

ЗАСТОСУВАННЯ НЕЧІТКОЇ КЛАСТЕРИЗАЦІЇ ДЛЯ УДОСКОНАЛЕННЯ КОНТРОЛЮ В ПОДАТКОВІЙ СЛУЖБІ

УДК 004.94: 336.2

ЗАБОЛОТНИКОВА Вікторія Сергіївна

асистент кафедри «Інформаційні системи в економіці» автомобільно-дорожнього інституту Донецького національного технічного університету.

Наукові інтереси: інформаційні технології.

e-mail: zabolotnikova@ukr.net

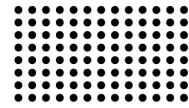
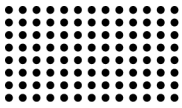
ВСТУП І ПОСТАНОВКА ЗАДАЧІ

Розвиток економіки України, її конкурентоспроможність і інтеграція в європейське співтовариство залежить від ефективності економічної політики держави, складовою частиною якої є податкова політика. Стратегічна мета податкової політики – створення в Україні стабільної податкової системи, яка б забезпечила достатній обсяг платежів у бюджеті всіх рівнів, ефективне функціонування економіки держави, слушний підхід до оподаткування всіх категорій платників податків. Становлення ринкових принципів господарювання, виникнення нових податкових відносин, необхідність зміцнення й стабілізації фінансової системи країни, забезпечення стійкого надходження бюджетних доходів об'єктивно обумовили зростання актуальності контролю в сфері оподаткування. У чинність цих об'єктивних причин відбулося відокремлення такого напрямку державного фінансового контролю, як контроль над своєчасною й повною сплатою платниками податків належних платежів у бюджет. Якісний та ефективний податковий контроль, виконання в повному обсягу всіх відповідних функцій повинен створювати позитивну динаміку податкових надходжень. На сьогоднішній день досягнуті значні результати в розвитку податкового контролю, які проводяться у рамках модернізації податкової служби. Однак загальний рівень податкової дисципліни, обсяг недоотриманих податків і зборів в бюджетну систему України, невідповідність положень податкового та

облікового законодавства, наслідком чого є збільшення числа і ускладнення схем ухилення від сплати податків, застосовуваних несумлінними платниками податків, змушують здійснювати пошук нових резервів підвищення результативності податкового контролю.

Для ефективного здійснення процесу контролю в податковій службі існує необхідність розподілу платників податків за категоріями уваги. На підставі такої класифікації податкові органи повинні вибирати методи і принципи спілкування з платниками податків, залежно від того, до якої категорії вони відносяться. Так, при роботі з сумлінними платниками податків необхідно звертати увагу в першу чергу на профілактику порушень податкового законодавства, то по відношенню до платників податків помірному ризику буде застосовуватися постійне спостереження, а також кабінетний аудит (камеральні перевірки). При роботі з платниками податків високого ризику доцільним буде здійснення виїзного аудиту разом з постійним спостереженням (виїзні перевірки) [3].

Виїзні перевірки є найбільш ефективною формою податкового контролю, в ході яких можуть бути виявлені правопорушення. Проте ця форма є дуже трудомісткою і забирає багато робочого часу працівників податкових органів. Тому здійснення даної форми контролю доцільно тільки тоді, коли витрати на її проведення значно вище отриманих результатів.



Враховуючи важливість цього аспекту, виникає необхідність у розробці ефективної математичної моделі розподілу об'єктів (платників податків) податкового контролю за категоріями уваги системі.

АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ

Проблемі контролю за виконанням податкового законодавства пильну увагу приділили такі вітчизняні вчені як І.А. Аврова, В.І. Братиків, Т.П. Брисіна, Є.В. Орлова. Питаннями податкового контролю в системі фінансових відносин держави і підприємств займалися А.В. Бризгалін, С.В. Котов, І.А. Орешкін, І.Р. Пайзулаєв. Тим не менш, розроблені методи побудови моделей не призводять до задовільним результатам у разі, коли вихідний опис, який підлягає вирішенню проблеми, свідомо є неточним або неповним.

Мета роботи – застосувати метод нечіткої кластеризації для удосконалення контролю в податковій службі. Для цього необхідно розподілити платників податків за категоріями уваги.

ОСНОВНИЙ МАТЕРІАЛ

Одним з найбільш цікавих і багатообіцяючих підходів до аналізу багатовимірних явищ і процесів, є методи кластерного аналізу [1]. Виходячи з цього, для розподілу платників податків за категоріями уваги буде застосований метод нечіткої кластеризації даних. Питаннями розробки та застосування методів нечіткої кластеризації займалися такі вчені як А.С. Тараскіна, яка займалася нечіткою кластеризацією за модифікованою методом середніх, С.Д. Штовба, який вивчав проектування систем керування, А.В. Леоненков акцентував увагу на нечіткє моделювання в середовищі MATLAB, Н.В. Куркіна досліджувала завдання динамічного програмування та нечіткої кластеризації, а так само О.М. Борисов, О.А. Крумберг, М.П. Деменков, Д.А. Вятченін. Серед зарубіжних вчених необхідно відзначити наступних вчених: Bezdek J., завдяки яким отримала розвиток і широке застосування нечітка кластеризація і його метод середніх, а також Л. Заде, Е. Распин, Дж. Дан, Д.Є. Густафсон, В.С. Кессель, Г. Ганн та інші.

Необхідність забезпечення наповнення дохідної частини бюджету ставить завдання щодо удосконалення системи адміністрування податків, що передбачає створення умов для постійного неупередженого конт-

ролю за своєчасністю та повнотою сплати суб'єктами господарювання податків і зборів до бюджетів та державних цільових фондів.

Для забезпечення виконання зазначеного та раціонального використання інтелектуальних, матеріальних та інформаційних ресурсів органів державної податкової служби України, забезпечення своєчасної та повної сплати податків і зборів (обов'язкових платежів) до бюджетів та державних цільових фондів є потреба у розподілі суб'єктів господарювання за категоріями уваги на основі факторів податкових ризиків [3].

Категорія уваги – характеристика суб'єктів господарювання відповідного рівня ризику, сформована на основі критеріїв оцінки та розподілу. Розрізняють такі категорії уваги:

- категорія 1 (зелений коридор) – сумлінні платники податків;
- категорія 2 (жовтий коридор) – платники податків помірному ризику;
- категорія 3 (червоний коридор) – платники податків високого ризику.

У цей час задачі обробки інформації, заданої в нечисловому виді, одержали широке поширення. Кластеризація даних відноситься до одних з основних методів інтелектуальної обробки інформації [4]. Аналіз і моделювання складних формалізованих процесів, які характеризуються великим числом враховуваних факторів, вимагає застосування спеціалізованих методів і інструментальних засобів. Використання кластеризації виправдане скрізь, де потрібен багатомірний аналіз різномірної інформації [5].

Кластеризація – це поділ елементів деякої множини на групи на основі їх схожості. Завдання кластеризації полягає в розбивці об'єктів з X на кілька підмножин (кластерів), у яких об'єкти більш схожі між собою, чим з об'єктами з інших кластерів. У метричному просторі «схожість» звичайно визначають через відстань [6].

Існує багато методів кластеризації, які можна класифікувати на чіткі й нечіткі [6].

Чіткі методи кластеризації розбивають вихідну множину об'єктів X на кілька непересічних підмножин. При цьому будь-який об'єкт із X належить тільки одному кластеру.

Нечіткі методи кластеризації дозволяють одному й тому самому об'єкту належати одночасно декільком

(або навіть усім) кластерам, але з різним ступенем належності. Нечітка кластеризація в багатьох ситуаціях більш «природня», чим чітка, наприклад, для об'єктів, розташованих на границі кластерів [6].

«Вимога знаходження однозначної кластеризації елементів досліджуваної проблемної області є досить грубим і твердим, особливо при рішенні погано або слабко структурованих завдань системного аналізу. Методи нечіткої кластеризації послаблюють цю вимогу. Ослаблення вимоги здійснюється за рахунок уведення в розгляд нечітких кластерів і відповідних їм функцій належності, що приймають значення з інтервалу $[0, 1]$ » [7, с.381]. А елементи матриці ступенів належності чіткої розбивки приймають значення із двоелементної множини $\{0,1\}$, а не з інтервалу $[0,1]$. У зв'язку з вищесказаним, стає очевидним застосування методу нечіткої кластеризації для розподілу платників податків за категоріями уваги.

Розвиток і широке застосування нечітка кластеризація одержала завдяки Бездеку та його методу нечітких c -середніх, але у базовому алгоритмі нечітких c -середніх відстань між об'єктом і центром кластера розраховується через стандартну Евклідову норму. В результаті алгоритмів кластеризації з фіксованою нормою форма всіх кластерів виходить однакова. Алгоритми кластеризації як би нав'язують даним невластиву їм структуру, що інколи приводить до неоптимальних результатів [6]. Для усунення цього недоліку існує кілька методів, серед яких виділимо алгоритм Густафсона-Кесселя.

Алгоритм Густафсона-Кесселя використовує адаптивну норму для кожного кластера, тобто для кожного i -го кластера існує своя норм-породжуюча матриця A_i . У цьому алгоритмі при кластеризації оптимізуються не тільки координати центрів кластерів та матриця нечіткої розбивки, але також і норм-породжуючі матриці для всіх кластерів. Це дозволяє виділяти кластери різної геометричної форми [8].

Досліджувана сукупність даних являє собою кінцеву множину елементів $A = \{a_1, a_2, \dots, a_n\}$, яке одержало назву множина об'єктів кластеризації. У розгляд також вводиться кінцева множина ознак або атрибутів $P = \{p_1, p_2, \dots, p_q\}$, кожний з яких кількісно представляє деяку властивість або характеристику елементів розглянутої проблемної області. При цьому натуральне n визначає загальну кількість об'єктів даних, а натураль-

не q – загальну кількість вимірних ознак об'єктів. Далі передбачається, що для кожного з об'єктів кластеризації деяким чином обмірювані всі ознаки множині P у деякій кількісній шкалі. Тим самим кожному з елементів $a_i \in A$ поставлений у відповідність деякий вектор $x_i = (x_1^i, x_2^i, \dots, x_q^i)$, де x_j^i – кількісне значення ознаки $p_j \in P$ для об'єкта даних $a_i \in A$.

Вектори значень ознак $x_i = (x_1^i, x_2^i, \dots, x_q^i)$ зручно представити у вигляді так названої матриці даних X розмірності $(n \times q)$, кожний рядок якої рівняється значенню вектора x_i .

Кластерна структура задається матрицею належності μ розмірності $c \times n$, де $\mu_{A_k}(a_i)$ – ступінь належності a_i -го елемента k -му кластеру.

Матриця належності повинна задовольняти наступним умовам:

$$a) \mu_{A_k}(a_i) \in [0,1], k = \overline{1, c}, i = \overline{1, n};$$

$$b) \sum_{k=1}^c \mu_{A_k}(a_i) = 1, \text{ тобто кожний об'єкт повинен бути розподілений між усіма кластерами};$$

не бути розподілений між усіма кластерами;

$$b) 0 < \sum_{i=1}^n \mu_{A_k}(a_i) < 1, \text{ тобто жоден кластер не повинен бути порожнім або містити всі елементи.}$$

Далі для кожного нечіткого кластера вводяться в розгляд так звані типові представники або центри v_k шуканих нечітких кластерів, які розраховуються для кожного з нечітких кластерів і по кожному з ознак за наступною формулою:

Далі для кожного нечіткого кластера вводяться в розгляд так звані типові представники або центри v_k шуканих нечітких кластерів, які розраховуються для кожного з нечітких кластерів і по кожному з ознак за наступною формулою:

$$v_j^k = \frac{\sum_{i=1}^n (\mu_{A_k}(a_i))^m \cdot x_j^i}{\sum_{i=1}^n (\mu_{A_k}(a_i))^m}, j = \overline{1, q}, \quad (1)$$

де $m \in (1, \infty)$ – експонентна вага, що визначає нечіткість, розмитість кластерів.

Чим більше значення експонентної ваги, тим матриця належності більш розмазана й при $m \rightarrow \infty$ елементи приймуть вид $\mu_{A_k}(a_i) = \frac{1}{c}$, що є поганим рішенням, тому що всі об'єкти будуть з однаковим ступенем розподілені по всіх кластерах. Експонентна вага дозволяє при формуванні координат центрів кластерів підсилити вплив об'єктів з більшими ступенями належності й

зменшити вплив об'єктів з малими ступенями належності. Теоретично обґрунтованого правила вибору ваги поки не існує, і зазвичай встановлюють $m = 2$ [9].

Для оцінки якості розбивки використовується критерій розкиду, що показує суму відстаней від об'єктів до центрів кластерів з відповідними ступенями належності:

$$f(A_k, v_j^k) = \sum_{i=1}^n \sum_{k=1}^c (\mu_{A_k}(a_i)) \sum_{j=1}^m (x_j^i - x_j^k)^2 \quad (2)$$

Задача нечіткої кластеризації щодо розподілу платників податків може бути сформульована в такий спосіб: розподілити платників податків за трьома категоріями уваги задавши функцію належності $\mu_{A_k}(a_i)$, і-рядок якої представляє інформацію про одного платника податків, а k-стовпець надає інформацію про всіх платників податків, які належать більшою чи меншою мірою до конкретного кластера, яка доставляє мінімум цільової функції і задовольняють вище перерахованим обмеженням.

Алгоритм Густафсона-Кесселя складається з наступних кроків.

Крок 1. Генеруємо матрицю нечіткої розбивки (табл. 1).

Таблиця 1 –

Матриця належності експертного оцінювання

№ п/п	Платник податків	Категорії уваги		
		Зелені	Жовті	Червоні
1	Коломійцева Галина Юріївна	0,0500	0,9000	0,0500
2	Мартиненко Олена Анатоліївна	0,9000	0,0500	0,0500
3	Юрченко Людмила Михайлівна	0,1000	0,2000	0,7000
4	Коноводов Юрій Миколайович	0,5500	0,3500	0,1000
97	Коваленко Валентина Василівна	0,5000	0,3000	0,2000
98	Глушко Лілія Володимирівна	0,7500	0,2000	0,0500
99	Кульбачна Тетяна Валентинівна	0,0500	0,0500	0,9000
100	Давідайтис Надія Миколаївна	0,1000	0,2000	0,7000

Крок 2. Розраховуємо центри кластерів за формулою (табл. 2).

Таблиця 2 –

Центри кластерів

Категорії уваги	Фактори							
	1	2	3	4	5	6	7	8
Зелені	0,23564	0,24754	0,26316	0,25478	0,2679	0,21748	0,28221	0,20789
Жовті	0,78414	0,76387	0,76809	0,80704	0,81388	0,80441	0,79142	0,77872
Червоні	0,52152	0,50933	0,50497	0,486	0,5043	0,47473	0,45326	0,44784

Крок 3. Визначаємо матриці коваріації для кожного кластера:

$$B_j^k = \frac{\sum_{i=1}^n (\mu_{A_k}(a_i))^m \cdot (x_j^i - v_j^k)^T \cdot (x_j^i - v_j^k)}{\sum_{i=1}^n (\mu_{A_k}(a_i))^m} \quad (3)$$

Матриця коваріації для третього кластеру, тобто для червоної категорії уваги має наступний вигляд:

$$B_j^k = \begin{pmatrix} 0,4924 & 0,4542 & 0,4808 & 0,4493 & 0,4152 & 0,4393 & 0,4544 & 0,4743 \\ 0,4542 & 0,4891 & 0,5079 & 0,4551 & 0,3935 & 0,4161 & 0,5092 & 0,5344 \\ 0,4808 & 0,5079 & 0,5972 & 0,4911 & 0,4311 & 0,4549 & 0,5449 & 0,5695 \\ 0,4493 & 0,4551 & 0,4911 & 0,5340 & 0,4769 & 0,5210 & 0,5874 & 0,6106 \\ 0,4152 & 0,3935 & 0,4311 & 0,4769 & 0,4927 & 0,5266 & 0,5404 & 0,5280 \\ 0,4393 & 0,4161 & 0,4549 & 0,5210 & 0,5266 & 0,6020 & 0,6190 & 0,6123 \\ 0,4544 & 0,5092 & 0,5449 & 0,5874 & 0,5404 & 0,6190 & 0,8236 & 0,7867 \\ 0,4743 & 0,5344 & 0,5695 & 0,6106 & 0,5280 & 0,6123 & 0,7867 & 0,8273 \end{pmatrix}$$

Крок 4. Розраховуємо відстані між об'єктами із X та центрами кластерів за наступною формулою:

$$D_{B_j^k} = (x_j^i - v_j^k) \cdot [(\det(B_j^k)) \cdot (B_j^k)^{-1}] \cdot (x_j^i - v_j^k)^T \quad (4)$$

Для третього кластеру матриця має наступний вид:

$$D_{B^3} \begin{pmatrix} 1,4151 & 3,5962 & 0,0567 & 0,4302 & 1,0251 & 0,7188 & 3,7478 & \dots & 0,2765 \\ 3,5962 & 16,5886 & 2,0982 & 1,1672 & 0,1606 & 2,3532 & 6,9467 & \dots & 2,8446 \\ 0,0567 & 2,0982 & 1,1471 & 1,0488 & 0,1982 & 1,3670 & 0,3431 & \dots & 1,0079 \\ 0,4302 & 1,1672 & 1,0488 & 6,9062 & 2,6100 & 0,5972 & 3,8626 & \dots & 0,9283 \\ 1,0251 & 0,1606 & 0,1982 & 2,6100 & 2,9383 & 0,4526 & 4,3004 & \dots & 0,2809 \\ 0,7188 & 2,3532 & 1,3670 & 0,5972 & 0,4526 & 3,4749 & 1,5800 & \dots & 1,1377 \\ 3,7478 & 6,9467 & 0,3431 & 3,8626 & 4,3004 & 1,5800 & 18,8696 & \dots & 1,4831 \\ 3,3776 & 11,1887 & 1,6619 & 2,4023 & 2,1277 & 3,4997 & 8,2875 & \dots & 1,9890 \\ 0,7372 & 0,8188 & 0,6980 & 1,2299 & 0,6961 & 0,6732 & 1,2256 & \dots & 0,7000 \\ 4,0245 & 13,3506 & 1,1373 & 1,2955 & 1,1556 & 3,2830 & 9,6540 & \dots & 1,7494 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0,2765 & 2,8446 & 1,0079 & 0,9283 & 0,2809 & 1,1377 & 1,4831 & \dots & 1,3108 \end{pmatrix}$$

Таблиця 3 –

Матриця належності нечіткої кластеризації алгоритмом Густафсона-Кесселя

№ п/п	Платник податків	Категорії уваги		
		Зелені	Жовті	Червоні
1	Коломійцева Галина Юріївна	0,4080	0,1126	0,8794
2	Мартиненко Олена Анатоліївна	0,0906	0,7381	0,1713
3	Юрченко Людмила Михайлівна	0,5812	0,3300	0,0888
4	Коноводов Юрій Миколайович	0,2026	0,2598	0,5376
...
96	Зайцева Вікторія Олександрівна	0,6164	0,2878	0,0958
97	Коваленко Валентина Василівна	0,2191	0,2033	0,5776
98	Глушко Лілія Володимирівна	0,0644	0,6655	0,2701
99	Кульбачна Тетяна Валентинівна	0,4795	0,4388	0,0817
100	Давідайтис Надія Миколаївна	0,5604	0,3260	0,1136

Крок 5. Перераховуємо елементи матриці нечіткої розбивки.

Якщо $D_{B^k} = 0$, то для відповідного нечіткого

кластера $\mu'_k(a_i) = 1$, а для інших $\mu'_k(a_i) = 0$.

Якщо $D_{B^k} > 0$, то проводимо розрахунок за

формулою:

$$\mu'_{A_k}(a_i) = \left(\frac{\sum_{l=1}^c \left(\frac{\left(\sum_{j=1}^q (x_j^i - v_j^k)^2 \right)^{\frac{1}{2}}}{\left(\sum_{j=1}^q (x_j^i - v_j^l)^2 \right)^{\frac{1}{2}}} \right)^{\frac{2}{m-1}}}{\sum_{l=1}^c \left(\frac{\left(\sum_{j=1}^q (x_j^i - v_j^l)^2 \right)^{\frac{1}{2}}}{\left(\sum_{j=1}^q (x_j^i - v_j^l)^2 \right)^{\frac{1}{2}}} \right)^{\frac{2}{m-1}}} \right)^{-1} \quad (5)$$

Крок 6. Якщо кількість виконаних ітерацій перевищує задане число s або ж модуль різниці $\left| f'(A_k, v_j^k) - f'(A_k, v_j^k) \right| \leq \varepsilon$, тобто не перевищує значення параметра збіжності алгоритму ε , то в якості шуканого результату нечіткої кластеризації прийняти нечітку розбивку μ' і закінчити виконання алгоритму. А якщо ні, то вважати поточною нечіткою розбивкою μ' і перейти на крок 3 алгоритму, збільшивши на 1 кількість виконаних ітерацій.

При виконанні алгоритму було зроблено 22 ітерації, так як різниця значень цільових функції на останній і передостанній ітераціях не перевищила значення параметра збіжності алгоритму, то нашою нечіткою розбивкою платників податків за категоріями уваги будемо вважати нечітку розбивку на 21 ітерації. Результати нечіткої розбивки платників податків за категоріями уваги на основі алгоритму Густафсона-Кесселя показані у табл. 3.

Результати кластеризації можуть бути сформовані наступним чином, наприклад, платник податків Глушко Лілія Володимирівна з більшою ймовірністю (приблизно 67%) відноситься до жовтої категорії уваги, тобто Глушко Лілія Володимирівна – платник податків помірною ризику.

Досвід вирішення прикладних завдань показує, що найбільш ефективний шлях отримання адекватних результатів полягає в багаторазовому виконанні алгоритму для різних вихідних розбивок. Отримані результати порівнюються значення цільової функції отриманих нечітких розбивок з метою ухвалення остаточного рішення про шукану нечітку кластеризацію [7].

Для оцінки якості кластеризації можна використовувати величину силуету S [10]. При нечіткої кластеризації номер кластера визначається за максимальним значенням ступеня належності. Значення силуету виражається для кожного об'єкта в такий спосіб:

$$S(x_i) = \frac{a(x_i) - b(x_i)}{\max(a(x_i), b(x_i))}, \quad (6)$$

де $a(x_i)$ – середня відстань між об'єктом x_i ($x_i \in k$, $k = \overline{1, c}$) і об'єктами того ж кластера k до якого належить x_i ;

$b(x_i)$ – мінімальна відстань між об'єктом x_i і об'єктами в кластері, який ближче всього до кластера k , тобто кластер до якого x_i не належить.

Значення силуету лежить в інтервалі $[-1; 1]$, якщо воно від'ємне, то платник податків вважається погано кластеризованим [11].

По алгоритму Густафсона-Кесселя значення силуету для першого платника податків буде дорівнювати: $S = (0,1213 - 0,0156) / 0,1213 = 0,8714$, так як наше значення додатне, це означає, що платник податків дуже добре кластеризован.

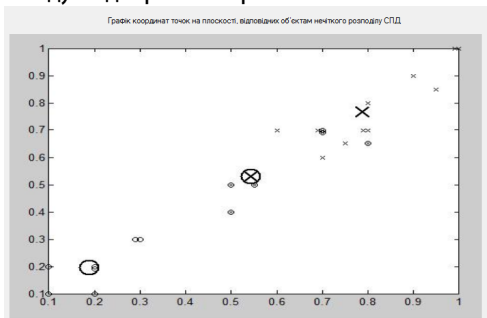


Рисунок 1 – Графік координат точок на площині, відповідних об'єктам нечіткого розподілу платників податків

ЛІТЕРАТУРА:

1. Вятчин Д.А. Нечеткие методы автоматической классификации: монография /Д.А. Вятчин. – Мн.: УП «Технопринт», 2004 – 219 с.
2. Жидкова Е.Ю. Налоги и налогообложение /Е.Ю. Жидкова; 2-е изд., перераб. и доп. – М.: Эксмо, 2009. – 480 с.
3. Наказ «Про затвердження Методичних рекомендацій щодо порядку розподілу платників податків за категоріями уваги N 442 14.04.2009» [Електронний ресурс] /В. Янукович. – 2009. – Режим доступу до наказу: http://search.ligazakon.ua/l_doc2.nsf/link1/GDPI7866.html.
4. Gan G. Data Clustering: theory, algorithms and applications /G. Gan, C. Ma, J. Wu. – Alexandria: ASA-SIAM, 2007. – 488 p.
5. Bezdek J.C. Pattern recognition with fuzzy objective function algorithms /J.C. Bezdek. – New York: Plenum Press, 1981. – 345 p.
6. Сокал Р.Р. Кластер-анализ и классификация: предпосылки и основные направления /Р.Р. Сокал; под ред. Дж. Вэн Райзина. – М.: Мир, 1990.
7. Леоненков А.В. Нечеткое моделирование в среде MATLAB и fuzzyTECH. – СПб.: БХВ-Петербург, 2005. – 736 с.
8. Gustafson D.E. Fuzzy clustering with a fuzzy covariance matrix /D.E. Gustafson W.C. Kessel //Proc. IEEE CDC. – 1979. – №7. – P.773-781.
9. Штовба С.Д. Проектирование нечетких систем средствами MATLAB. – М.: Горячая линия-Телеком, 2007. – 288 с.
10. Rousseeuw J.P. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis /J.P. Rousseeuw //J. Comp. Appl. Math. – 1987. – №20. – P.53-65.
11. Bhuyan P.K. FCM clustering using GPS data for defining level of service criteria of urban streets in Indian context /P.K. Bhuyan, K.V.K. Rao //Transport problems. – 2010. – №5. – P.105-113.
12. Заболотникова В.С. Учет субъектов предпринимательской деятельности в условиях неопределенности на основе метода нечеткой кластеризации /В.С. Заболотникова //Научные труды ДонНТУ. Серия «Информатика, кибернетика и вычислительная техника». – 2011. – №14 (188). – С.283-290.

Результати також можуть бути отримані в системі MathLab. Цей метод є більш трудомістким, але має гнучкість і можливість відображення функцій належності. Результат вирішення завдання нечіткої кластеризації системою MathLab для трьох нечітких кластерів представлений візуалізацією вихідних даних і результатів нечіткої кластеризації у двовимірному просторі на площині (рис. 1).

ВИСНОВКИ

Метод нечіткої кластеризації, а саме алгоритм Густафсона-Кесселя, є досить ефективним методом розподілу платників податків за категоріями уваги. Однак виконана з його допомогою розбивка не є остаточним кроком у поставленій меті на шляху організації виїзних податкових перевірок платників податків. Проведена нечітка кластеризація дозволила багатосторонньо проаналізувати природу процесів податкової служби та спростити обробку даних.

Для реалізації методів нечіткого управління більш природно описується характер людського мислення і хід його міркувань, ніж в формально-логічних системах, існує можливість будувати моделі, які найбільш адекватно відображають різні аспекти невизначеності, які постійно присутні у навколишній реальності.

Перспективи подальших досліджень спрямовані на розробку інформаційної технології автоматизованого контролю процесів податкової служби.