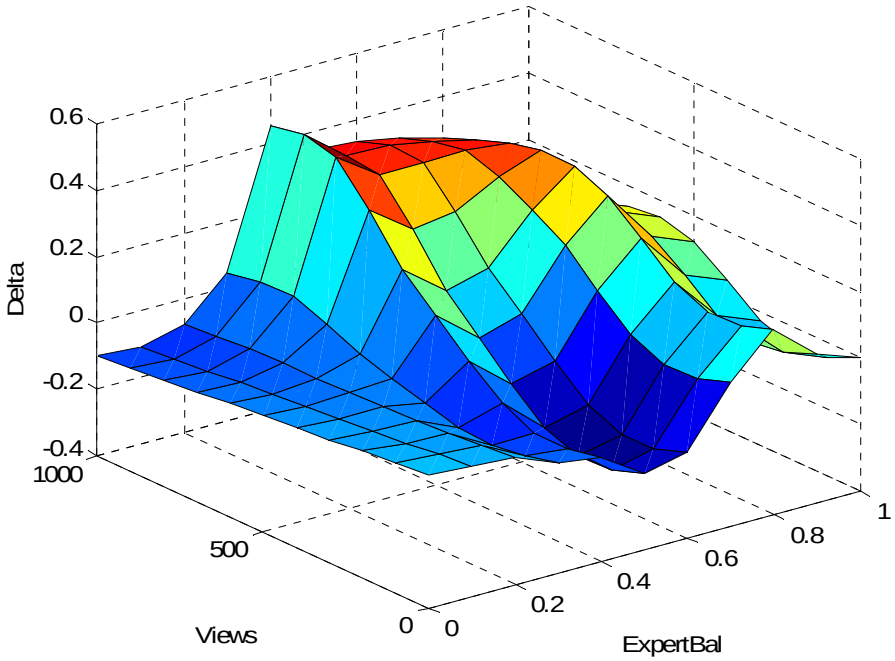

Рис. 5. Поверхня похибок нечіткої системи типу Сугено

Для навчання нечіткої системи Сугено необхідно дотримуватися двох умов:

1) кількість термів вихідної функції належності має бути рівною кількості складених правил;

2) тип функцій належності термів вихідної змінної має бути linear і їх параметри рівними 0.

Кількість, назви, тип і параметри термів вхідних функцій належності є аналогічними до попередніх систем. Після виконання програми ми одержали нову нечітку систему learn.fis (рис. 6) і поверхню неув'язок нової системи (рис. 7).

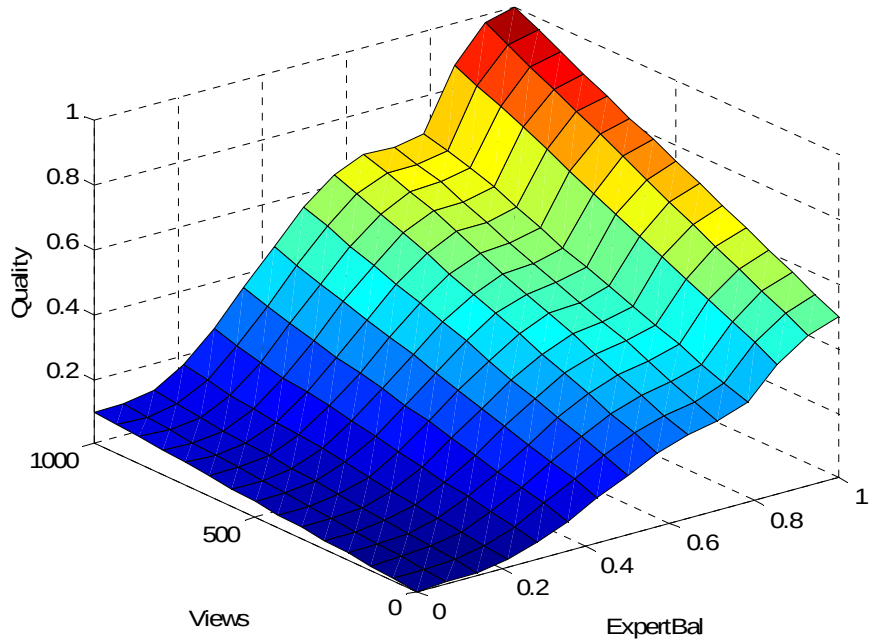

Рис. 6. Поверхня відгуку нечіткої системи типу Сугено, оптимізованої за допомогою алгоритму anfis

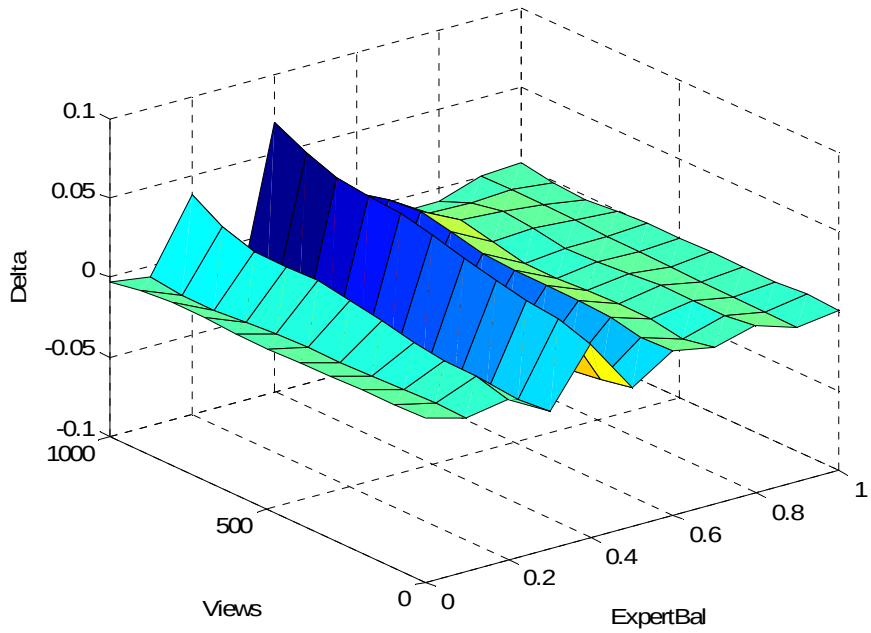

Рис. 7. Поверхня похибок нечіткої системи Сугено, оптимізованої за допомогою алгоритму anfis

Найменшу похибку під час відтворення заданої залежності має нечітка система типу Сугено, що навчена за допомогою алгоритму anfis (див. табл. 1). Саме на базі цієї системи складено нечіткий контролер в Simulink.

Таблиця 1

Похибки нечітких систем оцінки якості публікацій

Назва системи	Похибка
Нечітка система Сугено, що налаштована експертами	0,13265
Нечітка система Сугено, оптимізована за допомогою алгоритму *anfis*	0,013656
Нечітка система Мамдані, що налаштована експертами	0,12819
Нечітка система Мамдані після навчання за допомогою функції оптимізації *fmincon*	0,13686

Для побудови нечіткого контролера (рис. 8) використовуємо блок Fuzzy Logic Controller.

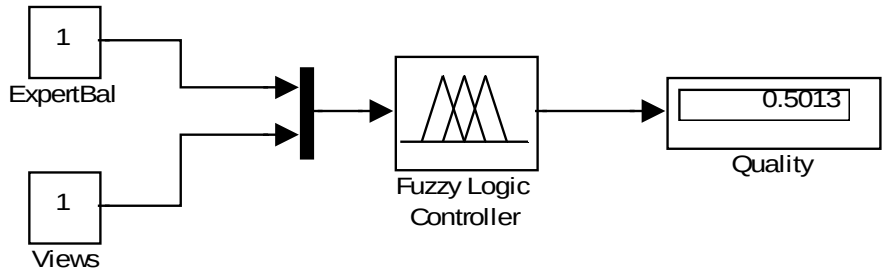

Рис. 8. Модель нечіткого контролера

У табл. 2 подані дані тестування контролера на тестовій вибірці, що наведена у вихідних даних.

Таблиця 2

Дані тестування нечіткого контролера

№ публікації	1	2	3	4	5	6	7	8	9	10
Експерт. оцінка	0,56	0,42	0,49	0,55	0,84	0,68	0,8	0,5	0,22	0,35
Кільк. переглядів	247	356	359	508	506	687	423	408	125	102
Якість публікації	0.322	0.245	0.31	0.417	0.606	0.557	0.49	0.337	0.047	0.118

Отже, побудована нечітка експертна система дозволяє провести оцінку якості публікацій, формалізація якої складна і неможлива в умовах неповної або неточної інформації. Більш точне врахування факторів, що впливають на якість публікацій, можна передбачити за рахунок додання вхідних параметрів моделі з відповідною зміною правил і їх вагового коефіцієнта залежно від важливості кожного з них у загальній базі правил.

Дослідження процесу створення і якості функціонування нечітких експертних систем показали високу їх ефективність, особливо під час реалізації алгоритму Сугено, завдяки відносно простим програмі навчання і практичній реалізації. Така оцінка якості може бути корисною для складання списку публікацій в архівах або на потужних Інтернет ресурсах. А також перспективою наступних досліджень є побудова автоматичної web-системи для оцінювання рейтингу публікацій, журналів, науковців і наукових організацій.

Список використаних джерел

1. Балагура І. В. Поєднання вебометричних методів та експертного оцінювання для побудови моделі рейтингу наукових публікацій на конференції «Інтелектуальний аналіз інформації» / І. В. Балагура // X Международная научная конференция имени Т. А. Таран «Интеллектуальный анализ информации ИАИ-2010», (Киев, 18–21 мая 2010 г.) сб. тр. / ред. кол.: гл. ред. С. В. Сирота. – К. : Просвіта, 2010. – 396 с.

2. Вища атестаційна комісія України. Вимоги до публікацій та фахових видань [Електронний ресурс]. – Режим доступу : www.vak.org.ua/protjournals1.php.

3. Корчемний М. О. Інтелектуальні технології управління та прийняття рішень. Ч.1. Розмита логіка / М. О. Корчемний, В. С. Федорейко. – Тернопіль : ТНПУ, 2007. – 140 с.

4. Круглов В. В. Нечеткая логика и искусственные нейронные сети / В. В. Круглов, М. Й. Дли, Р. Ю. Голунов. – М. : Издательство Физико-математической литературы, 2001. − 224 с.

5. Леоненков А. В. Нечеткое моделирование в среде MatLab і fuzzy TECH / А. В. Леоненков. – СПб. : БХВ – Петербург, 2003. – 736 с.

6. Штовба С. Д. Проектирование нечетких систем средствами MatLAB / С. Д. Штовба. – М. : Горячая линия. – Телеком, 2007. – 288 с.

7. Mamdani E. H. Application of fuzzy algorithms for the control of a simple dynamic plant / E. H. Mamdani // Proc. IEEE 121, 1974. – P. 1585–1588.

8. Sugeno M. Industrial applications of fuzzy control / M. Sugeno. R. Tong, R. R. Yager. – New York, John Wiley&Sons, 1995. – 359 p.

НЕЧЕТКОЕ МОДЕЛИРОВАНИЕ РЕЙТИНГА НАУЧНЫХ ПУБЛИКАЦИЙ

Цидыло Иван Николаевич, кандидат педагогических наук, доцент кафедры компьютерных технологий Тернопольского национального педагогического университета имени Владимира Гнатюка, г. Тернополь, e-mail: tsidylo@mail.ru.

Аннотация

Рассматривается проблема принятия решений в условиях неопределенности на основе применения нечетких множеств и нечеткой логики. Построена нечеткая экспертная система, что позволяет провести оценку качества публикаций, журналов, научных работников и научных организаций, формализация которой сложна и невозможна в условиях неполной или неточной информации. В процессе исследования показано, что нечеткие системы Сугено, настроенные с помощью алгоритмов обучения, имеют наивысшую точность благодаря относительно простым программе обучения и практической реализации. Описана структура и функции нечеткой экспертной системы для оценки качества научных публикаций. Проведено тестирование контролера на тестовой выборке, которая приведена в выходных данных.

Ключевые слова: нечеткая логика, экспертная оценка, нечеткая система Сугено, нечеткий контролер.

THE FUZZY MODELING OF SCIENTIFIC PUBLICATIONS RATING

Ivan M. Tsidylo, PhD (pedagogical sciences), associate professor of the Department of computer technologies of the Ternopil National Pedagogical University named after Volodymyr Gnatyuk, Ternopil, e-mail: tsidylo@mail.ru.

Resume

The problem of making decision is examined in the conditions of vagueness on the basis of application of fuzzy sets and fuzzy logic. The fuzzy consulting model that was built, allows to conduct the estimation of quality of publications, magazines, research workers and scientific organizations formalization of which is difficult and impossible in the conditions of incomplete or inexact information. It is shown in the process of research, that the fuzzy systems of Sugeno, adjusted by means of algorithms studies have the greatest exactness due to the simple program of studies and practical realization. A structure and functions of fuzzy estimation system for evaluation of quality of scientific publications. Controller testing was conducted on a test selection and all the results of the test are given below.

Keywords: fuzzy logic, expert estimation, Sugeno fuzzy system, fuzzy controller.

Матеріал надійшов до редакції 24.06.2011 р.