

УДК 658. 5+519

Сорокіна Л. В.,*кандидат економічних наук, доцент кафедри економіки підприємства
Київського міжнародного університету фінансів*

ПРОГНОЗУВАННЯ ГРОШОВИХ ПОТОКІВ ЗА ДОПОМОГОЮ УДОСКОНАЛЕНИХ НЕЧІТКИХ СИСТЕМ

У статті запропоновано методика удосконалення процедури проектування фаззи-алгоритмів для прогнозування руху грошових коштів. Сутність удосконалення полягає у поглибленому аналізі тренувальних даних на етапі побудови нечіткої бази знань на предмет частоти одночасної появи певних передумов та результатів.

Ключові слова: грошові потоки, операційні витрати, кластерний аналіз, нечіткий висновок, ваговий множник, частота спостережень.

В статті пропонується методика удосконалення процедури проектування фаззи-алгоритмів, призначених для прогнозування руху грошових коштів. Суть удосконалення полягає у поглибленому аналізі тренувальних даних на етапі побудови нечіткої бази знань на предмет частоти одночасного появи певних передумов та результатів.

Ключевые слова: денежные потоки, операционные затраты, нечеткий кластерный анализ, весовой множитель, частота наблюдений.

In the article is proposed the method of improvement of the procedure of the design of the fuzzy-algorithms, intended for predicting the motion of cash resources, is proposed. The essence of improvement consists of deep analysis of training given in the stage of the construction of illegible base knowledge to the object of frequency one temporary appearance of the specified conditions and results.

Key words: cash-flow, the operating expenditures, Fuzzy c-means clustering (FCM), weighting factor, the frequency of observations.

Постановка проблеми. Якість результатів, отримуваних внаслідок використання фаззи-алгоритмів та інших прийомів інтелектуального аналізу даних значною мірою визначається припущеннями, зробленими ще на етапі проектування технологій дейтамайнінгу. При цьому передпроцесування вхідного, тренувального масиву інформації має якнайточніше відображати відповідність вхідного набору даних меті та завданням інтелектуальної експертної системи. Істотне підвищення точності фаззи-алгоритмів стало можливим завдяки удосконаленню прийомів обґрунтування типу та параметрів функцій належності, проте процедура формування нечіткої бази знань у частині композиції передумов та висновків правил, агрегації нечітких продукцій

та дефазифікації потребує удосконалення. Актуальність окресленої проблеми підтверджується недостатнім поширенням його в практику фінансового та проектного менеджменту, яка супроводжується чималою невизначеністю станів внутрішнього та зовнішнього середовища підприємства.

Аналіз останніх досліджень та публікацій. Поглиблена статистична обробка масиву вхідних даних під час формування нечіткої бази знань ґрунтується на дослідженнях [3, 7]. Один із напрямів – нечітку кластеризацію як інструмент визначення параметрів функцій належності вхідних і вихідних змінних нами було успішно застосовано для обґрунтування надбавки до ставки дисконтування за внутрішньофірмові ризики [5] та для діагностики системи менеджменту будівельних підприємств [4]. В основу цих публікацій покладено ключові положення [7]. Стосовно проблеми удосконалення процесів проектування нечітких систем варто уваги дослідження [2], де запропоновано визначати функції належності вхідних і вихідних змінних за допомогою методу головних компонент.

Мета і завдання дослідження полягають у розробці та апробації методики удосконалення процедури проектування систем нечіткого висновку для прогнозування руху грошових коштів, яке спрямовано на підвищення точності значень вихідних змінних, і передусім стосується правильності визначення їх знаків.

Виклад основного матеріалу. У попередніх наших публікаціях [4, 6] було запропоновано визначати функції належності для всіх термів вхідних і вихідних змінних на підставі кластерного аналізу тренувальних даних, який має здійснюватися не одночасно за сукупністю критеріїв, а за кожним окремо. Його реалізація засобами програмного середовища MATLAB дозволяє не лише встановити типові значення вхідних і вихідних змінних, у яких функція належності спостереження певному терму наближається до одиниці, але і обчислює міру впевненості у належності кожного конкретного спостереження тому чи іншому кластеру. При цьому кластери формуються так, щоб міра впевненості щодо характеристик, властивих центру кластеру, не була нижчою за 0,5. Таким чином, завдяки використанню результатів нечіткої кластеризації зменшується трудомісткість та суб'єктивність початкового етапу формулювання правил нечіткої бази знань. Ми вважаємо, що такий підхід до формування алгоритму нечіткого висновку підвищить його гнучкість та забезпечить вищу точність результатів нечіткого висновку. На нашу думку, під час проектування нечітких систем перевагу варто віддавати гауссовим функціям належності, які за допомогою налаштувань програмного середовища MATLAB можуть застосовуватися і для так званих “крайніх” термів, для яких міру впевненості у належності певного спостереження графічно відображають S - або Z -подібні залежності. Техніку їх визначення у *fis-editori* MATLAB детально висвітлено в [1, 7].

Параметри функцій належності, в основу яких покладено гаусову криву, нами було запропоновано [4, 6] визначати як значення ознаки для центру кластеру та стандартне відхилення в межах певного кластеру.

Процес формування правил нечіткої бази знань, який дозволяє об'єктивно та із високим рівнем статистичної значимості сформулювати логічні імплікації та визначити їх ваговий множник, викладено в [3]. Його сутність полягає у поглибленому аналізі навчальної вибірки на предмет взаємної відповідності термів, яким із максимальною мірою упевненості належать вхідні змінні, із термами, до яких із максимальною мірою упевненості належить значення вихідної змінної. При цьому терми зазвичай визначаються шляхом поділу сукупностей значень вхідних та вихідних змінних на рівні інтервали, кількість яких суб'єктивно визначається дослідником, а функції належності обираються так, щоб у центрі інтервалу їх значення дорівнювало одиниці, але при підстановці абсциси центру інтервалу до суміжних функцій належності їх значення дорівнювали нулю. Потім тренувальні дані ставлять у відповідність тим термам, у яких вони досягають максимальних значень функції належності, а далі шляхом зіставлення лінгвістичних термів вхідних і вихідних змінних формується правило. Міра істинності кожного правила визначається як добуток значень функцій належності всіх його передумов. Серед правил, що мають однакові передумови, але різні висновки визначають такі, у яких міра істинності виявиться найбільшою. У такий спосіб розв'язується проблема усунення суперечливих правил та скорочення їх кількості у базі знань. Подальше удосконалення цієї процедури полягає у вивченні "типових" залежностей "термів" вихідних змінних від термів входів. Припускаючи відповідність випадкових змін внутрішнього та зовнішнього економічного середовища підприємства нормальному закону розподілу, ми вважаємо за доцільне проаналізувати відносну частоту належності вхідних змінних тому чи іншому терму в розрізі окремих груп спостережень, виділених на підставі нечітких кластерів-термів вихідної змінної. Інакше кажучи, для формування нечіткого правила та визначення його вагового коефіцієнта, потрібно проаналізувати лінгвістичну структуру передумов у межах кожного терму вихідної змінної. При цьому об'єднання передумов по окремих входах в єдине правило варто виконувати із використанням логічної зв'язки "ТА". Як відомо із теорії нечітких множин, подібна функція кількох змінних називається **T-нормою** [3], реалізація якої полягає в обчисленні мінімуму або добутку значень функцій належності вхідних передумов. Отже, для розрахунку міри істинності правила викладений вище спосіб формування правил нечіткої бази знань передбачає встановлення істинності кожного правила за допомогою T-норми у формі добутку. Запропоноване нами удосконалення процедури проектування алгоритму нечіткого висновку полягає у застосуванні T-норми у формі мінімуму. Міра впевненості у тому, що вихідна змінна належить певному терму, спочатку визначається за максимальними частотами термів входів, а міра істинності самого правила встановлюється за мінімальною частотою серед максимумів частот передумов. Іншими словами, вага кожного правила визначається

ся за “МІНІМАКСНОЮ” частотою термів його передумов, яка визнається за сукупністю навчальних (тренувальних) даних.

Формалізуючи виконані у вищезазначений спосіб процедури, пропонуване удосконалення прийомів формування бази знань має бути врахованим ваговим коефіцієнтом правил (α^*). Його величина, на нашу думку, повинна обчислюватися за формулою:

де $\max_{i=1, \dots, n} \alpha_{ij} = \beta_j$ – максимальна частота передумов, що призвели до терму вихідної змінної, для якого складається правило;

$\min_{j=1, \dots, m} \beta_j$ – мінімальна частота з m спостережень, для яких справедливе досягнення терму висновку, стосовно якого складається правило;

i – порядковий номер терму вхідної змінної;

n – кількість термів вхідної змінної;

j – порядковий номер передумови у правилі нечіткого висновку;

m – кількість передумов у правилі.

Деяка подібність формули (1) до рекомендацій теорії ігор щодо зменшення ризику під час прийняття управлінських рішень, на нашу думку, має усунути чи принаймні зменшити кількість не виправданих оптимістичних прогнозів, зокрема випадків, коли всупереч додатним результатам прогнозів, фактичні результати виявляються від’ємними.

Ефективність запропонованого удосконалення можна встановити після його застосування, порівнюючи отримані результати із фактичними даними підприємств. Для проектування фаззі-алгоритму за допомогою вищеподаної методики удосконалення було використано ретроспективні дані річної фінансової звітності 20 будівельних підприємств, організаційно-правова форма яких – відкрите акціонерне товариство. 11 із об’єктів дослідження присутні в переліках “Української біржі” та ПФТС, решту ж можна охарактеризувати як “не публічні компанії”, акції яких не можуть бути предметом купівлі-продажу на фондовому ринку. Інформація фінансових звітів за 2004 – 2008 роки склала тестову вибірку, а дані кризового 2009 року – контрольну.

Для формулювання правил нечіткого висновку нами було проаналізовано структуру передумов у межах кожного терму вихідної змінної, тобто визначалась питома вага випадків із високим, середнім або низьким проявом вхідної ознаки. Результати такого аналізу систематизовано у вигляді діаграм на рис. 1-2. Кількість значень питомої ваги секторів на діаграмах відповідають величині α_{ij} з формули (1).

Як показано на діаграмах на рис. 1., типові передумови для від’ємного грошового потоку можна сформулювати, по-перше, для “типового” варіанту структури передумов, тобто таких термів вхідних змінних, які мають місце найчастіше і мають найбільшу питому вагу. Пропонуємо вагу такого правила визначати за максимальним значенням частоти прояву за

сукупністю всіх ознак-передумов (β , з формули (1)). Цей вибір можна пояснити метою дослідження – зменшити кількість випадків помилкового прогнозування позитивного грошового потоку. Саме з цих міркувань пропонуємо сформулювати ще одне правило – для тих випадків, передумови яких відрізняються від типових, побудованих згідно з так званим “Міні-максним” підходом за формулою (1).

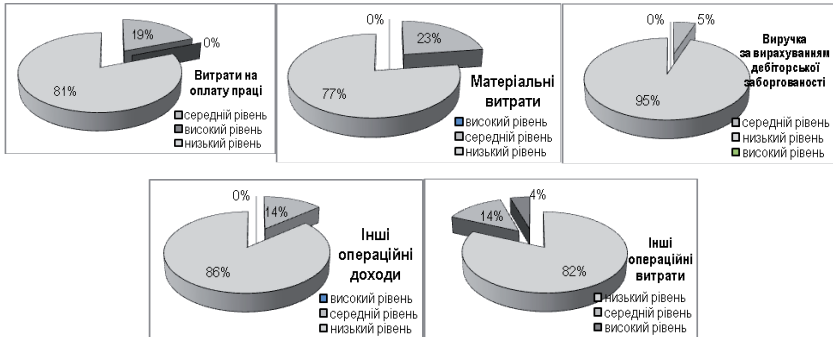


Рис. 1. Частота розподілу термів фактичних значень вхідних змінних, у випадках від’ємного грошового потоку будівельних підприємств

У такому правилі для кожної з вхідних змінних застосовано логічну зв’язку “НЕ”, а вага правила визначається як різниця між одиницею та

вагою правила для “типових” передумов: $\bar{\alpha}^* = 1 - \alpha^*$, (1)

Тобто до нечіткої бази знань включаються 2 взаємодоповнюючих правила, ваги яких доповнюють одна одну до одиниці (α^* та $\bar{\alpha}^*$ відповідно до формул (1) та (2)). Між собою передумови в обох правилах пов’язуються логічною зв’язкою “ТА”. Наприклад, для пропонованого алгоритму до бази знань вміщено такі правила із ваговими коефіцієнтами:

1. Якщо “витрати на оплату праці” “низькі” та “матеріальні витрати” “низькі”, та “виручка за вирахуванням дебіторської заборгованості” “низька”, та “інші операційні доходи” “низькі”, та “інші операційні витрати” “низькі”, грошовий потік від операційної діяльності “від’ємний”. Вага цього правила $\alpha_1^* = 0,7$ визначається відповідно до мінімального показника відсотка, встановленого порівнянням максимальних частот прояву термів окремих змінних – мінімальне значення з максимальних за розміром секторів діаграм на рис. 1.

2. Якщо “витрати на оплату праці” “НЕ низькі” та “матеріальні витрати” “НЕ низькі”, та “виручка за вирахуванням дебіторської заборгованості” “НЕ низька”, та “інші операційні доходи” “НЕ низькі”, та “інші операційні витрати” “НЕ низькі”, “грошовий потік від операційної діяльності” “від’ємний”. Вага цього правила $\bar{\alpha}_2^* = 0,23 = 1 - 0,77$.

Третє правило в базі знань сформовано на підставі других за частотою

прояву термів вхідних змінних – тобто другі за розміром сектори на діаграмах (рис. 1.). Логічна зв'язка між термами передумов також “ТА”, а численні експерименти на тестових та контрольних даних показали, що вагу правила варто обирати в межах від середнього арифметичного до максимальної величини частоти прояву передумов, які визначаються розміром другого за розміром сектору на кожній з діаграм (рис. 1.). Таким чином, третє правило у нечіткій базі знань має таке формулювання.

3. Якщо “витрати на оплату праці” “середні” та “матеріальні витрати” “середні”, та “виручка за вирахуванням дебіторської заборгованості” “середня”, та “інші операційні доходи” “середні”, та “інші операційні витрати” “середні”, “грошовий потік від операційної діяльності” “від’ємний”. Вага цього правила $\alpha_3^* = 0,23$ визначається відповідно до другого за розміром показника відсотка, встановленого порівнянням максимальних частот прояву термів окремих змінних (мінімальне значення з середніх за розміром секторів діаграм на рис. 1.).

Аналогічно було сформульовано три інші правила у базі знань та визначено їх вагові коефіцієнти:

4. Якщо “витрати на оплату праці” “низькі” та “матеріальні витрати” “низькі”, та “виручка за вирахуванням дебіторської заборгованості” “низька”, та “інші операційні доходи” “низькі”, та “інші операційні витрати” “низькі”, “грошовий потік від операційної діяльності” “низький”. Вага цього правила $\alpha_4^* = 1$.

5. Якщо “витрати на оплату праці” “середні” та “матеріальні витрати” “середні”, та “виручка за вирахуванням дебіторської заборгованості” “низька”, та “інші операційні доходи” “низькі”, та “інші операційні витрати” “низькі”, “грошовий потік від операційної діяльності” “середній”. Вага цього правила $\alpha_5^* = 0,46$ – визначається відповідно до другого за розміром показника відсотка, встановленого порівнянням максимальних частот прояву термів окремих змінних (мінімальне значення з найбільших за розміром секторів діаграм на рис. 2.).

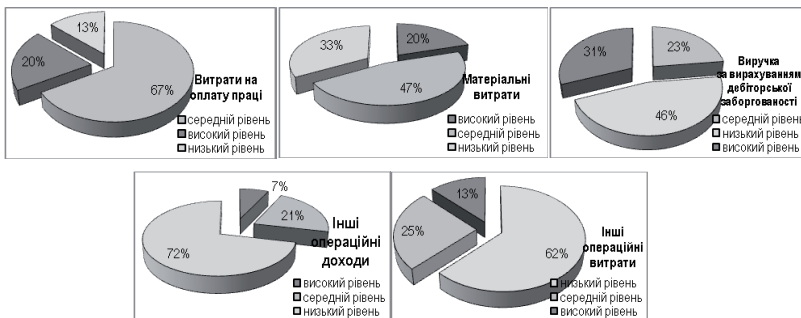


Рис. 2. Частота розподілу термів фактичних значень вхідних змінних, у випадках середнього за обсягом додатного грошового потоку будівельних підприємств

6. Якщо “витрати на оплату праці” “високі” та “матеріальні витрати” “високі”, та “виручка за вирахуванням дебіторської заборгованості” “висока”, та “інші операційні доходи” “низькі”, та “інші операційні витрати” “низькі”, “грошовий потік від операційної діяльності” “високий”. Вага цього правила $\alpha_6^* = 1$.

Завдяки об’єднанню в одне правило кількох, а інколи і всіх вхідних змінних, кількість правил у базі знань вдалося обмежити 6-ма. Це значно менше за максимально можливу $1024=(4(\text{терми вихідної змінної})^5(\text{кількість вхідних змінних}))$, і навіть менше на запропоновану у роботі [7, с. 48] мінімальну кількість правил, визначену як 3% від максимально можливої їх кількості, (тобто 31 для умов виконаного дослідження). Проте запропонована система дозволяє отримати результати прийнятної якості. У табл. 1. зведено показники грошового потоку від операційної діяльності, досягнуті аналізованими підприємствами у 2009 році, а також розраховані за допомогою fuzzy-алгоритму. Хоча розрахункові значення дещо відхиляються від фактичного результату, проте кількість випадків помилки у визначенні знаку результативного показника в результаті застосування fuzzy-алгоритму не перевищує 20%.

Таблиця 1
Результати прогнозування грошового потоку від операційної діяльності, отримані за допомогою удосконаленого Fuzzy-алгоритму, тис. грн.

Підприємство	Обсяг грошового потоку від операційної діяльності	
	фактично досягнутий у 2009 році	розрахований за допомогою удосконаленого fuzzy-алгоритму
Підприємства, не включені до переліку ПФТС та “Української біржі”		
ВАТ Дніпроводбуд	-133,0	-1510,0
ВАТ Луцьксантехмонтаж	1394,0	270,0
ВАТ Херсонводбуд	-87,0	281,0
ВАТ Чернігівводбуд	-2657,0	-869,0
ВАТ Укрводбуд	-1288,0	-876,0
ВАТ “Закарпатсько-галицьке ШБУ”	-117,0	-443,0
ВАТ Ковельське ШБУ 63	1359,0	291,0
ВАТ Хмельницьке ШБУ 56	-1418,0	17,1
ВАТ Львівське ШРБУ 65	1765,0	347,0
Підприємства, включені до переліку ПФТС та “Української біржі”		
ВАТ Будмеханізація	-32989,0	-1050,0
ВАТ ДБК-3	8259,0	4260,0
ВАТ ДБК-4	18715,0	27700,0
ВАТ Київелектромонтаж	-2932,0	-5000,0
ВАТ Київметробуд	13143,0	-64500,0

ВАТ Трест Київспецбуд	1498,0	-1480,0
ПАТ Мостобуд	-31520,0	30000,0
ВАТ Південьзахід-електромережбуд	5240,0	22800,0
ВАТ Трест-Житлобуд-1	-7235,0	-13900,0
ВАТ “ШБУ-41”	3756,0	7420,0
ПАТ ВАТ трест “Київміськбуд-1”	-11667,0	-28000,0

Висновки. Удосконалення процедури передпроцесування тренувальних даних значно поліпшує якість фаззи-алгоритмів та їх практичну значимість для розв’язання задач фінансового прогнозування в умовах невизначеності. Проте запропонована у статті методика є досить трудомісткою та потребує чималої вибірки тренувальних даних, яка має включати фінансові результати інших підприємств, подібних до суб’єкта управління. Зазначена обставина є спонукальним мотивом для подальших досліджень обраного напрямку, серед яких особливе місце посідає проблема автоматизації статистичної обробки та формування нечіткої бази знань програмними засобами.

Література:

1. Алгоритмы и системы нечеткого вывода при решении задач диагностики городских инженерных коммуникаций в бреде MATLAB/ Л. А. Демидова, В. В. Кираковский, А. Н. Пылькин. – М.: Радио и связь, Горячая линия – Телеком, 2005. – 365 с.: ил.
2. Вітлінський В., Піскунова О., Рибальченко Л. Дослідження діяльності страхових компаній України методами факторного аналізу та теорії нечітких множин // Економіка України – 2009. – № 5. – С. 46-58.
3. Рутковская Д., Пилинський М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы / Пер. с польск. И. Д. Рудинского. – М.: Горячая линия – Телеком, 2007. – 452 с.: ил.
4. Сорокіна Л. В. застосування інструментарію теорії нечітких множин для діагностики функціонування системи менеджменту будівельних підприємств // Фінансова система України: Збірник наукових праць. – Острого: Видавництво Національного університету “Острозька академія”, 2010. – Випуск 14. – С. 471– 481.
5. Сорокіна Л. В. Застосування теорії нечітких множин для оцінки вартості будівельного підприємства // Актуальні проблеми економіки. – 2010. – № 8 (110) – С. 145 – 154.
6. Сорокіна Л. В. Застосування алгоритму нечіткого висновку типу Сугено для управління фінансовими потоками будівельних підприємств // Актуальні проблеми економіки. – 2011. – № 1 (115).
7. Штовба С. Д. Проектирование нечетких систем средствами MATLAB. – М.: Горячая линия – Телеком, 2007. – 288 с., ил.