

**БІОЛОГІЧНІ ТА МЕДИЧНІ ПРИЛАДИ І СИСТЕМИ**

УДК 004.032.26

О. К. КОЛЕСНИЦЬКИЙ, Ю. О. ЖУРАВСЬКА

Вінницький національний технічний університет, Вінниця

**ЗАСТОСУВАННЯ НЕЙРОННОЇ МЕРЕЖІ КОХОНЕНА ДЛЯ МЕДИЧНОГО ДІАГНОСТУВАННЯ ПАЦІЄНТІВ ЗА АНАЛІЗОМ КРОВІ**

**Анотація:** В даній статті запропоновано нейронну мережу Кохонена, в якій центри кластерів визначаються заздалегідь згідно з нормальними та відхиленими від норми значеннями показників загального аналізу крові, що підвищує достовірність попереднього діагностування пацієнтів.

**Ключові слова:** нейронна мережа Кохонена, кластеризація, медична діагностика.

**Анотация:** В данной статье была предложена нейронная сеть Кохонена, в которой центры кластеров определяются заранее в соответствии с нормальными и отклонёнными от нормы значениями показателей общего анализа крови, что повышает достоверность предварительной диагностики пациентов.

**Ключевые слова:** нейронная сеть Кохонена, кластеризация, медицинская диагностика.

**Abstract:** A Kohonen neural network, in which the cluster centers are determined in advance in accordance with normal and deviations from the normal values of general blood count parameters that increases the reliability of preliminary diagnostics of patients, is proposed in current article.

**Keywords:** Kohonen neural network, clustering, medical diagnostics.

**Вступ**

На даний час глобальний процес автоматизації зачепив практично всі сфери людської діяльності, включаючи медичну галузь. Існуючий процес діагностування пацієнтів лікарень є досить довгим та багатостадійним – хворий повинен пройти реєстрацію, консультацію у сімейного лікаря (терапевта), отримати направлення до лікаря, спеціалізація якого має ймовірність виявитись хибною відносно реальної хвороби пацієнта. Тому щоб підтвердити або спростувати діагноз потрібно здати низку аналізів, що може призвести через довготривалість даного процесу до загострення хвороби або поставлення хибного діагнозу. У даній ситуації раціонально використовувати можливості сучасних інформаційних технологій та широке розповсюдження мобільних інформаційних пристроїв у населення.

Існуючі автоматизовані методи діагностики також є недостатньо ефективними – більшість систем орієнтована на визначення конкретного захворювання, що через специфіку медичної галузі, має високу ймовірність виявитись хибним; мають невиправдану складність алгоритмів та їх програмної реалізації; можуть використовуватись лише медичними фахівцями. [1]

Таким чином, існує проблема недосконалості сучасних методів діагностування, що не забезпечують високу точність виявлення захворювання, швидкість (оперативність) та зручність використання із залученням персональних мобільних засобів (телефони, смартфони, планшети та ін.), не можуть бути використані пацієнтами дистанційно без відвідування лікарні.

Для усунення даної проблеми доцільно застосувати підхід, заснований на визначенні показників загального аналізу крові, що проводиться при більшості захворювань і профілактичних обстеженнях, і включає в себе 13 показників, кожен з яких при відхиленні від норми передбачає схильність до певних хвороб. Діагностування за допомогою аналізу отриманих показників аналізу крові допоможе виявити тип захворювання пацієнта та надати рекомендації щодо лікування. [2] Для реалізації даного підходу раціонально використати кластеризацію даних, яка зможе забезпечити високу точність їх обробки, при цьому мати нескладну програмну структуру. [3]

**Задачею дослідження** є обрання методу кластеризації, який забезпечує найбільш якісну реалізацію сформульованого підходу до діагностування, що повинен гарантувати високу точність та швидкість виявлення типу захворювання. При цьому на основі значень показників загального аналізу крові потрібно визначити попередній діагноз пацієнту, тобто групу, до якої відноситься можливе захворювання.

Таким чином, **мета цієї статті** – аналіз відомих методів кластеризації даних та обрання найточнішого для забезпечення достовірного попереднього діагностування пацієнтів на основі аналізу крові з метою надання рекомендацій щодо лікування.

**Аналіз відомих методів кластеризації**

Основними методами кластеризації є група ієрархічних алгоритмів, алгоритм k-means та його похідні, дерево мінімального покриття, метод найближчого сусіда, алгоритм нечіткої кластеризації, нейронні мережі та генетичні алгоритми.

Ієрархічні алгоритми спираються на вихідну «гіпотезу компактності»: у просторі об'єктів всі близькі об'єкти повинні ставитися до одного кластеру, а всі різні об'єкти відповідно повинні знаходитися в різних кластерах, що може призвести до великої похибки при кластеризації даних. Основним недоліком

методу k-means є те, що потрібно заздалегідь задавати  $k$  - кількість кластерів та еталонів, що не завжди можливо зробити раціонально. Метод є дуже чутливим до початкових наближень значень центрів. Дерево мінімального покриття може виділяти кластери довільної форми, але воно має громіздку структуру та ненаочність подання результатів кластеризації. Алгоритм найближчого сусіда простий у реалізації, швидко виконується, але, як і інші «жадібні» алгоритми, може видавати неоптимальні рішення. Нечіткі алгоритми недоцільно використовувати, якщо заздалегідь невідомо число кластерів, або необхідно однозначно віднести кожен об'єкт до одного кластеру. Для реалізації кластеризації даних мережею Кохонена використовується універсальний кластеризатор – нейронна мережа, передбачається навчання мережі без вчителя та самоорганізація мережі, існує суттєва простота реалізації та гарантоване отримання відповіді після проходження даних по шару нейронів. Недоліками є робота тільки з числовими даними, мінімізація розмірів мережі та необхідність задання кількості кластерів. Генетичні алгоритми не гарантують виявлення глобального рішення за прийнятний час, не гарантують, що знайдене рішення буде оптимальним та мають переваги перед іншими алгоритмами лише при дуже великих розмірах завдань і відсутності впорядкованості у вихідних даних, коли альтернативою є метод повного перебору варіантів. [4]

Оскільки для кластеризації результатів показників загального аналізу крові потрібні лише числові значення та велика точність в їх визначенні, для реалізації даної задачі була обрана мережа Кохонена, найефективніша нейронна мережа кластеризації числових даних. [5]

Оскільки стандартна реалізація мережі Кохонена не передбачає визначеної кількості кластерів, та їх початкові центри приймають випадкові малі значення, була проведена модифікація даної мережі, що полягає у попередньому визначенні кількості кластерів та їх центрів, які відповідають нормальним та відхиленням від норми значенням показників загального аналізу крові.

#### Модернізація нейронної мережі Кохонена

Задача, що розглядається, запропонована до вирішення модифікованою мережею Кохонена. Дана мережа використовує неконтрольоване навчання та навчальна множина складається лише із значень вхідних змінних.

Мережа розпізнає кластери в навчальних даних і розподіляє дані до відповідних кластерів. Якщо в наступному мережа зустрічається з набором даних, несхожим на жодний із відомих зразків, вона відносить його до нового кластеру. Якщо в даних містяться мітки класів, то мережа спроможна вирішувати задачі класифікації.

Шар Кохонена складається з деякої кількості  $n$  паралельно діючих лінійних елементів. Всі вони мають однакову кількість входів  $m$  та отримують на свої входи один і той самий вектор вхідних сигналів  $x=(x_1...x_m)$ . На виході  $j$ -го лінійного елемента отримуємо сигнал, що розраховується за формулою (1).

$$y_j = w_{j0} \sum_{i=1}^m w_{ji} x_i \quad (1)$$

де  $w_{ji}$  — ваговий коефіцієнт  $i$ -го входу  $j$ -го нейрону,  $w_{j0}$  — пороговий коефіцієнт.

Після проходження шару лінійних елементів сигнали посилаються на обробку за правилом «переможець забирає все»: серед вихідних сигналів  $y_j$  шукається максимальний; його номер  $j_{\max} = \arg \max_j \{y_j\}$ . Остаточо, на виході сигнал з номером  $j_{\max}$  дорівнює одиниці, всі інші — нулю. Якщо максимум одночасно досягається для декількох  $j_{\max}$ , то приймають всі відповідні сигнали рівними одиниці.

Мережа Кохонена має два прошарки: вхідний і вихідний, що називають самоорганізованою картою. Елементи карти розташовуються в деякому просторі - як правило двовимірному, який передбачено в реалізації запропонованого підходу (рис. 1).[6]

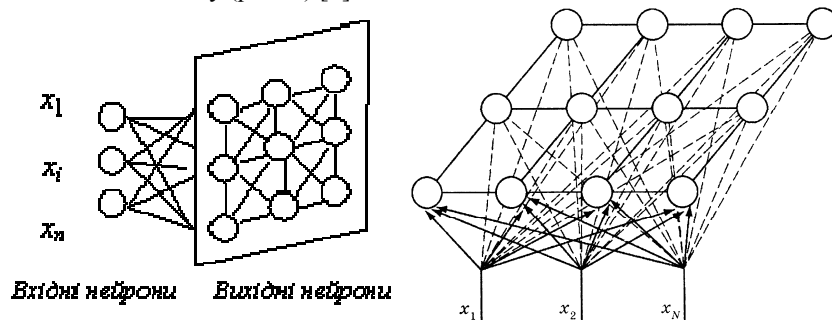


Рисунок 1 – Двовірна структура мережі Кохонена

Навчання розроблюваної мережі відбувається наступним чином. На початку роботи визначається кількість кластерів та їх центри. Після цього деякий вхідний вектор з набору навчальних вибирається і встановлюється на вході нейронної мережі. На цьому етапі відмінності між вхідним вектором та всіма векторами обчислюються за формулою (2).

$$D_{ij} = |X^l - W_{ij}| = \sqrt{(x_1 - w_{ij1})^2 + \dots + (x_n - w_{ijn})^2}, \quad (2)$$

де  $i$  та  $j$  – показники нейронів у вихідному шарі. Після цього нейронна мережа обирає нейрон-переможця з переліку визначених центрів кластерів, тобто такий, щоб його ваговий вектор був схожий на вхідний за формулою (3).

$$D(k_1, k_2) = \min_{i,j} D_{i,j}, \quad (3)$$

де  $k_1$  та  $k_2$  – показники нейрона-переможця. Після цього проводиться корекція вагових векторів переможця та сусідніх нейронів. Близні нейрони до переможця визначаються топологічною функцією сусідства, яка розраховується за формулою 4.

$$h(p, t) = \exp\left(-\frac{p^2}{\sigma^2(t)}\right) \left(1 - \frac{2}{\sigma^2(t)} p^2\right), \quad (4)$$

де  $p$  – відстань до нейрона переможця, яка знаходиться за формулою 5.

$$p = \sqrt{(k_1 - i)^2 + (k_2 - j)^2} \quad (5)$$

де  $\sigma$  – функція, що визначає радіус сусідства. На початку функціонування програмного модулю вона включає весь простір сенсорного поля (сітки), але з часом, значення її зменшується.

У якості функції сусідства була обрана функція «Мексиканський капелюх», що згідно з наведеними результатами експериментів у таблиці 1, забезпечує більшу точність розподілу даних на кластери.

Після обчислення топологічної функції ваги всіх нейронів переобчислюються за формулою (6).

$$W_{ij}(t+1) = W_{ij}(t) + \alpha(t)h(p, t)(X^l(t) - W_{ij}(t)) \quad (6)$$

де  $\alpha(t)$  – функція швидкості навчання, яка також змінюється з часом. Якщо нейрон є переможцем, або сусіднім до нього, його вектор ваг оновлюється або залишається незмінним в іншому випадку. На кожному кроці нейронна мережа визначає нейрон, чий ваговий вектор найбільш схожий до вхідного, та коригує його ваги та ваги сусідів, щоб наблизити їх до вхідного вектора (рис. 2).

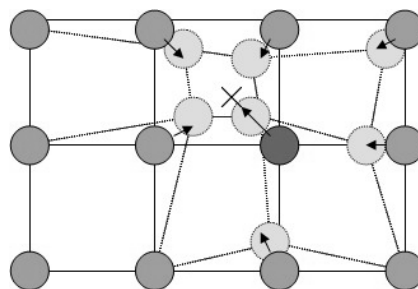


Рисунок 2 – Оновлення нейрону-переможця та його сусідів та підштовхування в сторону вхідного вектору, що на рисунку позначений «X». Суцільні та пунктирні лінії визначають ситуацію до та після оновлення, відповідно

Кожен вхідний вектор з навчальної вибірки представляється нейронній мережі, і навчання триває або деяке фіксоване число циклів, або доки різниця між вхідним і ваговим векторами досягне заданого значення  $\epsilon$ . Різниця між сусідніми нейронів зменшується з плином часу, і, отже, вони організуються в групи (кластери), які відповідають одному з класів з навчальної множини. [7]

Для оптимальних розрахунків мережі була також введена процедура попередньої обробки вхідних даних, при якій значення ознак, що утворюють вхідний вектор, приводяться до деякого заданому діапазону. Нормалізація необхідна, оскільки вихідні значення ознак змінюються в достатньо великому діапазоні, тому існує ймовірність некоректної роботи нейронної мережі з такими даними. Так, в одному вхідному векторі можуть міститися значення, що відрізняються один від одного на кілька порядків: стандартний рівень гемоглобіну – 120-180г/л, колірного показнику – 0,85-1,15%. Такий дисбаланс між значеннями показників може викликати нестійкість роботи моделі, погіршити результати навчання і уповільнити його процес. Після нормалізації всі значення вхідних ознак будуть приведені до деякого вузького діапазону, а саме - [0 ... 1] що дозволяє мережі працювати з даними більш коректно.

Для реалізації запропонованого підходу до діагностування була використана нормалізація вхідних даних за формулою (7).

$$y = \frac{(x - x_{\min})(d_2 - d_1)}{x_{\max} - x_{\min}} + d_1 \quad (7)$$

де:  $x$ - значення, що підлягає нормалізації,  $[d_1, d_2]$ - інтервал значень  $x$ ,  $[x_{\min}, x_{\max}]$ - інтервал, до якого буде зведене значення  $x$ . [8]

### Програмна реалізація діагностування на основі нейронної мережі Кохонена

Описаний підхід до діагностування був реалізований за допомогою мови програмування C# (рис.3). Вхідні дані представлені у вигляді стовпців даних, де перший стовпець відповідає імені пацієнту, другий – його статі, стовпці з третього по одинадцятий відповідають показникам загального аналізу крові: гемоглобін, еритроцити, колірний показник, ретикулоцити, тромбоцити, ШОЕ, лейкоцити, паличкоядерні, сегментоядерні, еозинофіли, базофіли, лімфоцити, моноцити.

Вхідні дані	Результати
A1 ж 120 3.7 0.93 0.5 242 8 6 2 61 4 0,4 29 2	Пацієнт A1 - здорова
A2 ч 141 4.1 0.97 0.9 216 8 7 3 57 7 0,6 37 8	Пацієнт A2 - можлива астма
A3 ч 157 4.1 0.90 0.9 299 5 4 1 55 6 0,5 24 6	Пацієнт A3 - можлива астма
A4 ч 153 4.9 0.91 1.2 346 15 11 7 72 5 0,4 24 8	Пацієнт A4 - можливе запалення в організмі
A5 ж 140 4.2 0.99 1.2 301 7 4 3 58 2 0,6 19 4	Пацієнт A5 - здорова
A6 ч 144 3.8 1.1 0.7 232 5 3 1 52 2 0,5 19 8	Пацієнт A6 - можливі проблеми з селезінкою
A7 ч 158 4.0 0.91 0.2 206 4 4 2 66 5 0,7 22 3	Пацієнт A7 - здоровий
A8 ч 145 4.2 1.05 1.3 192 4 4 1 53 3 0,3 39 9	Пацієнт A8 - можливі проблеми з нирками
A9 ч 146 3.7 1.04 0.4 380 16 10 7 81 5 0,4 36 7	Пацієнт A9 - можливе запалення в організмі
A10 ч 159 5.7 1.14 0.2 225 6 5 6 71 2 0,2 27 7	Пацієнт A10 - здоровий
B1 ч 151 3.9 0.99 1.0 186 5 3 6 72 3 0,5 35 3	Пацієнт B1 - можливі проблеми з селезінкою
B2 ч 159 3.7 1.02 0.5 215 5 6 5 56 2 0,3 43 11	Пацієнт B2 - ймовірність захворювання туберкульозом
B3 ч 157 4.2 0.92 0.9 271 4 4 3 58 2 0,2 19 9	Пацієнт B3 - здоровий
B4 ж 133 3.9 1.12 0.6 256 10 4 2 48 6 0,8 21 2	Пацієнт B4 - можлива астма
B5 ч 188 5.7 1.2 0.2 300 6 6 4 62 4 0,6 21 3	Пацієнт B5 - можливе зневоднення організму
B6 ч 151 4.0 1.12 0.3 239 6 5 2 48 3 0,5 26 5	Пацієнт B6 - здоровий
B7 ч 141 3.9 0.92 0.2 205 5 4 3 57 7 0,8 38 9	Пацієнт B7 - можлива астма
B8 ч 157 4.3 1.03 0.2 252 6 6 4 67 4 0,4 31 3	Пацієнт B8 - здоровий
B9 ч 142 4.8 0.94 0.3 218 5 5 2 51 2 0,3 29 7	Пацієнт B9 - здоровий
B10 ч 134 5.1 0.74 1.3 172 12 4 1 62 3 0,4 20 5	Пацієнт B10 - можлива анемія
C1 ч 140 4.7 1.01 0.6 214 7 4 1 55 2 0,2 37 7	Пацієнт C1 - можливі проблеми з цитоподібною залозою
C2 ч 157 4.9 0.98 0.9 209 6 5 2 66 4 0,4 23 5	Пацієнт C2 - здоровий
C3 ж 157 4.6 1.28 0.5 312 12 4 1 55 2 0,3 36 4	Пацієнт C3 - можливе зневоднення організму
C4 ч 130 4.6 0.85 0.6 244 5 4 1 55 2 0,3 25 4	Пацієнт C4 - здоровий
C5 ч 145 4.1 0.92 0.1 301 8 6 2 65 4 0,6 30 6	Пацієнт C5 - можливі проблеми з нирками
C6 ч 133 4.1 0.94 1.0 247 9 7 3 70 5 0,9 34 6	Пацієнт C6 - здоровий
C7 ч 133 3.7 1.05 1.1 315 13 8 2 68 -1 -0,3 21 8	Пацієнт C7 - можливе запалення в організмі, можлива алергія
C8 ч 119 3.0 0.91 1.0 312 14 4 2 57 2 0,3 29 5	Пацієнт C8 - можлива анемія
C9 ч 143 4.7 0.97 0.1 221 6 5 3 63 4 0,7 19 5	Пацієнт C9 - можливі проблеми з нирками
C10 ч 135 4.7 0.79 1.1 320 5 3 7 76 5 0,5 32 8	Пацієнт C10 - здоровий
D1 ч 150 4.2 0.87 0.7 207 8 8 6 73 5 1,1 19 10	Пацієнт D1 - можлива інфекція в організмі
D2 ч 139 4.1 1.00 1.1 245 6 6 3 63 4 0,5 34 5	Пацієнт D2 - здоровий

Рисунок 3 – Приклад роботи програми

### Результати тестування програми попереднього діагностування пацієнтів

При тестуванні програми перевірялась достовірність її роботи при виборі різних функцій сусідства[6, 9]:

1) Функція Гауса (8);

$$h(p, t) = \exp\left(-\frac{p^2}{2\sigma^2(t)}\right) \quad (8)$$

2) Функція «Мексиканський капелюх» (9);

$$h(p,t) = \exp\left(-\frac{p^2}{\sigma^2(t)}\right)\left(1 - \frac{2}{\sigma^2(t)} p^2\right) \quad (9)$$

3) Функція «Французький капелюх» (10);

$$h(p) = \begin{cases} 1, & |p| \leq \alpha \\ -\frac{1}{3}, & \alpha < |p| \leq 3\alpha \\ 0, & |p| > 3\alpha \end{cases} \quad (10)$$

Для вибору найефективнішої з представлених функцій сусідства, проведемо порівняння їх точності на основі вибірки у 1000 вхідних векторів, де 800 – навчальна вибірка, 200 – тестова;  $\epsilon=10^{-4}$  (табл.1), для проведення експериментів була створена база даних, основана на безкоштовних загальнодоступних базах даних показників аналізів крові та відповідних хвороб. [10-11].

Таблиця 1 – Функції сусідства та досягнена достовірність діагностики

Функція сусідства	Кількість правильно розпізнаних вхідних векторів (з 200)	Кількість неправильно розпізнаних вхідних векторів (з 200)	Достовірність діагностики
Функція Гауса	166	34	83%
Функція «Мексиканський капелюх»	186	14	93%
Функція «Французький капелюх»	180	20	90%

Таким чином, з результатів експерименту видно, що обрана для подальшої реалізації функція «Мексиканський капелюх» дає найбільшу точність кластеризації.

Проведемо порівняння кластеризації звичайною мережею Кохонена та запропонованою модифікованою мережею з наперед заданими кластерами та їх центрами на основі точності розбиття вибірки у 10, 100, 500 та 1000 вхідних векторів тестової вибірки та  $\epsilon=10^{-4}$  (табл.2).

Таблиця 2 – Досягнена достовірність діагностування звичайною мережею Кохонена та запропонованою модифікованою мережею

Обсяг вибірки	Класична нейронна мережа Кохонена			Модифікована нейронна мережа Кохонена		
	кількість правильно розпізнаних вхідних векторів	кількість неправильно розпізнаних вхідних векторів	Достовірність діагностики	кількість правильно розпізнаних вхідних векторів	кількість неправильно розпізнаних вхідних векторів	Достовірність діагностики
10	7	3	70%	8	2	80%
100	73	27	73%	82	18	82%
500	355	145	71%	394	106	78,8%
1000	716	284	71,6%	803	197	80,3%

У результаті експериментальних досліджень було встановлено, що запропонований метод має більш високі показники точності роботи у порівнянні із класичним, що дає можливість значно підвищити якість діагностування пацієнтів.

Існуюча інтерактивна система діагностування Diagnos.ru, що має найбільшу в світі базу захворювань (близько 240) заснована на методі нечіткої логіки. При цьому точність виявлення захворювання складає в середньому 68% [12], що нижче зазначеного отриманого результату.

### Висновки

У роботі запропоновано модифіковану нейронну мережу Кохонена, в якій центри кластерів визначаються заздалегідь згідно з нормальними та відхиленними від норми значеннями показників загального аналізу крові з використанням функції сусідства «Мексиканський капелюх». Це рішення покладено в основу підходу до діагностування пацієнтів на основі загального аналізу крові.

Отже, була вирішена задача знаходження підходу та його реалізації до визначення типу хвороби згідно результатів загального аналізу крові, що характеризується простотою числової реалізації, більш високою точністю роботи у порівнянні з аналогами.

### Список літератури

1. Алгоритми діагностування пацієнтів [Електронний ресурс]. Режим доступу: <http://www.vidal.by/vracham/Informatsiya-dlya-spetsialistov/Nevrologiya-psihiatriya/Algoritmy-diagnostiki-i-vedeniya-patsientov/> - Назва з екрану
2. Загальний аналіз крові [Електронний ресурс]. Режим доступу: <http://ukrhealth.net/zagalnij-analiz-krovi/>
3. Обзор методов кластеризации [Електронний ресурс]. Режим доступу: [http://www.dialog21.ru/Archive/2001/volume2/2\\_26.htm](http://www.dialog21.ru/Archive/2001/volume2/2_26.htm) - Назва з екрану
4. Симахин В.А., Кашинцев А.В. Кластеризация медицинских данных с помощью нейросетей [Електронний ресурс]. Режим доступу: <http://neurocomp.ru/klasterizaciya-medicinskix-dannyx-s-pomoshhyu-nejrosetej/> - Назва з екрану
5. О С. Амосов. Интеллектуальные информационные системы. Нейронные сети и нечеткие системы: Учеб. Пособие. - Комсомольск-на-Амуре: ГОУВПО «КнАГТУ», 2004. -104 с.
6. Руденко О.Г., Бодянский Е.В. Искусственные нейронные сети – Харьков, 2005. – 407с.
7. Дяченко В.А., Михаль О.Ф. Адаптивное параллельное обучение модифицированной саморганизующейся карты Кохонена [Електронний ресурс]. Режим доступу: [http://archive.nbuv.gov.ua/portal/natural/bionint/2012\\_1/Mihal1.pdf](http://archive.nbuv.gov.ua/portal/natural/bionint/2012_1/Mihal1.pdf)
8. Дмитро Парфенович – «Нейронні мережі – від теорії до практики»[Електронний ресурс]. Режим доступу - <http://www.mql5.com/ru/articles/497> - Назва з екрану
9. Применение модернизированной вейвлет-функции «Французская шляпа» для аппроксимации продольного распределения магнитного поля в магнитных реверсивных фокусирующих системах. Сравнение её показателей. [Електронний ресурс]. Режим доступу: <http://www.moluch.ru/archive/44/5367/> - Назва з екрану.
10. Laboratory Procedure Material. Complete Blood Count using HMX – NHANES 2007-200810. Mayo clinic. Diseases and conditions [Електронний ресурс]. Режим доступу: <http://www.mayoclinic.org/diseases-conditions/high-blood-cholesterol/basics/tests-diagnosis/con-20020865> - Назва з екрану.
11. CLL Patient Databases [Електронний ресурс]. Режим доступу: <https://patientdatabases.org/wp/about>
12. Диагноз.ру. Инновационный медицинский центр [Електронний ресурс]. Режим доступу: <http://www.diagnos.ru/> - Назва з екрану.

### Відомості про авторів

**Колесницький Олег Костянтинович** - кандидат технічних наук, доцент, доцент кафедри комп'ютерних наук, Вінницький національний технічний університет.

**Журавська Юлія Олександрівна** – студентка групи ІКН-10 Вінницький національний технічний університет.

## ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

УДК 004.912

О.В. ЛОЗИНСЬКА, М.В. ДАВИДОВ, В.В. ПАСІЧНИК

Національний університет «Львівська політехніка», м. Львів

## МАШИННИЙ ПЕРЕКЛАД НА ОСНОВІ ПРАВИЛ ДЛЯ ПЕРЕКЛАДУ НА УКРАЇНСЬКУ ЖЕСТОВУ МОВУ

**Анотація.** Розглянуто сучасні досягнення у розв'язанні проблеми комп'ютерного перекладу з української словесної мови на українську жестову. Описано правила обходу дерева синтаксичного розбору речень українською мовою для синтезу речень жестовою мовою. На основі прикладів дерев синтаксичного розбору речень української словесної мови визначено основні правила перекладу на українську жестову. Розглянуто стверджувальні, заперечні й запитальні речення.

**Ключові слова:** українська жестова мова, корпус паралельних текстів, машинний переклад на основі правил, дерева синтаксичного розбору.

**Аннотация.** Рассмотрены современные достижения в решении проблемы компьютерного перевода с украинского словесного языка на украинский жестовый. Описаны правила обхода дерева синтаксического разбора предложений на украинском языке для синтеза предложений на жестовом языке. На основе примеров деревьев синтаксического разбора предложений украинского словесного языка определены основные правила перевода на украинский жестовый. Рассмотрены утвердительные, отрицательные и вопросительные предложения.

**Ключевые слова:** украинский жестовый язык, корпус параллельных текстов, машинный перевод на основе правил, дерева синтаксического разбора.

**Abstract.** Modern achievements in solving the problem of computer translation from Ukrainian Spoken Language into Ukrainian Sign Language are considered. Rules of traversing tree parsing sentences in Ukrainian for the synthesis of Sign Language sentences are described. On the basis of the examples of tree parsing sentences of Ukrainian Spoken Language, the rules for Sign Language translation are defined. Declarative, negative and interrogative sentences are considered.

**Keywords:** Ukrainian Sign Language, bilingual corpora, rule-based machine translation, parsing trees.

## Вступ

Розвиток сучасної науки, комп'ютеризація суспільства, використання мультимедійних та інтернет-технологій створили умови для розроблення комп'ютеризованих систем комунікації людей із особливими потребами, зокрема людей з вадами слуху. В Україні проживає близько 400 тисяч людей з вадами слуху. Для такої категорії громадян функціонує 59 спеціалізованих шкіл, 20 університетів (зокрема Національний технічний університет України «Київський поліграфічний інститут»). Актуальним завданням є розроблення програмних засобів, які допомагають у вивченні та перекладі української жестової мови.

Жестова мова (ЖМ) – це комунікативна система, висловлювання якої побудовані не на звуковій, а на жестикуляційно-мімічній основі (жести рук, артикуляція губ, міміка та емоції). Українська жестова мова (УЖМ) є природною візуально-просторовою мовою, яка використовується українським товариством глухих, заснованому в 1933 році. Ця мова незалежна від української словесної мови (УСМ), і не слід її плутати з калькованою жестовою мовою (КЖМ), яка повністю відповідає УСМ.

Дослідження української жестової мови розпочав Р. Г. Краєвський [1]. Вчений здійснив лінгвістичний опис мови жестів на основі українознавчого матеріалу і створив унікальний жестовий словник у вигляді посібника «Мова жестів глухих» 1964 р. (усього описано більше 1500 жестів). На сьогоднішній день в Україні вивченням синтаксичних особливостей УЖМ займаються Н.Б. Адамюк, І.І. Чепчина [2].

## Постановка проблеми

Проблема перекладу українською жестовою мовою та навпаки є складною з двох причин. Перша з них — недосконалість жестових інтерфейсів комп'ютера, що ускладнює процес введення та виведення жестів. Друга причина — значна лінгвістична відмінність жестової мови від словесної. На відміну від калькованої жестової мови, яка повністю повторює порядок слів речення словесної мови, українська жестова мова має свій синтаксис та правила побудови речення.

Текстові та жестові словосполучення, пов'язані семантичними відношеннями, далеко не завжди перекладаються дослівно. Найчастіше зустрічається випадок, коли словосполучення, яке подає певний смисловий концепт, передається одним жестом або кількома жестами.

Метою статті є систематизація правил обходу дерева синтаксичного розбору речень українською мовою для синтезу речень жестовою мовою. На основі прикладів дерев синтаксичного розбору речень УСМ визначено основні правила перекладу на УЖМ. Розглянуто стверджувальні, заперечні й запитальні речення.

### Огляд відомих підходів

Серед відомих підходів до машинного перекладу жестової мови найбільш успішними виявилися підходи на основі методу трансферу та інтерлінгви [3], [4], [5].

Переклад на жестову мову передбачає синтаксичний розбір речень словесною мовою, усунення омографії, уточнення змісту та синтез правильних конструкцій жестовою мовою.

Не зважаючи на велику кількість наукових робіт, пов'язаних із синтаксичним розбором речень українською мовою [6], [7], [8], досі відсутні відкриті для використання програми автоматичного синтаксичного аналізу речень українською мовою.

Проблемами синтезу УЖМ займаються українські вчені Ю.В. Крак, О.В. Бармак та ін. з Інституту кібернетики ім. В.М. Глушкова НАН України [9]. При реалізації комп'ютерної технології для перекладу з української на жестову мову вчені стикнулися з проблемами, пов'язаними з наявністю двох жестових мов спілкування – калькуючої та української жестових мов, з тим, що немає однозначної відповідності слів жестам. У роботі [10] вчені запропонували механізм узагальнених граматичних конструкцій простих речень української та жестової мов на основі теорії лексикографічних систем [11] і побудували 764 узагальнених граматичних конструкцій із 6559 речень. У роботі не наведено результатів експериментів, які оцінюють правильність перекладу даної системи.

Не зважаючи на те, що лінгвістичні дослідження української жестової мови вже тривалий час проводяться науковцями, відсутні роботи, які описують усі аспекти граматики УЖМ та основних принципів перекладу на УЖМ.

Для англійської мови доступний цілий ряд синтаксичних аналізаторів [12], [13]. Граматичний розбір англійських речень може бути виконаний з допомогою граматики зв'язків (Links Grammar) [14], граматики синхронних приєднаних дерев (synchronous tree-adjointing grammar, STAG) [15], [16], лексико-функційної граматики (Lexical Functional Grammar, LFG) [5] тощо.

Для перекладу на жестову мову відомі різні методи, зокрема методи перекладу на основі прикладів [17], методи перекладу на основі правил [18], методи перекладу на основі даних [19] та статистичні методи перекладу [20], [21]. Іспанськими вченими Р. Сан-Сегундо та ін. [18] порівняно підходи на основі правил, на основі фраз та статистичний підхід для перекладу на іспанську жестову мову. За даними, наведеними у [18] найбільш точним виявився підхід на основі правил. Серед іноземних наукових розроблень для перекладу на жестову мову найбільш відомими є проекти ViSiCAST [22], TEAM [4], ASL Workbench [5], iCommunicator [23] та Artemis [24].

Для того, щоб розробити автоматизовану систему перекладу з української словесної мови на українську жестову та навпаки потрібно розробити синтаксичний аналізатор речень для обох мов, створити систему правил перекладу однієї мови на іншу, а також розробити корпус паралельних речень (УСМ – УЖМ) для порівняння якості перекладу.

### Правила синтезу речень українською жестовою мовою

Структурну основу двоскладного речення утворюють підмет і присудок. Двоскладні речення подані в кожній мові кількома структурними схемами залежно від граматичного оформлення підмета і присудка і залежно від характеру зв'язку між ними. Нижче розглянуто основні групи речень УСМ і наведено дерева синтаксичного розбору для кожної з них.

**Стверджувальні речення.** Дерево синтаксичного розбору для речення УСМ «Я дивлюсь цікавий фільм» зображено на рис. 1.

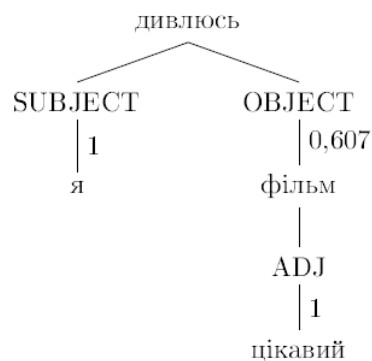


Рисунок 1 – Дерево синтаксичного розбору речення «Я дивлюсь цікавий фільм»

Оскільки українська словесна мова є флективною, то це ж саме речення ми можемо побудувати іншим чином: «Я цікавий фільм дивлюсь» або «Я цікавий дивлюсь фільм». Зміст речення при цьому не зміниться. Це не можна стверджувати щодо української жестової мови, в якій порядок слів строго визначений та підпорядковується правилам граматики жестової мови. Найчастіше слова в українській жестовій розташовуються у порядку SVO («Subject + Verb + Object»), що є прямим порядком розташування слів. Отже, враховуючи особливості української жестової мови, вищенаведене речення на УЖМ буде відтворене таким чином: «Я дивитися фільм цікавий».

На основі правил граматики української жестової мови визначено правила перекладу, за допомогою яких отримуємо речення УЖМ за допомогою дерева синтаксичного розбору речення УСМ.

**Правило №1.** Якщо потрібно відтворити розповідне стверджувальне речення на УЖМ, тоді першим відтворюється суб'єкт (підмет), потім предикат (присудок) і об'єкт (додаток). Після додатка відтворюється його ознака (прикметник). Це саме стосується і ознаки підмета. Отже, набір правил для стверджувального речення УЖМ буде мати такий вигляд:

*VERB->show (Subject, \* , Object),*

*SUBJECT->show (\* , Adj1, Adj2,..., AdjN),*

*OBJECT->show (\* , Adj1, Adj2,..., AdjN).*

**Позначення часу в УЖМ.** Для перекладу речення, важливо знати, який час мають основні дієслова. Якщо йдеться про теперішній час – переклад здійснюється без будь-яких уточнень часу. Якщо ж дієслова в минулому чи майбутньому часі, переклад ведеться з додаванням часових жестів, тому що у жестовій мові для позначення дії предмету вживається інфінітив, а також допоміжні дієслова (Verb<sub>a</sub>, auxiliary verb). Наприклад, «читав» ми перекладемо як «читати + був», а «читатиму» як «читати + буду».

Це розглянуто на прикладі речення УСМ «Я подивилася цікавий фільм» (рис. 2). У цьому реченні дієслово знаходиться у минулому часі.

**Правило 2.** Для позначення часу в УЖМ використовуються допоміжні жести «буду» (майбутній), «був», «вже» (минулий), які додаються до основного жесту. Таким чином, уточнені правила для минулого та майбутнього часу будуть мати такий вигляд:

*VERB[FUTURE]->show (Subject, \* , буду, Object),*

*VERB[PAST]->show (Subject, \* , був, Object).*

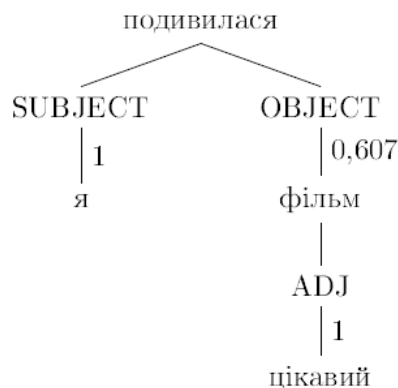


Рисунок 2 – Дерево синтаксичного розбору речення «Я подивилася цікавий фільм»

**Правило №3.** При побудові розповідного стверджувального речення прислівники використовуються після тих слів, до яких вони стосуються. Отже, набір правил для наведеного речення УЖМ буде мати такий вигляд:

*VERB->show (Subject, \* , ADV, Object),*

*VERB[FUTURE]->show (Subject, \* , буду, ADV, Object),*

*VERB[PAST]->show (Subject, \* , був, ADV, Object).*

Розглянемо дерево синтаксичного розбору для речення УСМ «Я нарешті прочитав цей текст» (рис. 3).

**Заперечні речення.** В УЖМ заперечні одиниці «ні», «не буде», «ніколи», «не хочу» тощо, які є основою заперечних жестових конструкцій, застосовуються лише після жестів, які вони заперечують. Розглянемо це на прикладі речення УСМ «Я ніколи не читав цей текст» (рис. 4).

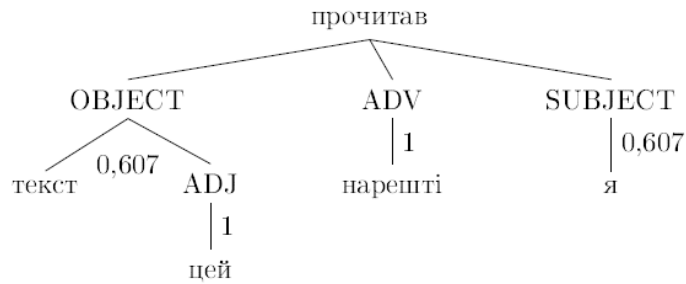


Рисунок 3 – Дерево синтаксичного розбору речення «Я нарешті прочитав цей текст»

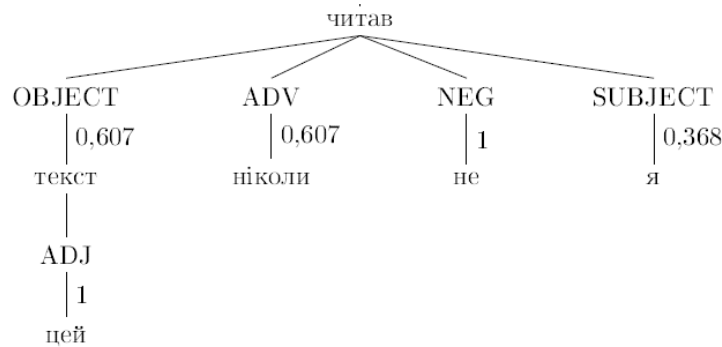


Рисунок 4 – Дерево синтаксичного розбору речення «Я ніколи не читав цей текст»

**Правило №4.** Для відтворення розповідного заперечного речення, заперечні жести «ні», «не буде», «ніколи», «не хочу», використовуються після тих жестів, які вони заперечують. Отже, набір правил для наведеного речення УЖМ буде мати такий вигляд:

- VERB[NEG]->show (Subject, \*, не, ADV, Object),*
- VERB[FUTURE, NEG]->show (Subject, \*, буду, не, ADV, Object),*
- VERB[PAST, NEG]->show (Subject, \*, був, не, ADV, Object).*

**Питальні речення.** В українській жестовій мові для побудови питального речення застосовується таке правило: питальні слова (наприклад, «як», «коли», «чому», «скільки»), завжди ставляться в кінці речення.

Питальні слова можуть належати до різних частин мови, тому для якісного перекладу класифікуємо їх на такі групи: питальні займенники («хто», «що», «який», «чий», «котрий», «скільки», «де», «куди», «звідки», «як», «коли», «доки», «чому», «навіщо» та ін.), питальні частки («чи», «хіба», «невже»)[6]. Крім того, враховано типи питальних речень:

- 1) загальне питання – до всього речення (наприклад, речення «Ти студент?»);
- 2) спеціальні питання – до другорядних членів речення (наприклад, «Де ти працюєш?»).

Для того, щоб описати правила трансформації питального речення з УСМ на УЖМ наведено дерево синтаксичного розбору для речення УСМ «Коли ти купив машину?» (рис. 5):

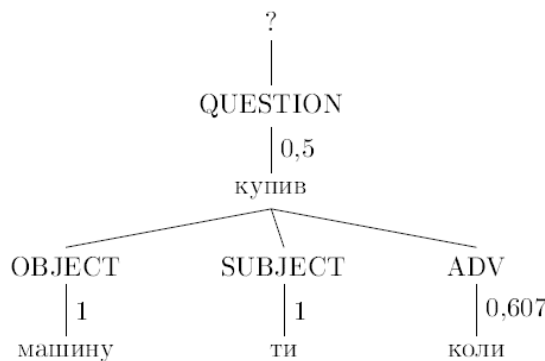


Рисунок 5 – Дерево синтаксичного розбору речення «Коли ти купив машину?»

**Правило №5.** Якщо потрібно відтворити питальне речення на УЖМ, тоді першим відтворюється суб'єкт, потім предикат, об'єкт і питальне слово. Для показу питального речення у жестовій мові використовується питальна міміка обличчя (question face expression, QFE) (у словесній мові застосовується знак питання). Оскільки речення жестовою мовою ми записуємо, то і знак питання будемо використовувати. Отже, набір правил для питального речення УЖМ буде мати такий вигляд:

- питальне слово – займенник (Pron , pronomen):  
**VERB[QUESTION]->show (QFE, Subject, \* , Object, Pron),**  
**VERB[FUTURE, QUESTION]->show (QFE,Subject, \* , буду, Object, Pron),**  
**VERB[PAST, QUESTION]->show (QFE,Subject, \* , був, Object, Pron).**
- питальне слово – частка (Pt, particle):  
**VERB[QUESTION]->show (QFE, Subject, \* , Object, Pt),**  
**VERB[FUTURE, QUESTION]->show (QFE, Subject, \* , буду, Object, Pt),**  
**VERB[PAST, QUESTION]->show (QFE, Subject, \* , був, Object, Pt).**

**Окличні речення.** Як і в українській словесній мові, та й у жестовій, окличні речення можуть виражати: здивування, іронію, розгубленість, сумнів, невпевненість, упевненість; жах, злість, гнів, жаль, ніжність, любов, захоплення, докір, осуд, погрозу, наказ, заклик, прохання та ін. Ці речення вимовляються підвищеним тоном з особливою інтонацією, в кінці яких в УСМ ставиться знак оклику. В українській жестовій мові окличні речення передаються за допомогою виразу обличчя (міміки) чи сплеску руками, тобто будь-яким природним жестом.

Розглянуто синтаксичний розбір окличного речення УСМ «Я нарешті прочитав цей текст!» (рис. 6):

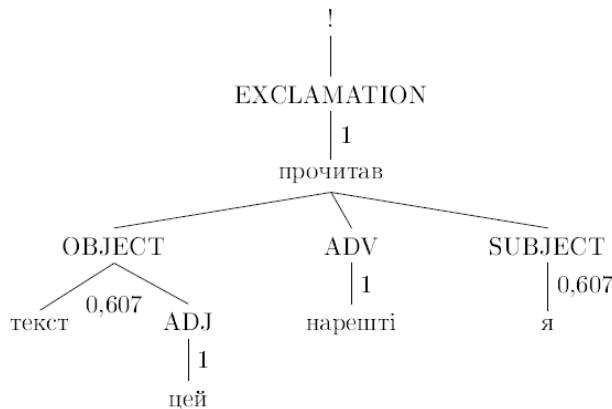


Рисунок 6 – Дерево синтаксичного розбору речення «Я нарешті прочитав цей текст!»

**Правило №6.** Для відтворення окличного речення на УЖМ, першим відтворюється суб'єкт, потім предикат, об'єкт. Для показу окличного речення у жестовій мові використовується міміка обличчя (exclamation face expression, EFE) (у словесній мові застосовується знак оклику). Оскільки речення жестовою мовою ми записуємо, то і знак оклику будемо використовувати. Отже, набір правил для окличного речення УЖМ буде мати такий вигляд:

- VERB[EXCLAMATION]->show (EFE, Subject, \* , ADV, Object),**
- VERB[FUTURE, EXCLAMATION]->show (EFE, Subject, \* , буду, ADV, Object),**
- VERB[PAST, EXCLAMATION]->show (EFE, Subject, \* , був, ADV, Object).**

#### Аналіз результатів

Для тестування перекладу було взято прості речення з корпусу паралельних текстів «Українська словесна мова – Українська жестова мова», який містить 220 речень, 932 слова УСМ і 839 слова УЖМ. Основні характеристики корпусу наведені у таблиці 1.

Таблиця 1 – Статистика корпусу паралельних текстів «Українська словесна мова – Українська жестова мова»

	УСМ	УЖМ
Кількість речень	220	
Кількість слів	932	839
Словник	432	398

Переклад здійснювався на основі правил, виведених за допомогою дерев синтаксичного розбору речень. Для оцінки якості перекладу з української жестової на українську словесну мову використано

оцінки WER (word error rate) – оцінка помилок перекладу, PER (position-independent word error rate) – позиційно-незалежна оцінка помилок. Результати оцінки якості перекладу наведені у таблиці 2.

Таблиця 2 – Оцінка якості перекладу

	WER, %	PER, %
УСМ→УЖМ	42,83	18,35

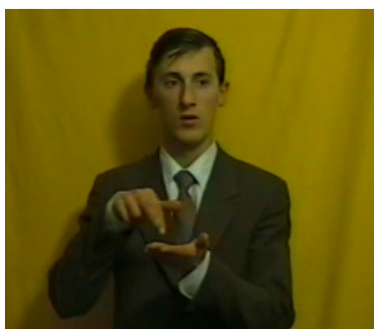
При перекладі з української словесної на українську жестову мову враховувалось, що усі іменники, займенники, прикметники та числівники в українській жестовій мові не мають закінчень, по яких можна ідентифікувати відмінок, рід та число, як у словесній мові. Також в УЖМ відсутні прийменники, а напрямок руху та розміщення жестів вже закладені у самих жестах. В українській словесній мові по закінченнях дієслів визначаються число, особа, рід, час, вид та спосіб, а у жестовій мові всі дієслова є у неозначеній формі дієслова.

Проте, враховуючи вищеописані правила, при перекладі виникали помилки. Нижче наведено кілька прикладів машинного перекладу з УСМ на УЖМ (табл. 3):

Таблиця 3 – Приклади перекладу УСМ→УЖМ

№	Українська словесна мова	Українська жестова мова (машинний переклад)	Українська жестова мова
1.	Дорога йде прямо.	Дорога йти прямо.	Дорога прямо.
2.	Серце б'ється швидко.	Серце битися швидко.	Серцебиття (швидкий темп).
3.	Як тебе звати?	Ти називати як?	Ти ім'я?

У першому, другому та третьому реченнях помилки виникли через те, що одному жесту може відповідати декілька слів. Тому для правильного генерування перекладу треба створити словник таких жестів та їх відповідних перекладів українською словесною мовою. Крім того, відповідно до контексту речення, в УЖМ використовуються різні жести на позначення дії. Наприклад, дієслово «йти» у різних контекстах позначається різними жестами: «я йду» і «час йде» (рис. 7). Для вирішення цього завдання потрібно створити двомовний словник концептів та правил визначення, до якого концепту належить слово у певному контексті для української жестової мови.



а



б

Рисунок 7 – Зображення жестів: а – «я йду»; б – «час йде»

### Висновки

Аналіз одержаних результатів показав доцільність використання машинного перекладу на основі правил для української жестової мови. Зокрема, були застосовані правила обходу дерев для синтезу речень українською жестовою мовою. Експерименти проводились на корпусі паралельних речень «Українська словесна мова – Українська жестова мова», який містить 220 речень. Подальшим дослідженням стане побудова правил перекладу з УСМ на УЖМ, враховуючи граматичні особливості української жестової мови. Результати цієї роботи можна використати під час розроблення програм перекладу з однієї мови на іншу, що полегшить спілкування осіб із вадами слуху між собою і з тими, хто не володіє жестовою мовою.

### Список літератури

1. Краєвський Р. Г. Мова жестів глухих / Р. Г. Краєвський. – К.: Рад. Школа, 1964. – 219 с.
2. Адамюк Н. Б. Синтаксичні особливості УЖМ: на прикладі простого речення / Н.Б. Адамюк, І.І. Чепчина // Жестова мова й сучасність: зб. наук. праць. – Випуск 4. – К.: Педагогічна думка, 2009. – С. 170-191.

3. Bangham J. A. Virtual Signing: Capture, Animation, Storage and Transmission – an Overview of the ViSiCAST Project / J. A. Bangham, S. J. Cox, R. Elliott, J. R. W. Glauert, I. Marshall, S. Rankov, M. Wells // *Speech and Language Processing for Disabled and Elderly People*, 2000.
4. Zhao L. A Machine Translation System from English to American Sign Language / L. Zhao, K. Kipper, W. Schuler, C. Vogler, N. Badler, M. Palmer // *A In Envisioning Machine Translation in the Information Future: Proceedings of the Fourth Conference of the Association for Machine Translation (AMTA-00)*. – Cuernavaca, Mexico. –2000. – pp. 293–300.
5. Speers d’A. L. Representation of American Sign Language for Machine Translation. PhD. dissertation, Graduate School of Arts and Sciences, Georgetown University, 2001.
6. Вихованець І. Р. Граматика української мови. Синтаксис / І. Р. Вихованець. – К., 1993. – 368 с.
7. Загнітко А.П. Основи українського теоретичного синтаксису. Частина 1 / А.П. Загнітко. – Горлівка, 2004. – 227с.
8. Дарчук Н.П. Автоматичний синтаксичний аналіз текстів корпусу української мови / Н.П. Дарчук // *Українське мовознавство*. – КНУ ім. Т. Шевченка, 2013.–№ 43.–С. 11-19.
9. Кривонос Ю.Г. Інформаційна технологія для моделювання української мови жестів / Ю.Г. Кривонос, О.В. Бармак, Ю.В. Крак, А.С. Тернов, Б.А. Троценко // «Штучний інтелект» 3. – 2009. – С. 186-197.
10. Бармак О. Комп’ютерна технологія для перекладу з української на жестову мову / О.Бармак, С. Романишин // *Вісник ТНТУ*. – 2012. – Том 65. – № 1. – С.195-208.
11. Широков В.А. Інформаційна теорія лексикографічних систем [Текст] / В.А. Широков. – К.: Довіра, 1998. – 331с.
12. Програма CMU-парсер [Електронний ресурс]. – Режим доступу: <http://www.link.cs.cmu.edu/link/>
13. Програма LFG-парсер [Електронний ресурс]. – Режим доступу: <http://nlp.ioperm.org/lfg-parser.html>
14. Marshall I. Extraction of semantic representations from syntactic SMU link grammar linkages. In G. Angelova, editor, *Proceedings of Recent Advances in Natural Lanugage Processing*, Tzgov Chark, Bulgaria. – 2001. – pp. 154-159,
15. Shieber S.M. Synchronous tree-adjointing grammars / S.M. Shieber, Y.Schabes // *In Proceedings of the 13th International Conference on Computational Linguistics*, volume 3, Helsinki. – 1990. – pp. 253–258.
16. Zhao L. Machine Translation System from English to American Sign Language / L. Zhao, K. Kipper, W. Schuler, C. Vogler, N. Badler, and M. Palmer // *Association for Machine Translation in the Americas*. – 2000.
17. Morrissey S. An Example-based Approach to Translating Sign Language / S. Morrissey, A. Way // *In Proceedings of the Workshop in Example-Based Machine Translation (MT Summit X)*, Phuket, Thailand. – 2005. – pp. 109–116.
18. San-Segundo R. [Evaluation of Alternatives on Speech to Sign Language Translation](#) / R. San-Segundo, A. Pérez, D. Ortiz, L. F. D’Haro, M. I. Torres, F. Casacuberta // *Proc. of Interspeech 2007*. – Antwerp, Belgium. – 2007. – pp. 2529-2532.
19. Morrissey S. Joining hands: developing a sign language machine translation system with and for the deaf community / S. Morrissey, A. Way // *In Proceedings of the Conference and Workshop on Assistive Technologies for People with Vision and Hearing Impairments - Assistive Technology for All Ages (CVHI-07)*. – Granada, Spain.
20. Stein D. Hand in Hand: Automatic Sign Language to Speech Translation / D. Stein, P. Dreuw, H. Ney, S. Morrissey, A. Way // *In Proceedings of Theoretical and Methodological Issues in Machine Translation (TMI-07)*. – Sweden. – 2007. – pp. 214-220.
21. Bungeroth J. Statistical Sign Language Translation / J. Bungeroth, H. Ney // *Proc. LREC 2004, Workshop proceedings: Representation and Processing of Sign Languages*. – Lisbon, Portugal. – 2004. – pp. 105–108.
22. Проект ViSiCAST [Електронний ресурс]. – Режим доступу: <http://www.visicast.co.uk/>
23. Проект «Communicator» [Електронний ресурс]. – Режим доступу: <http://www.icommunicator.com>.
24. Проект Artemis [Електронний ресурс]. – Режим доступу: <http://www.sim.int-evry.fr/Artemisю>

#### Відомості про авторів

**Лозинська Ольга Володимирівна** – аспірант кафедри інформаційних систем та мереж, Національний університет «Львівська політехніка».

**Давидов Максим Володимирович** – канд. техн. наук, доцент кафедри інформаційних систем та мереж, Національний університет «Львівська політехніка».

**Пасічник Володимир Володимирович** – д-р. техн. наук, Лауреат державної премії України в галузі науки і техніки, професор кафедри інформаційних систем та мереж, Національний університет «Львівська політехніка».

УДК 004.056.5 : 519.728

В. А. ЛУЖЕЦЬКИЙ, Т. М. ЧЕБОРАКА

Вінницький національний технічний університет, м. Вінниця

**МЕТОДИ УЩІЛЬНЕННЯ ДАНИХ НА ОСНОВІ ВІДКИДАННЯ ПОСЛІДОВНОСТЕЙ  
НУЛІВ ТА ОДИНИЦЬ**

**Анотація.** Запропоновано методи ущільнення даних без втрат, що базуються на використанні методів відкидання послідовностей однакових символів у старших, молодших, внутрішніх та старших і молодших розрядах. Характерною особливістю цих методів є розбиття вхідної послідовності даних на блоки та їх подальша обробка на бітовому рівні, внаслідок чого зменшується їх залежність від типу оброблюваних даних. Розроблено програмний засіб для проведення експериментального дослідження запропонованих методів. Наведено результати дослідження ефективності ущільнення при використанні файлів такого типу: \*.doc, \*.txt, \*.bmp, \*.gif, \*.jpg, \*.au, \*.mp3, \*.exe, \*.mdb.

**Ключові слова:** ущільнення даних, послідовності однакових символів.

**Аннотация.** Предложены методы сжатия данных без потерь, основанные на использовании методов отбрасывания последовательностей одинаковых символов в старших, младших, внутренних, старших и младших разрядах. Характерной особенностью этих методов является разбиение входной последовательности данных на блоки и их дальнейшая обработка на битовом уровне, в результате чего уменьшается их зависимость от типа обрабатываемых данных. Разработан программный продукт для проведения экспериментального исследования предложенных методов. Приводятся результаты исследования эффективности сжатия при использовании файлов следующего типа: \*.doc, \*.txt, \*.bmp, \*.gif, \*.jpg, \*.au, \*.mp3, \*.exe, \*.mdb.

**Ключевые слова:** сжатие данных, последовательности одинаковых символов.

**Abstract.** The lossless methods of data compression, based on the truncation of the same characters sequences in low, high, internal and low and high order positions are proposed. A characteristic feature of these methods is a partitioning of input data sequence into blocks and further processing at bit level that leads to decreasing their dependence on the type of data being processed. A software tool for the experimental investigation of the proposed methods is developed. The research results of the effectiveness of data compression with use of the file types \*.doc, \*.txt, \*.bmp, \*.gif, \*.jpg, \*.au, \*.mp3, \*.exe, \*.mdb are presented.

**Keywords:** data compression, sequences of the same characters.

**Вступ**

Обсяги інформації, яка зберігається і передається, з кожним роком значно зростають, що вимагає збільшення обсягів пам'яті і наявності високошвидкісних каналів передавання в сучасних інформаційних системах. Це призводить до постійного збільшення вартості таких систем. Тому економічно вигідним є застосування методів ущільнення для зменшення обсягів інформації, що зберігається і передається.

Існує декілька підходів до ущільнення інформації, що породжують цілу низку методів ущільнення [1, 2], для реалізації яких, у свою чергу, використовується величезна кількість алгоритмів. Одні з них мають науковий характер, а інші знайшли практичну реалізацію у вигляді архіваторів.

**Актуальність**

Теоретичні дослідження і практика застосування архіваторів показали, що не існує універсального методу ущільнення, що забезпечував би однаковий ступінь ущільнення для різних типів даних. Тому наукові дослідження спрямовані на створення ефективних методів ущільнення певних типів даних. Однак дані навіть одного типу, з погляду ущільнення, мають різні властивості і характеристики.

**Мета**

Метою роботи є підвищення значення коефіцієнта ущільнення даних без втрат шляхом створення методів ущільнення на основі відкидання послідовностей однакових символів та засобів, що їх реалізують.

**Задачі**

1. Розробити методи ущільнення даних без втрат на основі відкидання послідовностей нулів та одиниць.
2. Провести експериментальне дослідження ефективності ущільнення запропонованих методів на тестових файлах різного формату та обсягу.
3. Виконати порівняльний аналіз отриманих результатів.

**Метод ущільнення даних на основі відкидання послідовності однакових символів у старших розрядах**

Пропонується метод відкидання послідовностей нулів та одиниць у старших розрядах. Для спрощення викладання матеріалу у подальшому використовуватиметься скорочена назва даного методу – метод СТ.

Суть методу полягає в тому, що підраховується кількість  $q$  однакових символів, що містяться в старших розрядах і залежно від цієї кількості, початковий блок або перетворюється відкиданням послідовності цих однакових символів, або залишається незмінним.

Оскільки після відкидання  $q$  нулів (одиниць) наступним символом завжди буде протилежний символ, то цей символ також можна відкинути. Для того, щоб відновити символи, що відкидаються

необхідно додатково вказати кількість відкинутих символів  $q$ , у вигляді двійкового коду розрядності  $\log_2 n$ , де  $n$  – розрядність початкового блоку.

Якщо виконується умова  $q < \log_2 n + 1$ , то блок даних зберігається без змін, до якого дописується один додатковий символ  $p$  зі значенням 0, – ознакою збереження початкового блоку даних без змін наведено на рис. 1.

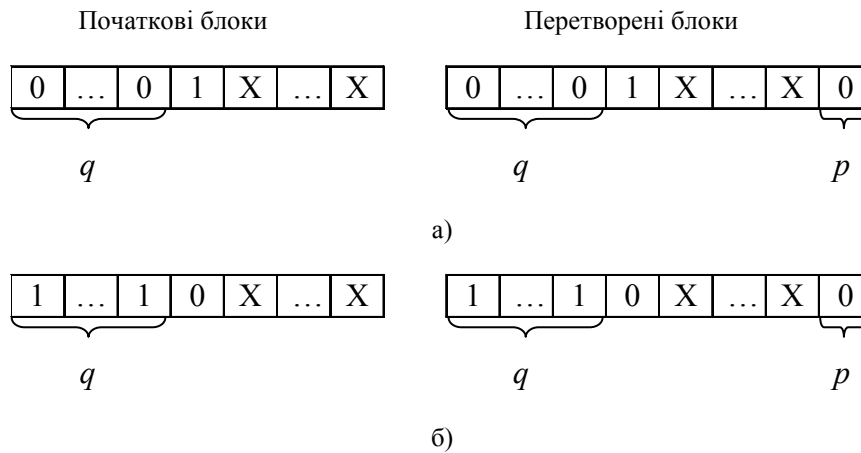


Рисунок 1 – Структури початкових та перетворених за методом СТ блоків (тип 0):  
а)  $q$  нулів; б)  $q$  одиниць

При формуванні блоків типу 0 відбувається збільшення розрядності на один додатковий розряд порівняно з початковою структурою блоків.

При виконанні умови  $q \geq \log_2 n + 1$ , перетворений блок буде мати структуру, що наведена на рис. 2.

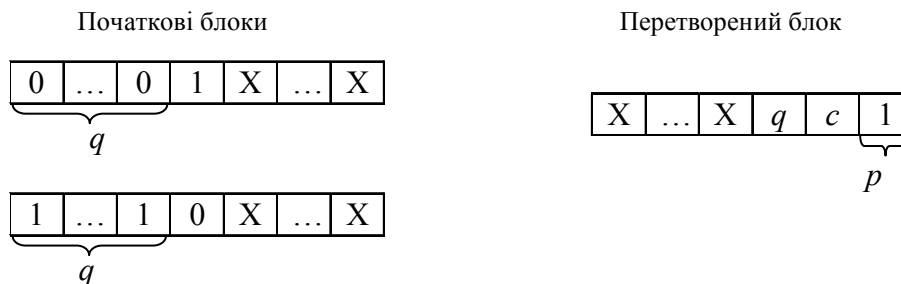


Рисунок 2 – Структури початкових та перетвореного за методом СТ блоку (тип 1)

Тут молодший розряд містить символ 1, що означає перетворення початкового блоку шляхом відкидання послідовності  $q$  однакових символів. Символ 2-го розряду  $c$  – вказує на тип символу, що складає послідовність, яка відкидається. Якщо  $c = 1$ , то це означає, що відкидається послідовність одиниць, а якщо  $c = 0$ , то – послідовність нулів. Поле  $q$  – двійковий код кількості однакових символів послідовності, що відкидається (розрядність коду  $\log_2 n$ ). Поле  $X...X$  – код, що залишається без змін.

Метод ущільнення на основі відкидання послідовностей однакових символів у старших розрядах реалізується шляхом ініціалізації та створенням порожньої множини для значень ущільненої послідовності  $P$ .

Наступним кроком є зчитування блоків вхідної послідовності  $D = \{d_1, d_2, \dots, d_K\}$ .

Далі для кожного блоку підраховується кількість однакових символів, що містяться в старших розрядах. При виконанні умови  $q < \log_2 n + 1$  формуються блоки типу 0 за методом СТ ущільненої послідовності  $p_i^0$ . При виконанні умови  $q \geq \log_2 n + 1$ , формуються блоки типу 1 за методом СТ ущільненої послідовності  $p_i^1$ .

**Метод ущільнення даних на основі відкидання послідовності однакових символів у молодших розрядах**

Алгоритм ущільнення за методом відкидання послідовностей однакових символів у молодших розрядах (метод МР) є аналогічним методу СТ. Відмінність полягає лише у тому, що відбувається підрахунок кількості однакових символів у молодших розрядах блоку, що наведена на рис. 3.

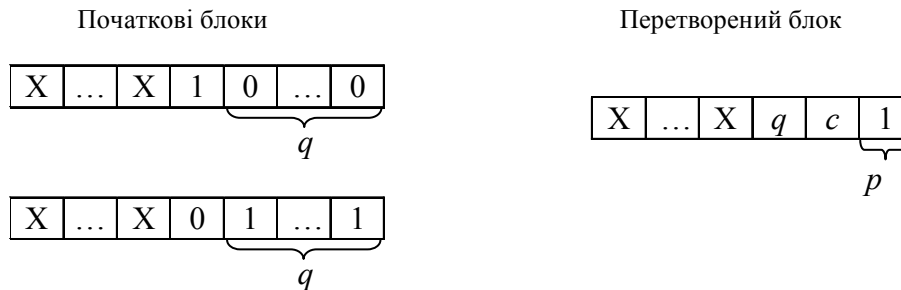


Рисунок 3 – Структури початкових та перетвореного за методом МР блоку (тип 1)

**Метод ущільнення даних на основі відкидання послідовності однакових символів у внутрішніх розрядах**

Пропонується метод відкидання послідовностей нулів та одиниць у внутрішніх розрядах. Для спрощення назви цього методу у подальшому використовуватиметься скорочена назва – метод ВР.

При виконанні умови  $q < 2 \log_2 n$  блок даних зберігається без змін, до якого дописується один додатковий символ  $p$  зі значенням 0, структури блоків наведені на рис. 4.

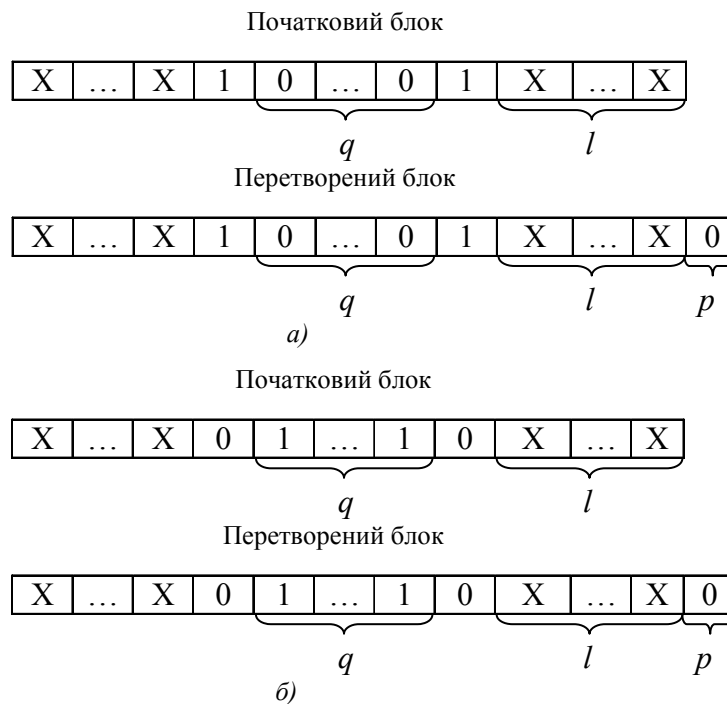


Рисунок 4 – Структури початкових та перетворених за методом ВР блоків (тип 0): а)  $q$  нулів; б)  $q$  одиниць

При виконанні умови  $q \geq 2 \log_2 n$  перетворений блок матиме структуру, що наведена на рис. 5.

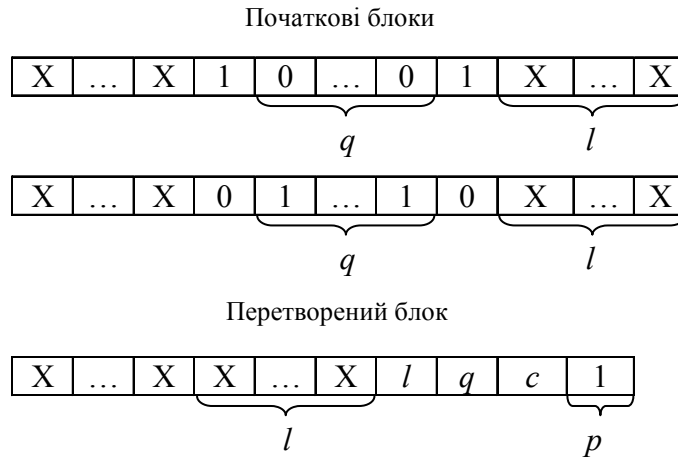


Рисунок 5 – Структури початкових та перетвореного за методом ВР блоків (тип 1)

Ця структура містить:

- $p = 1$  – ознаку перетворення;
- $c$  – тип символу, що відкидається;
- поле  $q$  – двійковий код кількості однакових символів послідовності, що відкидається (розрядність коду  $\log_2 n$ );
- поле  $l$  – двійковий код кількості символів, що залишається без змін у молодших розрядах (розрядність коду  $\log_2 n$ );
- поле  $X...X$  – код, що залишається без змін.

**Метод ущільнення даних на основі відкидання послідовності однакових символів у старших і молодших розрядах**

Пропонується метод відкидання послідовностей нулів та одиниць у молодших і старших розрядах. Для спрощення викладання матеріалу у подальшому використовуватиметься скорочена назва цього методу – метод МСТ.

Суть методу полягає в тому, що підраховуються кількості  $q_l$  і  $q_h$  однакових символів, що містяться в молодших і старших розрядах відповідно, і залежно від цих кількостей початковий блок або перетворюється відкиданням послідовності однакових символів, або залишається незмінним.

Загальна структура початкового блоку, що підлягає ущільненню за методом МСТ наведена на рис. 6.

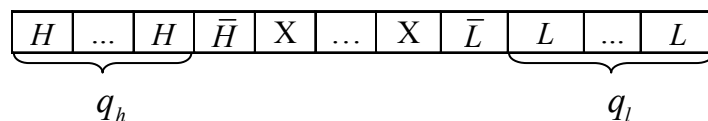


Рисунок 6 – Загальна структура початкового блоку, що підлягає ущільненню за методом МСТ

У наведеній структурі  $q_h$  – кількість однакових символів  $H$ , що відкидаються у старших розрядах,  $q_l$  – кількість однакових символів  $L$ , що відкидаються в молодших розрядах. Символи  $\bar{H}$  та  $\bar{L}$  протилежні символам  $H$  і  $L$  відповідно. Поле  $X...X$  – код, що залишається без змін.

Можливі варіанти структур початкового блоку наведено на рис. 7.

Для кожної структури, в свою чергу, можливі такі співвідношення кількостей однакових символів:

- 1)  $q_h > q_l$ ;
- 2)  $q_h < q_l$ ;
- 3)  $q_h = q_l$ .

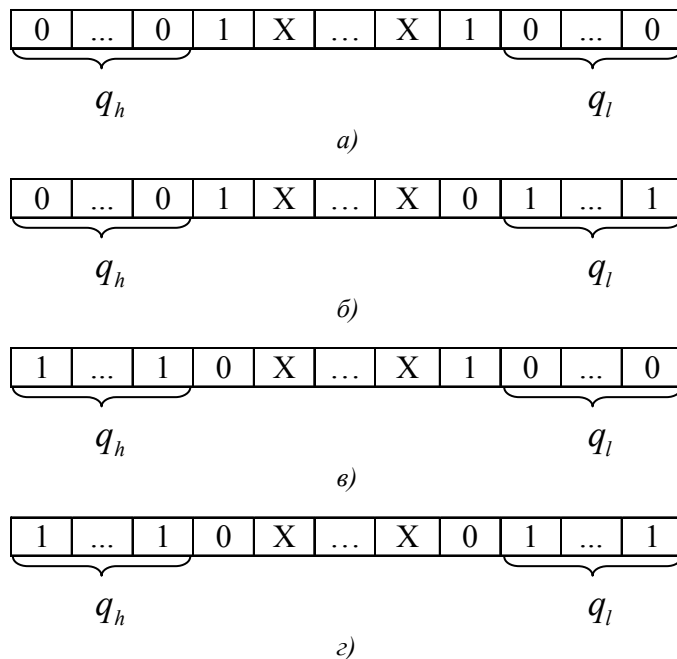


Рисунок 7 – Структури початкових блоків, що підлягають ущільненню за методом МСТ:

- а)  $q_l$  і  $q_h$  нулів у молодших і старших розрядах відповідно; б)  $q_l$  одиниць у молодших і  $q_h$  нулів у старших розрядах; в)  $q_l$  нулів у молодших і  $q_h$  одиниць у старших розрядах; г)  $q_l$  і  $q_h$  одиниць у молодших і старших розрядах відповідно

При виконанні умови  $q_h + q_l \geq ((2 \log_2 n) + 3)$  перетворений блок буде мати структури, що наведені на рис. 8.

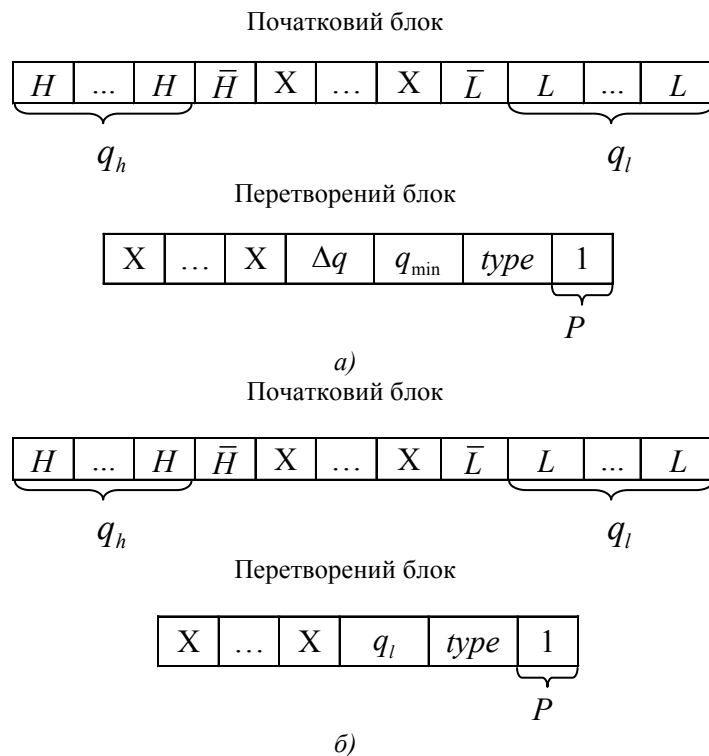


Рисунок 8 – Структури початкових та перетворених блоків, що підлягають ущільненню за методом МСТ:

- а)  $q_h > q_l$  або  $(q_h < q_l)$  ; б)  $q_h = q_l$

Ці структури перетворених блоків містять:

- $p = 1$  – ознаку перетворення;
- поле *type* – 4-розрядний код типу перетворення;
- поле  $q_{\min}$  розрядністю  $\log_2 n$ , в яке записується значення  $\min(q_l, q_h)$ ;
- поле  $\Delta q$  розрядністю  $\log_2 n$ , в яке записується значення  $|q_l - q_h|$ ;
- поле X...X – код, що залишається без змін.

У разі, коли виконуються співвідношення 1) і 2), структура перетвореного блоку має вигляд, наведений на рис. 8, а. Якщо  $q_h = q_l$ , то достатньо вказати одне із значень (нехай це буде  $q_l$ ) і не потрібно вказувати  $\Delta q$ . З урахуванням цього маємо структуру перетвореного блоку, наведену на рис. 8, б.

Кількість можливих типів початкових структур, що підлягають перетворенню дорівнює 12. Коди типу структури (*type*) і відповідні умови формування цих структур наведено в табл. 1.

Таблиця 1 – Позначення правил формування перетворених структур

<i>H</i>	<i>L</i>	Співвідношення $q_h$ і $q_l$	Тип початкового блоку
0	0	$q_h > q_l$	0010
0	0	$q_h < q_l$	0001
0	0	$q_h = q_l$	0011
0	1	$q_h > q_l$	0110
0	1	$q_h < q_l$	0101
0	1	$q_h = q_l$	0111
1	0	$q_h > q_l$	1010
1	0	$q_h < q_l$	1001
1	0	$q_h = q_l$	1011
1	1	$q_h > q_l$	1110
1	1	$q_h < q_l$	1101
1	1	$q_h = q_l$	1111

**Методика дослідження і програмний засіб для проведення досліджень**

Вихідні дані, що підлягають ущільненню, будемо розглядається як послідовність символів 0 і 1. Ця послідовність розбивається на блоки, що містять однакову кількість символів, тобто виконується рівномірна розбивка [3]. Тоді структура вихідних даних матиме вигляд, наведений на рис. 9 [4, 5].

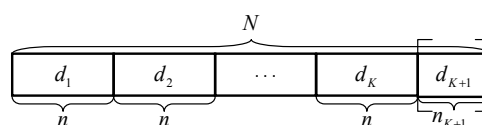


Рисунок 9 – Структура вихідних даних

Тут вихідна послідовність  $D$  із  $N$  символів розбита на блоки  $d_1, d_2, \dots, d_K [d_{K+1}]$  по  $n$  символів, де  $n$  кратне байту:  $D = \{d_1, d_2, \dots, d_K [d_{K+1}]\}$ .

Кількість блоків розрядністю  $n$  дорівнює  $K = \lfloor N/n \rfloor$ , де  $\lfloor \cdot \rfloor$  означає округлення до меншого цілого. Якщо  $N$  ділиться на  $n$  точно, то всі блоки будуть мати однакову довжину  $n$ , і блок  $d_{K+1}$  буде відсутнім. Якщо результат ділення є нецілим числом, то блок  $d_{K+1}$  буде мати довжину  $n_{K+1} = N \bmod n$ .

Для того, щоб сформувати ущільнену послідовність  $P$ , необхідно поставити у відповідність кожному початковому блоку перетворений блок  $d_i \rightarrow p_i$ . Результатом перетворення вхідного потоку даних є ущільнена послідовність:  $P = \{p_1, p_2, \dots, p_K [p_{K+1}]\}$ .

Над блоком  $d_{K+1}$  перетворення не виконуються і він залишається без змін –  $p_{K+1} = d_{K+1}$ .

Для автоматизації експериментального дослідження запропонованих методів ущільнення на тестових файлах розроблено програмний засіб.

Під час проведення дослідження розрядність початкового блоку даних може приймати значення в діапазоні: 8, 16, 32, ..., 8192. Дослідження відбувалося за такими показниками: коефіцієнт ущільнення  $k$ , тривалість ущільнення  $t_c$ , тривалість відновлення  $t_{dc}$ .

Результати дослідження запропонованих методів ущільнення за показниками  $k$ ,  $t_c$ ,  $t_{dc}$  виводяться у табличному та графічному вигляді (рис. 10).

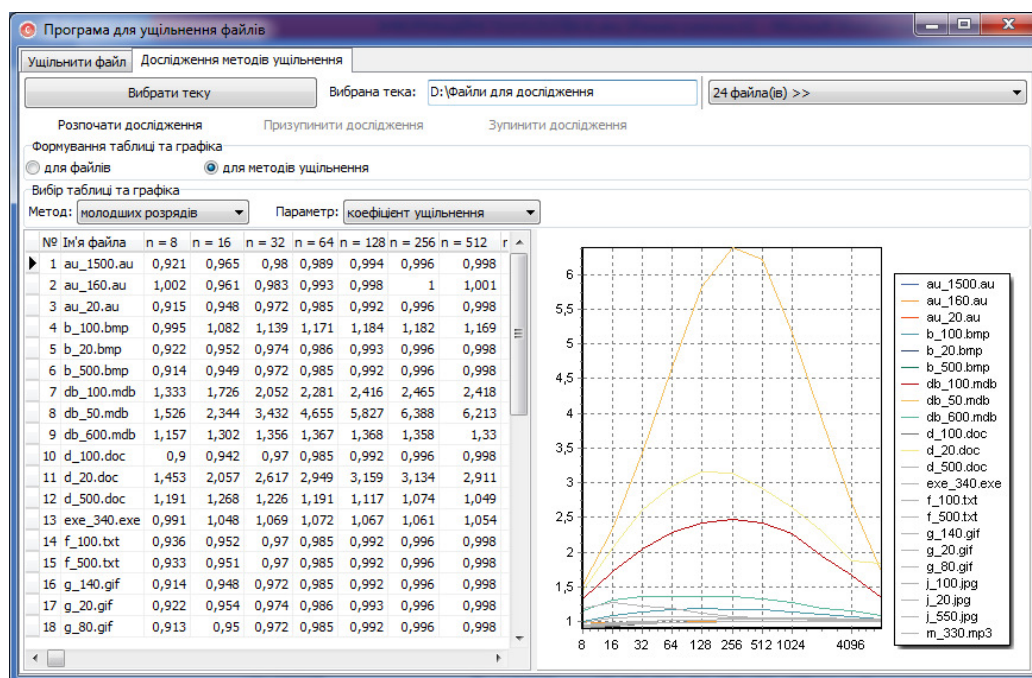


Рисунок 10 – Головне вікно програмного засобу

У процесі ущільнення вхідного файлу формується вихідний файл, що містить усю необхідну інформацію для відновлення початкового файлу.

Для проведення експериментальних досліджень запропонованих методів ущільнення було сформовано тестову вибірку із 24 файлів таких форматів та розмірів:

- текстового формату \*.doc – 20 кБ, 100 кБ, 500 кБ;
- текстового формату \*.txt – 100 кБ, 500 кБ;
- графічного формату \*.bmp – 20 кБ, 100 кБ, 500 кБ;
- графічного формату \*.gif – 20 кБ, 80 кБ, 140 кБ;
- графічного формату \*.jpg – 20 кБ, 100 кБ, 550 кБ;
- музичного формату \*.au – 20 кБ, 160 кБ, 1500 кБ;

- музичного формату \*.mp3 – 40 кБ, 330 кБ, 500 кБ;
- формату виконуваних файлів \*.exe – 340 кБ;
- формату бази даних \*.mdb – 50 кБ, 100 кБ, 600 кБ.

Під час експериментальних досліджень вміст вихідних файлів розбивався на блоки розрядності  $n = 8, 16, \dots, 2048$ .

Результати досліджень коефіцієнта ущільнення наведено у табл. 2 та представлено у вигляді графіка на рис. 11. У таблиці для кожного значення коефіцієнта ущільнення також зазначено в дужках розрядність блоків, при якій було досягнуто цей коефіцієнт.

Таблиця 2 – Результати дослідження коефіцієнта ущільнення

Формат та обсяг файлу	Метод ущільнення			
	MP	CT	BP	MCT
au, 160 кБ	1,002 (8)	1,187 (16)	1,012 (32)	1,007(16)
bmp, 100 кБ	1,184 (128)	1,181 (128)	1,174 (128)	1,198(256)
mdb, 100 кБ	2,465 (256)	2,484 (256)	2,485 (256)	2,651(1024)
mdb, 50 кБ	6,388 (256)	7,23 (512)	6,468 (256)	8,388(1024)
mdb, 600 кБ	1,368 (128)	1,374 (64)	1,414 (128)	1,386(256)
doc, 20 кБ	3,159 (128)	3,327 (128)	3,262 (256)	3,389(256)
doc, 500 кБ	1,268 (16)	1,43 (16)	1,25 (16)	1,196(16)
exe, 340 кБ	1,072 (64)	1,082 (32)	1,081 (512)	1,082(64)
mp3, 40 кБ	1,033 (1024)	1,031 (256)	1,037 (512)	1,034(2048)

З табл. 2 та рис. 11 видно, що для більшості файлів, що піддаються ущільненню запропонованими методами, найкращі результати демонструє метод ущільнення – MCT. Лише для файлів (au, 160 кБ) та (doc, 500 кБ) метод CT в незначній мірі дав кращі результати, ніж метод відкидання у старших і молодших розрядах: 1,187 проти 1,007 і 1,43 проти 1,196 відповідно.

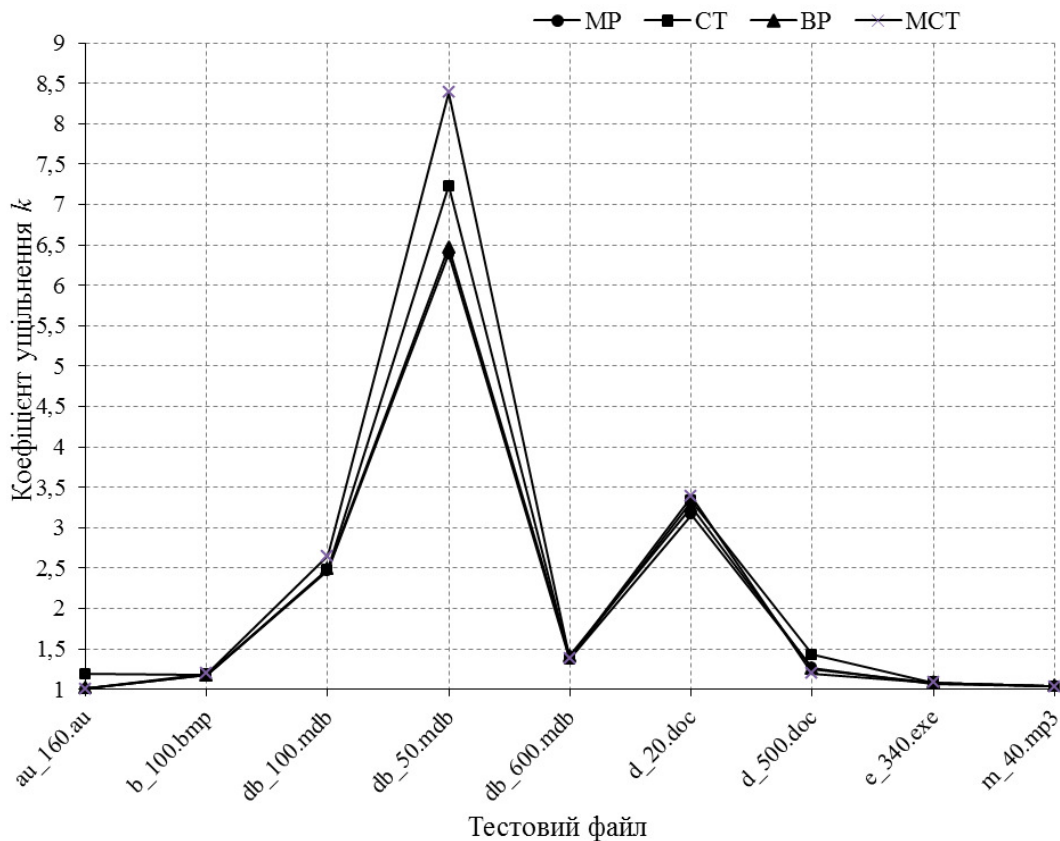


Рисунок 11 – Графіки результатів дослідження коефіцієнта ущільнення

Таким чином, проведені експериментальні дослідження запропонованих методів ущільнення показали, що ці методи дозволяють ущільнювати файли різних форматів та розмірів. В результаті досліджень найбільший коефіцієнт ущільнення із запропонованих методів продемонстрував метод відкидання послідовностей однакових символів у старших і молодших розрядах, який дав максимальний приріст значення коефіцієнта ущільнення у 16,01% порівняно із найкращими значеннями інших методів.

Щодо тривалості ущільнення та відновлення файлів запропонованими методами, то експериментальне дослідження на комп'ютері типу IBM/PC з процесором Intel Core i3 M370 (2.4GHz) і ОЗП розміром у 3 ГБ показало, що для найбільшого файлу (au, 1500 кБ) процедура ущільнення найдовше тривала при найменшій розрядності блоків: від 0,53 с. для методу МР та методу СТ до 0,5 с., – а процедура відновлення – 0,46 с. для методу ВР та 0,39 с. для методу МСТ.

#### Висновки

1. Запропоновано методи ущільнення даних на основі відкидання послідовностей нулів та одиниць у молодших, старших, внутрішніх та старших і молодших розрядах. Ці методи на відміну від існуючих обробляють вихідну послідовність даних на бітовому рівні, що дозволяє зменшити їх залежність від типу даних.

2. Проведено експериментальне дослідження коефіцієнта ущільнення запропонованих методів на тестових файлах різного формату та обсягу за допомогою розробленого програмного засобу. Це дослідження показало, що запропоновані методи ущільнення дозволяють ущільнювати файли різних форматів та розмірів.

3. Найбільший коефіцієнт ущільнення із запропонованих методів продемонстрував метод відкидання послідовностей однакових символів у старших і молодших розрядах, який дав максимальний приріст значення коефіцієнта ущільнення у 16,01% порівняно із найкращими значеннями інших методів.

#### Список використаних джерел

1. Ватолин Д. Методы сжатия данных. / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин – М.: ДИАЛОГ-МИФИ, 2002. – 384 с.
2. Salomon D. Handbook of Data Compression / D. Salomon, G. Motta. – London: Springer, 2010. – 1361 p.
3. Лужецький, В. А. Узагальнена модель адаптивного ущільнення даних / В. А. Лужецький, Л. А. Савицька, Ш. А. Хок // Інформаційні технології та комп'ютерна інженерія. – 2009. – № 1. – С. 56-63.
4. Лужецький В. А. Дослідження методу ущільнення даних на основі методу відкидання послідовностей нулів та одиниць / В. А. Лужецький, Т. М. Алексеева // Методи та засоби кодування, захисту й ущільнення інформації: міжнар. наук.-практ. конф., 20–22 квітня 2011 р.: тези доп., – Вінниця, 2011. – С.164-165.
5. Лужецький В. А. Дослідження числових моделей даних / В. А. Лужецький, А. В. Кульчицький, Т. М. Алексеева // Інформаційні технології та комп'ютерна інженерія. – 2010. – № 3. – С. 50–56.

#### Інформація про авторів

**Лужецький Володимир Андрійович** – д.т.н., професор, завідувач кафедри захисту інформації, Вінницький національний технічний університет, вул. Хмельницьке шосе 95, м. Вінниця.

**Чеборака Тетяна Михайлівна** – аспірант кафедри захисту інформації, Вінницький національний технічний університет, вул. Хмельницьке шосе 95, м. Вінниця.

УДК 004.42

А.М. ПЕТУХ, О.М. РЕЙДА, В.В. ВОЙТКО

Вінницький національний технічний університет, м. Вінниця

**МОДЕЛЬ ІННОВАЦІЙНОЇ СТРУКТУРИ УНІВЕРСИТЕТУ**

**Анотація.** В статті розглянуто особливості впровадження інноваційної складової в навчальний процес вищого навчального закладу. Розроблено модель інноваційної структури університету та розглянуто особливості створення комерціалізованого простору для розвитку і забезпечення підприємницької та інноваційної діяльності студентів і молодих науковців університету.

**Ключові слова:** інноваційна діяльність, модель інноваційної структури університету.

**Анотация.** В статье рассмотрены особенности внедрения инновационной составляющей в учебный процесс вуза. Разработана модель инновационной структуры университета и рассмотрены особенности создания коммерциализированного пространства для развития и обеспечения предпринимательской и инновационной деятельности студентов и молодых ученых университета.

**Ключевые слова:** инновационная деятельность, модель инновационной структуры университета.

**Abstract.** In article features the introduction of innovative component in the educational process of higher education. The model of the innovative structure of the university and the features of commercialized create space for the development and maintenance of enterprise and innovation activities of students and young scientists of the University.

**Keywords:** innovation, innovation model university structure.

**Вступ**

Сьогодні потреба економічного зростання на сучасному етапі розвитку суспільства обумовлює необхідність забезпечення умов та створення відповідних механізмів, спрямованих на стимулювання інноваційної діяльності ще на етапі отримання знань і практичних навичок у вищій школі. Такий підхід спрямований на підтримку розвитку економіки країни інноваційним шляхом. Закон України «Про інноваційну діяльність» визначає коло правових, економічних та організаційних заходів, підпорядкованих загальнодержавному регулюванню інноваційної діяльності в Україні [1].

**Актуальність**

Правова підтримка інноваційних процесів вищої школи визначена законами України «Про наукову і науково-технічну діяльність», «Про інноваційну діяльність» та «Про пріоритетні напрями інноваційної діяльності в Україні» [1-3]. Активне впровадження інноваційної складової в навчальні плани вищої школи забезпечить комерціалізацію результатів наукових досліджень та підсилить практичне значення освітянської діяльності [4-7]. У зв'язку з цим актуальною є розробка та реалізація моделей інноваційної структури вищих навчальних закладів з метою подальшого впровадження результатів інноваційної діяльності у господарську практику.

**Мета**

Метою роботи є комерціалізація навчального процесу вищої школи шляхом впровадження інноваційної складової в освітянську діяльність вищих навчальних закладів. Під об'єктом дослідження розуміємо інноваційні технології та інноваційні процеси в системі освіти. Предметом дослідження постають моделі інноваційної структури вищого навчального закладу.

**Задачі**

Основними задачами є розробка моделі інноваційної структури університету та створення простору для розвитку і забезпечення підприємницької та інноваційної діяльності студентів і молодих науковців університету.

**Визначення сфери діяльності моделі інноваційної структури університету**

Сферою діяльності моделі є організація, розвиток і забезпечення підприємницької та інноваційної діяльності студентів та молодих науковців університету.

Основними результатами роботи моделі вбачаємо:

- фінансування і виконання інноваційних робіт;
- комерціалізацію, організацію та забезпечення виробництва продукції на власних університетських або дочірніх виробничих потужностях;
- координацію наукової, інноваційної та комерційної діяльності;
- матеріально-технічне забезпечення інноваційної діяльності студентів і молодих науковців;
- отримання консультацій учасниками інноваційних робіт;
- доступ до ресурсів інноваційної структури;
- участь у виставках, презентаціях, аукціонах, семінарах, симпозиумах, конкурсах, які організує базовий вищий навчальний заклад.

Модель орієнтована на забезпечення можливості підвищення ефективності інноваційної діяльності в університеті, спрощення та прискорення процесів інноваційної діяльності.

**Визначення завдань та переваг у використанні моделі для кожного з її об'єктів**

Під об'єктами моделі розуміємо сам університет (У), інноваційну структуру університету (ІСУ) та ІТ-компанії (ІТК), з якими співпрацює університет у процесі здійснення інноваційної діяльності.

Сформулюємо головну мету інноваційної діяльності університету як об'єкту моделі:

- підвищення ефективності та якості процесів інноваційної діяльності в університеті;
- підтримка і розвиток інноваційної діяльності студентів і співробітників університету.

Головними завданнями університету в цьому ключі постає:

- визначення напрямків діяльності для підвищення ефективності та якості процесів з інноваційною складовою;
- розробка методик підтримки і розвитку інноваційної діяльності.

До загальних переваг функціонування університету як об'єкту моделі інноваційної структури відносимо:

- підвищення ефективності та якості процесів інноваційної діяльності;
- підвищення кількості інноваційних продуктів та послуг;
- підвищення регіонального, національного та міжнародного рейтингів університету.

Головна мета впровадження інноваційної структури університету полягає в:

- підвищенні якості підготовки учасників інноваційних проектів;
- координації наукової, інноваційної та комерційної діяльності;
- комерціалізації, організації та забезпеченні реалізації продуктів інноваційної діяльності.

Головними завданнями вбачаємо:

- визначення напрямку та методів підвищення якості підготовки учасників інноваційних проектів;
- визначення методів координації наукової, інноваційної та комерційної діяльності;
- розробку методик для комерціалізації, організації та забезпечення реалізації продуктів інноваційної діяльності.

Серед переваг впровадження інноваційної структури університету виділяємо:

- індивідуальну цільову підготовку учасників проекту;
- участь співробітників університету в робочих групах по розробці інноваційних проектів;
- комерціалізацію, організацію та забезпечення виробництва продукції на власних або дочірніх виробничих потужностях;
- залучення партнерів з метою реалізації спільних проектів;
- матеріально-технічне забезпечення інноваційної діяльності студентів та молодих науковців.

Головна мета діяльності ІТ-компанії як об'єкту моделі полягає у використанні інноваційних продуктів і послуг для підвищення ефективності виробничих процесів і якості продукції.

Головним завданням ІТ-компанії в рамках інноваційної освітньої діяльності постає необхідність визначення інноваційних продуктів і послуг для підвищення ефективності виробничих процесів і якості вихідної продукції.

Переваги інноваційної діяльності ІТ-компанії полягають у:

- створенні якісних інноваційних продуктів;
- отриманні додаткових прибутків.

### **Структура моделі**

На рис. 1 наведено модель інноваційної структури університету. Для реалізації моделі необхідно створити ІСУ. Інноваційна структура університету створюється як структурна одиниця університету шляхом об'єднання, реорганізації існуючих підрозділів, що мають відповідний профіль, або створення нового структурного підрозділу для організації, координації, контролю процесу розробки і виконання проектів. Крім того, слід розробити нормативно-правову базу для функціонування ІСУ, розробити та реалізувати систему взаємодії об'єктів моделі: ІСУ, У та ІТК.

Інноваційна структура університету створюється як структурна одиниця університету шляхом об'єднання, реорганізації існуючих підрозділів, що мають відповідний профіль, або створення нового структурного підрозділу для організації, координації, контролю процесу розробки і виконання проектів.

Метою ІСУ є розвиток науково-технічної та інноваційної діяльності для ефективного та раціонального використання наявного наукового потенціалу, матеріально-технічної бази, комерціалізації результатів наукових досліджень і їх впровадження.

ІСУ визначає пріоритетні напрями інноваційної діяльності, спрямовує свою діяльність на підвищення якості підготовки учасників інноваційних проектів; координує наукову, інноваційну та комерційну діяльність; виконує комерціалізацію та забезпечує реалізацію продуктів інноваційної діяльності.

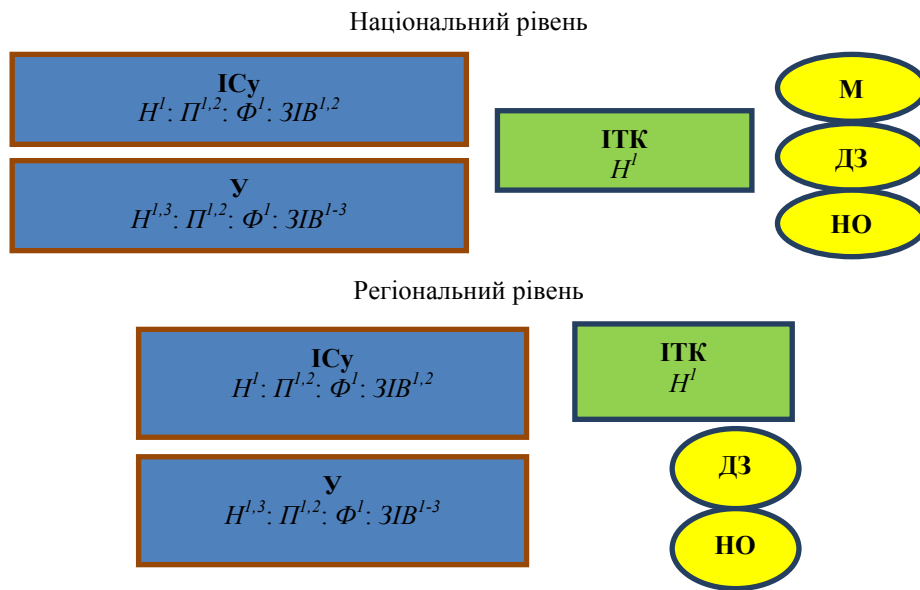


Рисунок 1 – Модель інноваційної структури університету

#### Огляд моделі

Модель безпосередньо пов'язана з діяльністю і послугами університету. Розглянемо зв'язки інноваційної структури як складової частини університету.

Для реалізації моделі обов'язковими об'єктами стають:

1. Університет з реалізованими процесами -  $H^1, H^2, H^3, P^1, P^2, ZIB^1, ZIB^2, ZIB^3$ .
2. Інноваційна структура університету –  $H^1, P^1, P^2, ZIB^1, ZIB^2$ .
3. IT- Компанія –  $P^1$ .

Процес  $P^1$  пов'язує об'єкти моделі, що дозволяє проводити аналіз зовнішнього інноваційного середовища з метою визначення необхідних напрямків інноваційної діяльності.  $P^1$  – це організаційна підтримка інноваційної діяльності в У, ІСУ. Об'єкт ІТК проводить організаційні заходи щодо допомоги у визначенні пріоритетних напрямків інноваційної діяльності ІСУ.

Об'єкт У реалізує такі процеси:

$H^1$ - студент має навчитися підходам для створення принципово нових продуктів та сервісів, шукати та аналізувати інформацію щодо пріоритетних напрямків в інформаційних технологіях та науці, вмінню висловлювати, описувати та презентувати власні ідеї, відрізнити хибні твердження та пропозиції;

$H^2$  – студент має розуміти, які види підприємств можуть бути створені для реалізації інновації, у чому полягає сам процес інвестування, як самостійно розробити бізнес план інновації;

$H^3$  – студент має навчитися шукати джерела фінансування інноваційного проекту, подавати інформацію інвесторам;

$P^1$  студент має отримати організаційну підтримку;

$P^2$  студент має отримати економічну підтримку;

$ZIB^1$  – студент має отримати допомогу щодо реєстрації ним створеної інтелектуальної власності як на державному, так і на міжнародних рівнях.

$ZIB^2$  - студент має отримати допомогу щодо впровадження вже зареєстрованої інтелектуальної власності.

$ZIB^3$  - студент має отримати заохочувальні та мотиваційні заходи.

Об'єкт ІСУ має реалізувати процеси  $H^1, P^1, P^2, ZIB^1, ZIB^2$ .

Послідовність дій у моделі описується алгоритмом:

1. Університет має власну інноваційну структуру, яка реалізує процеси  $H^1, P^1, P^2, ZIB^1, ZIB^2$ .
2. Студенти, аспіранти, співробітники університету визначають напрямок інноваційної діяльності. ІСУ проводить консультації щодо актуальності інноваційного напрямку, формується робоча група.
3. На наступному етапі відбувається організаційна та економічна підтримка діяльності групи розробників.
4. Група розробників реалізує інноваційний проект, отримавши підтримку ІСУ в реалізації результатів проекту.
5. Готовий інноваційний проект передається ІСУ.

6. ІСУ проводить реєстрацію створеної інтелектуальної власності як на державному, так і на міжнародних рівнях і впроваджує, за необхідності, результати зареєстрованої інтелектуальної власності.

#### **Висновок**

Запропонована модель інноваційної структури університету призначена для розвитку науково-технічної та інноваційної діяльності вищого навчального закладу, ефективного та раціонального використання наявного наукового потенціалу, матеріально-технічної бази за рахунок комерціалізації результатів наукових досліджень і їх впровадження у господарську практику.

#### **Список літератури**

1. Закон України «Про інноваційну діяльність» // Відомості Верховної Ради України, 2002, N36. — 266с.
2. Закон України «Про наукову і науково-технічну діяльність» // Відомості Верховної Ради України (ВВР), 1992, №12. — С. 165.
3. Закон України «Про пріоритетні напрями інноваційної діяльності в Україні» // Цінні відомості Верховної Ради (ВВР), 2003, №13. — С. 160.
4. Закон України Про авторське право і суміжні права // Відомості Верховної Ради України (ВВР), 1994, №13. — С. 64.
5. Ионас В. Я. Произведения творчества в гражданском праве. / Ионас В. Я. — М.: Юридическая литература, 1972. — 168 с.
6. Попович О. С. Науково-технологічна та інноваційна політика: основні механізми формування та реалізації / Попович О. С. — К.: Фенікс, 2005. — 226 с.
7. Липчик Д. Авторское право и смежные права / Липчик Д.— М.: Радомир; Издательство ЮНЕСКО, 2002. — 788 с.

#### **Відомості про авторів**

**Петух Анатолій Михайлович** – д.т.н., професор, Україна, м. Вінниця, Вінницький національний технічний університет, кафедра програмного забезпечення.

**Рейда Олександр Миколайович** – к.т.н., доцент, Україна, м. Вінниця, Вінницький національний технічний університет, кафедра програмного забезпечення.

**Войтко Вікторія Володимирівна** – к.т.н., доцент, Україна, м. Вінниця, Вінницький національний технічний університет, кафедра програмного забезпечення.

УДК 621.39

О. М. ТКАЧЕНКО, О. Ф. ГРІЙО ТУКАЛО

Вінницький національний технічний університет, м. Вінниця

### ПІДХІД ДО ОЦІНЮВАННЯ ТРИВАЛОСТІ ФРАГМЕНТА ДЛЯ ПОШУКУ МУЗИЧНОГО ТВОРУ ЗА ЗАДАНИМ ШАБЛОНОМ

**Анотація.** Стаття належить до області інформаційних технологій, зокрема ідентифікації музичного твору на основі аудіо контенту. Теоретично обґрунтовано можливість ідентифікації музичного твору за його фрагментом. Застосовано кластерний аналіз під час формування шаблонів музичних творів в БД, що дозволяє зменшити обсяги пам'яті для їх зберігання. Запропоновано критерій порівняння фрагменту музичного твору з шаблонами БД. Визначено мінімальну тривалість фрагменту, що дозволяє суттєво зменшити складність обчислень в процесі ідентифікації музичного твору. Експериментальні результати підтвердили коректність теоретичних положень.

**Ключові слова:** ідентифікація за фрагментом аудіо запису, параметризація, мел-частотні кепстральні коефіцієнти, кластерний аналіз, Евклідова відстань.

**Аннотация.** Статья относится к области информационных технологий, в частности идентификации музыкального произведения по аудио контенту. Теоретически обоснована возможность идентификации музыкального произведения по его фрагменту. Применен кластерный анализ при формировании шаблонов музыкальных произведений в БД, что позволяет уменьшить объемы памяти для их хранения. Предложен критерий сравнения фрагмента музыкального произведения с шаблонами БД. Определена минимальная продолжительность фрагмента, что позволяет существенно уменьшить сложность вычислений в процессе идентификации музыкального произведения. Экспериментальные результаты подтвердили корректность теоретических положений.

**Ключевые слова:** идентификация по фрагменту аудио записи, параметризация, мел-частотные кепстральные коэффициенты, кластерный анализ, Евклидово расстояние.

**Abstract.** Article relates to the field of information technology, in particular the content-based identification of songs. It is theoretically justified the ability to identify a piece of music by its fragment. Cluster analysis is applied when forming templates of music in the database, which reduces the amount of memory to store them. We propose the criterion for comparing the song fragment with templates database. The minimal length of the fragment is determined, which significantly reduces the computational complexity in the process of the song identifying. The experimental results confirm the correctness of theoretical positions.

**Keywords:** audio identification by fragment, parameterization, mel-frequency cepstral coefficients, cluster analysis, Euclidean distance.

#### Вступ

Дана стаття належить до області інформаційних технологій, зокрема автоматичної ідентифікації музичного твору за фрагментом аудіо запису.

В сучасних комп'ютерних мережах основний об'єм трафіку припадає на мультимедійну, зокрема, аудіо інформацію. Зростання обсягу мультимедійної інформації, що передається і обробляється в комп'ютерних системах, зумовила необхідність автоматизації процесів аналізу і пошуку даних. Більшість сучасних систем пошуку аудіо інформації використовуює метадані (текстові описи про виконавця, назву музичного твору, рік запису тощо) та текстові анотації аудіо контенту [1]. Недоліком пошуку виключно на основі метаданих є те, що користувач пошукової системи повинен мати досить чітке уявлення про зміст того, що він шукає. Тому метадані доповнюють анотаціями (тегами) змісту аудіо запису [2]. Генерація таких анотацій аудіо контенту є трудомістким і тривалим процесом [1, 3]. Таким чином, в сучасних системах обробки аудіо інформації виникає необхідність автоматичного швидкого пошуку музичних творів на основі аудіо контенту у віддалених базах даних (БД) великого розміру на сервері. Виходячи з вищесказаного, суть пошуку музичного твору на основі аудіо контенту полягає в тому, щоб автоматично отримувати файли аудіо записів музичних творів, подібних до заданого аудіо запису під час запиту. При цьому, враховуючи великі обсяги аудіо інформації в БД, велике значення має швидкість пошуку. В зв'язку з цим в даній статті йде мова про можливість ідентифікації музичного твору саме за коротким фрагментом. Важливо, щоб тривалість фрагменту була якомога меншою, оскільки це дозволить: 1. збільшити швидкість пошуку; 2. зменшити час завантаження та мережевий трафік. Разом з тим зменшення тривалості фрагменту може зумовити зростання ймовірності помилки під час ідентифікації музичного твору. Тому головне питання, на яке потрібно дати відповідь: яка мінімальна тривалість фрагменту є достатньою, щоб характеризувати весь музичний твір.

#### Мета та задачі статті

Метою даної статті є теоретичне обґрунтування можливості ідентифікації музичного твору за його фрагментом та мінімальної тривалості фрагменту, що дозволяє зменшити складність обчислень в процесі автоматизованої ідентифікації музичного твору.

Для реалізації системи ідентифікації музичного твору необхідно розв'язати такі задачі: 1) обрати параметри, які дозволили б однозначно та компактно описати музичний твір; 2) створити швидкий та надійний метод порівняння за обраними параметрами вхідного фрагменту музичного твору та попередньо створених еталонів (шаблонів), що зберігаються в базі даних; 3) обґрунтувати мінімальну тривалість фрагменту, який дозволяє ідентифікувати музичний твір; 4) провести експериментальну

перевірку запропонованих теоретичних положень. Результатом розпізнавання буде шаблон БД з мінімальним розходженням відносно вхідного аудіо запису.

Далі в роботі вважається, що аудіо запис, який треба ідентифікувати, точно збігається з одним із записів, що містяться в БД. Проте всі зроблені надалі висновки залишаються справедливими і в тому випадку, коли вхідний аудіо запис та відповідний йому шаблон БД не є ідентичними в силу таких факторів як наявність шумів, зміна темпу, частоти дискретизації тощо.

### Математична модель аудіо сигналу (MFCC)

Одним з центральних понять інформаційного пошуку музики (MIR) є схожість/подібність музичних творів. Моделювання подібності музики – це основа програм для автоматичної організації та обробки баз даних музики. Схема визначення відповідності музичних творів на основі контенту базується на використанні аудіо файлу для побудови моделі аудіо сигналу.

Порівнювати безпосередньо звукові сигнали в часовій області — довго і не дуже ефективно, тому відліки аудіо сигналу ділять на невеликі фрагменти (фрейми), для яких характеристики сигналу залишаються відносно стійкими (стаціонарний випадковий процес), з перекриттям фреймів. Тривалість одного фрейму, як правило, лежить у межах 10 – 30 мс. Для кожного фрейму виконується спектральний аналіз, на основі якого (тим чи іншим чином) обчислюється значення вектора параметрів (параметризація). Обрані параметри повинні мати такі властивості:

1. мінімізація обсягу інформації, необхідного для опису аудіо запису (за рахунок логарифмування);
2. некорельованість параметрів (за рахунок DCT);
3. однорідність параметрів, тобто однакова дисперсія в середньому;
4. можливість застосовувати прості метрики (Евклідова метрика) для визначення близькості між наборами параметрів;

Багато різних параметрів запропоновано в літературі [4, 5]. Зазначеним вище властивостям найкраще відповідають мел-частотні кепстральні коефіцієнти (MFCC – Mel Frequency Cepstral Coefficients), які вперше було запропоновано як параметри в системах розпізнавання мовлення та диктора [6], а в подальшому отримали широке використання в процесі інформаційного пошуку музики (MIR) [7, 8], зокрема, під час класифікації за жанрами, визначенні подібності аудіо тощо.

Обравши MFCC як параметри, ми отримуємо опис музичного твору у вигляді файлу з параметрами MFCC. Таким чином, якщо тривалість музичного твору в середньому складає 3хв (180с), частота дискретизації аудіо файлу – 44,1кГц, довжина фрейму – 20мс з перекриттям 0,5 фрейма, то маємо:

$$\begin{aligned} 44100 \text{ відліків/с} \cdot 180\text{с} &= 7\,938\,000 \text{ відліків} \\ 44100 \text{ відліків/с} \cdot 20\text{мс/фрейм} &= 882 \text{ відліків/фрейм} \\ 7\,938\,000 \text{ відліків} / 882 \text{ відліків/фрейм} &= 9000 \text{ фреймів} \\ (\text{або } 180\text{с} / 20\text{мс/фрейм} &= 9000\text{фреймів}) \end{aligned}$$

Отже, середньостатистичний музичний твір характеризують, виходячи з наведених розрахунків, приблизно 8 млн. відліків або з врахуванням перекриття фреймів приблизно 18тис. фреймів (9000 фреймів\*2–1), кожен з яких описується вектором параметрів MFCC розмірності 13. Тобто параметризація дозволяє зменшити кількість інформації, необхідної для опису музичного твору, в десятки разів:

$$\frac{7\,938\,000 \text{ відліків}}{17999 \text{ MFCC} \cdot 13} = \frac{7\,938\,000}{233\,987} \approx 34 \text{ рази}$$

Таким чином, коефіцієнти MFCC є компактним представленням спектральної обвідної, що під час розпізнавання музичного твору дозволяє успішно замінити мільйони відліків аудіо файлу.

### Вибір методу порівняння невідомого музичного твору з шаблонами БД

Після визначення параметрів для опису музичних творів, необхідно перейти до етапу пошуку музичного твору серед шаблонів БД, тобто порівняння невідомого музичного твору з шаблонами та визначення шаблону, розходженням з яким буде мінімальним.

Для того, щоб ідентифікувати невідомий аудіо запис, необхідно мати критерій порівняння. Як правило, таким критерієм є відстань  $D$ . При цьому наголошуємо, що підхід порівняння за обраними параметрами має забезпечити високий рівень розрізнення власного шаблону  $\tilde{X}$  та шаблонів інших творів  $\tilde{Y}$ . Тобто в результаті порівняння невідомого музичного твору  $X$  з шаблоном власного твору похибка між ними має бути мінімальною, і навпаки максимальною – для шаблонів інших творів:

$$\begin{cases} D(X, \tilde{X}) \rightarrow \min \\ D(X, \tilde{Y}) \rightarrow \max \end{cases} \Rightarrow D(X, \tilde{X}) \ll D(X, \tilde{Y}), \quad (1)$$

де  $\tilde{X}$  – множина векторів параметрів шаблону власного музичного твору;  $\tilde{Y}$  – множина векторів параметрів шаблону іншого музичного твору.

Загальну схему ідентифікації музичного твору наведено на рисунку 1.

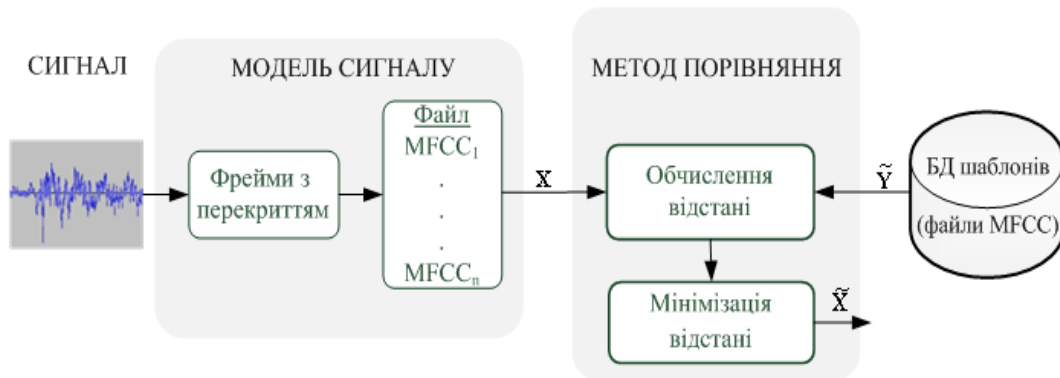


Рисунок 1 – Загальна схема ідентифікації музичного твору

Найпростішим і очевидним підходом для визначення близькості між наборами параметрів MFCC невідомого музичного твору та еталонів БД є порівняння MFCC на основі найбільш розповсюдженої Евклідової метрики, точніше квадрату Евклідової відстані  $D_{Eu}^2$  (щоб надати велику вагу більш віддаленим об’єктам). Формула незваженої Евклідової відстані між вектором параметрів музичного твору, який треба ідентифікувати,  $x = (x_1, x_2, \dots, x_d)$  та вектором шаблону БД  $\tilde{y} = (\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_d)$ :

$$D_{Eu}^2(x, \tilde{y}) = \sum_{i=1}^d (x_i - \tilde{y}_i)^2. \quad (2)$$

Відповідно відстань між файлами параметрів невідомого твору та шаблону БД можна знайти за формулою:

$$D(X, \tilde{Y}) = \sum_{j=1}^n D_{Eu}^2(x_j, \tilde{y}_j) = \sum_{j=1}^n \sum_{i=1}^d (x_{ji} - \tilde{y}_{ji})^2. \quad (3)$$

В ідеальному випадку при такому підході відстань між файлами параметрів MFCC одного і того ж музичного твору буде рівна нулю, для різних творів – відмінною від нуля:

$$\begin{cases} D(X, \tilde{X}) = 0, \\ D(X, \tilde{Y}) > 0. \end{cases} \quad (4)$$

Проте варто зауважити, що навіть для одного і того ж музичного твору аудіо записи можуть відрізнятися, наприклад: на початку запису може йти тиша, мелодія іншого музичного твору тощо; записи можуть мати різний темп, тривалість. Це означає, що у разі зсуву фреймів в часі відстань до власного шаблону  $D(X, \tilde{X}) \neq 0$ , тобто умова чіткого розрізнення (1) не виконується, отже, безпосереднє порівняння файлів параметрів за Евклідовою відстанню не підходить для задачі ідентифікації власного шаблону. В цьому випадку придатним для ідентифікації музичного твору є алгоритм динамічної трансформації шкали часу (DTW - Dynamic Time Warping), який дозволяє працювати з послідовностями векторів, що мають певний зсув в часі [9]. Однак використання DTW призведе до зростання кількості операцій порівняння, що є неприйнятним.

Розрахуємо складність алгоритму DTW за кількістю операцій порівняння  $N_{op}$  фрагменту та шаблону. Якщо вважати, що тривалість шаблону  $T_{song} = 3\text{хв}$ , що відповідає  $\approx 18000$  фреймів, а тривалість фрагменту  $\tau = 5\text{с}$ , тобто  $\approx 500$  фреймів, тоді:

$$\begin{aligned} N_{op} &= (T_{song} - \tau) \cdot \tau; \\ N_{op} &= (18000 \text{ фр} - 500 \text{ фр}) \cdot 500 \text{ фр} = 8750000 \approx 9 \cdot 10^6 \text{ операцій}. \end{aligned}$$

Очевидно, що в більшості випадків музичний твір характеризується певною періодичністю, що полягає в наявності ідентичних або дуже схожих за текстом та характером мелодії фрагментів. Відповідно можна говорити про надлишковість даних, якими описується музичний твір, і можливість скоротити кількість параметрів для його опису. З огляду на це доцільним є застосування методів кластерного аналізу. Використання кластеризації для формування еталонів, що містяться в БД, дозволить зменшити обсяги пам'яті, необхідні для їх зберігання.

Задача кластеризації даних є важливим елементом загальної проблеми обробки даних. Кластеризацію часто використовують, зокрема, під час статистичного аналізу даних, векторної квантизації, розпізнавання образів тощо. Кластеризація — це поділ множини вхідних даних (в нашому випадку векторів параметрів MFCC) на групи (кластери) за мірою «схожості» один на одного, тобто таким чином, щоб кожен кластер містив найбільш схожі об'єкти, а об'єкти різних кластерів відрізнялися між собою. Задачу кластеризації можна сформулювати так: заданий набір з  $n$  векторів, кожен з яких має розмірність  $d$ , необхідно розбити на підмножини відповідно до заданого критерію оптимізації. Як правило, таким критерієм є мінімізація спотворення  $e_i^2 \rightarrow \min$ . Існують різні шляхи оцінювання спотворення, але в більшості прикладних реалізацій використовують суму середньоквадратичних Евклідових відстаней між центром кластеру (центроїдом)  $c_i$  і векторами параметрів, які до нього належать  $X_i = \{x\}, X_i \subset X$  [10, 11], тобто:

$$e_i^2 = \{c_i : \sum_{j=1}^{N_i} D_{Eu}^2(x_j, c_i) \mid x \in X_i \leq \sum_{j=1}^{N_i} D_{Eu}^2(x_j, c) \mid x \in X_i\},$$

$$X_i \subset X, \forall c \in X \setminus X_i, e_i^2 \rightarrow \min.$$

де  $N_i$  — кількість точок, що належать центроїду  $c_i$ .

Таким чином, шаблони музичних творів в БД можна описати кластерами параметрів MFCC, показаних на рисунку 2 у вигляді кіл (для двомірного випадку).

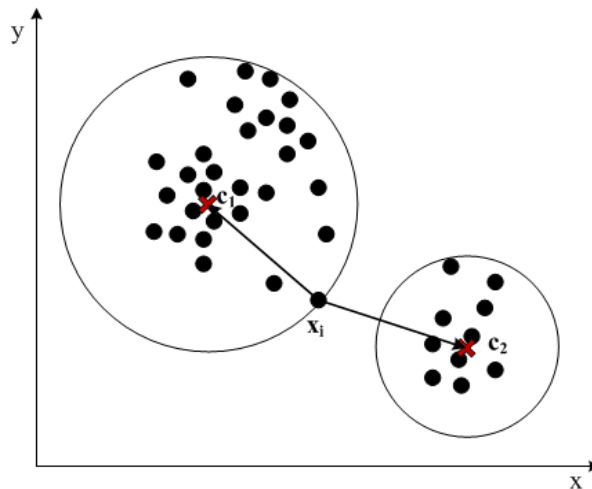


Рисунок 2 – Розбиття векторів параметрів на кластери

Кожен шаблон було представлено 1000 кластерів (замість 18 тис. векторів параметрів MFCC для твору тривалістю 3 хв.), тобто в середньому на кластер припадає близько 20 векторів. При подальшому збільшенні числа кластерів похибка кластеризації ще зменшиться, проте в такому випадку виникне ситуація, коли на кластер припадатиме менше 10 векторів, в результаті центроїди почнуть “підлаштовуватись” під конкретні вектори. Це призведе до погіршення результатів у випадку, якщо немає точного співпадіння між аудіо записами шуканого музичного твору та його власного шаблону. Зменшення кількості кластерів небажане, оскільки це призведе до зростання похибки кластеризації.

Відповідно, якщо, наприклад, шаблон представити 1000 кластерами, то обчислювальні затрати для порівняння його з фрагментом тривалістю  $\tau=5s$  (500фреймів) становитимуть:

$$N_{op} = T_{song} \cdot \tau = 1000 \text{ кл} \cdot 500 \text{ фп} = 5 \cdot 10^5 \text{ операцій} .$$

Тобто такий підхід порівняно з DTW дозволяє зменшити кількість операцій порівняння в десятки разів.

Основні етапи порівняння файлів параметрів невідомого музичного твору з певним шаблоном БД:

1. Пошук мінімальної евклідової відстані  $D_{\min}^2$  між поточним вектором параметрів  $\mathbf{x} = (x_1, x_2, \dots, x_d)$  з множини параметрів  $\mathbf{X} = \{\mathbf{x}_j, |\mathbf{X}| = n$  музичного твору, який треба ідентифікувати, та множиною векторів-кластерів  $\tilde{\mathbf{Y}} = \{\tilde{\mathbf{y}}_j, |\tilde{\mathbf{Y}}| = m$ ,  $\tilde{\mathbf{y}} = (\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_d)$  шаблону:

$$D_{\min}^2 = \min_j (D_{Eu}^2(\mathbf{x}, \tilde{\mathbf{y}}_j)) = \min_j \left( \sum_{i=1}^d (x_i - \tilde{y}_{ji})^2 \right), j = \overline{1, m}. \quad (5)$$

2. Обчислення оцінки відстані в цілому до шаблону як суми квадратів мінімальних відстаней  $D_{\min}^2$ :

$$D(\mathbf{X}, \tilde{\mathbf{Y}}) = \sum_{l=1}^n D_{\min}^2 = \sum_{l=1}^n \min_j \left( \sum_{i=1}^d (x_i - \tilde{y}_{ji})^2 \right), j = \overline{1, m}. \quad (6)$$

Повертаючись до рисунку 2 з векторами параметрів MFCC, згрупованих в кластери, видно, що між кожним кластером та векторами, які належать до нього, є похибка  $e_i^2$ , що є сумою квадратів відстаней між ними. Звідси випливає, що відстань між файлами параметрів, що описують один і той же музичний твір до кластеризації і після (навіть якщо аудіо записи були ідентичними), буде додатною і рівною величині сумарної похибки кластеризації  $E^2$ . Таким чином, відстань до власного шаблону дорівнюватиме:

$$D(\mathbf{X}, \tilde{\mathbf{X}}) = \sum_{l=1}^n D_{\min}^2 = E^2 = \sum e_i^2, E^2 \rightarrow \min \quad (7)$$

#### Оцінювання похибки за приведеною власною відстанню

Очевидно, що оскільки твори мають різну тривалість, кожний твір характеризується власною кількістю фреймів, представлених параметрами MFCC. Під час кластеризації обчислюється однакова кількість кластерів для усіх шаблонів. Це призводить до того, що різні твори знаходяться в нерівних умовах, тобто для творів, тривалість яких більша, початкова похибка (між файлами того ж твору до і після кластеризації) теж буде більшою, оскільки в цьому випадку на кожен кластер буде припадати більше векторів параметрів. Позбутися цього можна за рахунок ділення відстані до власного шаблону БД, визначеної в формулі (7), на кількість фреймів музичного твору  $n$ . Назвемо цю величину приведеною власною відстанню ( $D_{ПВ}$ ) музичного твору:

$$D(\mathbf{X}, \tilde{\mathbf{X}}) = D_{ПВ} = \frac{\sum_{l=1}^n D_{\min}^2}{n} = \frac{\sum_{l=1}^n \min_j \left( \sum_{i=1}^d (x_i - \tilde{y}_{ji})^2 \right)}{n} = \frac{E^2}{n}, j = \overline{1, m}, \quad (3.1-8)$$

Таким чином,  $D_{ПВ}$  – по суті є математичним очікуванням (МО) похибки кластеризації:

$$D_{ПВ} = M_E = \frac{E^2}{n}. \quad (3.2-9)$$

Як можна побачити з формули (8) характеристика  $D_{ПВ}$  не залежить від кількості фреймів (тривалості запису). Таким чином, її можна використовувати як критерій прийняття рішення як для запису в цілому, так і для його окремого фрагменту. Проте це твердження буде справедливим, тільки якщо для різних фрагментів ця характеристика буде змінюватися незначно, тобто за умови стаціонарності процесу.

#### Обґрунтування можливості ідентифікації музичного твору за його фрагментом

Змінення значень похибки, що виникає при порівнянні певного запису музичного твору з шаблоном, є випадковою функцією (процесом), що протікає в часі  $E(t)$ . Значення похибки кожного музичного твору є окремими реалізаціями випадкової функції  $E(t)$ .

Виходячи з самого принципу формування кластерів можна очікувати, що процес змінення значень похибки кластеризації в часі буде носити стаціонарний характер. Проте це припущення потребує статистичної перевірки.

Згідно з визначенням стаціонарним процесом є такий, значення якого незалежно від часу коливаються біля їх середнього значення (оцінка МО). Відповідно у разі підтвердження стаціонарності похибка повинна бути більш-менш постійною протягом усього музичного твору, це, в свою чергу, є підставою для ідентифікації музичного твору за  $D_{ПВ}$ , що є МО похибки кластеризації, на основі фрагменту, причому обраному незалежно від проміжку часу.

Отже, в формалізованому вигляді випадкова функція  $E(t)$  називається стаціонарною, якщо всі її ймовірнісні характеристики не залежать від часу  $t$ : математичне очікування  $m_E(t)$ , кореляційна функція  $K_E(t, t+r)$  (включає дисперсію  $D_E(t)$ ).

$$\begin{aligned} m_E(t) &= M_E = \text{const}; \\ K_E(t, t+r) &= K_E(r), D_E(t) = K_E(t, t) = k_E(0) = \text{const}. \end{aligned} \quad (4.1-10)$$

Перевірка (10) здійснювалась для 20 музичних творів. Результати наведені для 4 з них. Проаналізуємо отримані дані з точки зору ймовірної стаціонарності похибки кластеризації  $E(t)$ . Математичне очікування  $m_E(t)$  на фрагментах  $\tau=15\text{с}$  відхиляється від математичного очікування, розрахованого для всього музичного твору  $\tilde{M}_E, \tilde{m}_{E_{\min}} \leq \tilde{M}_E \leq \tilde{m}_{E_{\max}}$ , незначно, наприклад: 1 –  $0,28 \leq 0,31 \leq 0,32$ ; 2 –  $0,36 \leq 0,38 \leq 0,40$ ; 3 –  $0,40 \leq 0,45 \leq 0,48$ ; 4 –  $0,37 \leq 0,43 \leq 0,50$ . Графіки кореляційної функції (показаної на рис. 3), отримані для фрагментів з початку, середини і кінця музичного твору (а саме другого, дев'ятого і шістнадцятого фрагментів по 15с), мають подібний характер.

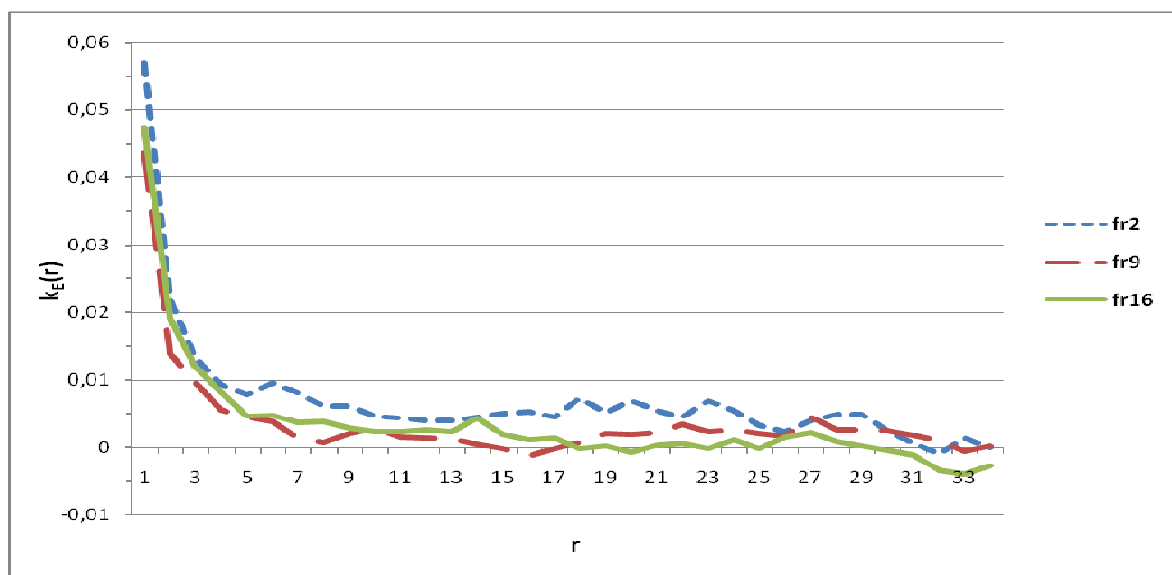


Рисунок 3 – Кореляційна функція для другого (fr2), дев'ятого (fr9) і шістнадцятого (fr16) фрагментів по 15с музичного твору

Таким чином, можна говорити про те, що припущення про стаціонарний характер процесу зміни значень похибки кластеризації  $E(t)$  підтверджено, відповідно ідентифікацію невідомого музичного твору доцільно здійснювати за його фрагментом.

Слід підкреслити, що особливо важливим є МО похибки  $M_E$ , оскільки ця характеристика є визначальною для ідентифікації музичного твору (згідно з формулами (3.1), (3.2)).

Хоча отримані статистичні дані підтвердили стаціонарність процесу зміни похибки (в результаті кластеризації), значення МО від фрагменту до фрагменту  $\tilde{m}_E(t_i), i = \tau, 2\tau, \dots, T_{\text{song}}$  все ж таки дещо коливаються відносно МО в цілому для музичного твору  $\tilde{M}_E = D_{ПВ}$  в межах від  $\tilde{m}_{E_{\min}} = \min(\tilde{m}_E(t_i)), i = \tau, 2\tau, \dots, T_{\text{song}}$  до  $\tilde{m}_{E_{\max}} = \max(\tilde{m}_E(t_i)), i = \tau, 2\tau, \dots, T_{\text{song}}$ :

$$\tilde{m}_{E_{\min}} \leq \tilde{M}_E \leq \tilde{m}_{E_{\max}}, \text{ або } \tilde{m}_{E_{\min}} \leq D_{ПВ} \leq \tilde{m}_{E_{\max}}.$$

Для характеристики цих коливань введемо коефіцієнти  $k_{\max \tau}, k_{\min \tau}$ , які розраховуються для кожного з еталонів музичних творів, що містяться в БД, за формулами:

$$k_{\max \tau} = (\tilde{m}_{E_{\max}} - \tilde{M}_E) / \tilde{M}_E = (\tilde{m}_{E_{\max}} - D_{ПВ}) / D_{ПВ}, \quad (4.2-11)$$

$$k_{\min \tau} = (\tilde{m}_{E_{\min}} - \tilde{M}_E) / \tilde{M}_E = (\tilde{m}_{E_{\min}} - D_{ПВ}) / D_{ПВ}.$$

Виходячи з цього сформулюємо умову, згідно з якою можна ідентифікувати невідомий фрагмент музичного твору, тобто визначити шаблон БД, що є його власним. Фрагмент аудіо запису можна вважати розпізнаним, якщо для відстані між фрагментом та певним шаблоном БД (що відповідно є власним шаблоном фрагмента) виконується нерівність:

$$(1 + k_{\min \tau}) \cdot D_{ПВ} \leq D(\mathbf{X}, \tilde{\mathbf{X}}) \leq (1 + k_{\max \tau}) \cdot D_{ПВ} \quad (4.3-12)$$

Відзначимо, що відповідно до формул (11) та (12), на основі яких має здійснюватись ідентифікація фрагмента музичного твору в БД музики, крім самих шаблонів музичних творів, для кожного музичного твору БД також мають зберігатись значення  $\tilde{m}_{E_{\min}}, \tilde{m}_{E_{\max}}$  та  $D_{ПВ}$ .

### Мінімальна і достатня тривалість фрагменту

В розділі 4 було доведено можливість ідентифікувати музичний твір за його фрагментом, зокрема тривалістю 15с. Зменшення тривалості фрагменту може зумовити зростання ймовірності помилки під час ідентифікації музичного твору. Однак крім надійності пошуку, велике значення має швидкість пошуку. Тому важливо, щоб тривалість фрагменту була якомога меншою. Відповідно важливою задачею є визначення мінімальної тривалості фрагменту, що дозволяє ідентифікувати музичний твір за шаблоном.

З цією метою було проведено дослідження, аналогічні описаним у розділі 4, на фрагментах тривалістю 1с та 5с.

В розділі 3 було введено поняття приведеної власної відстані музичного твору  $D_{ПВ}$ , на основі якої здійснюється ідентифікація фрагменту музичного твору за шаблоном згідно з формулою (12). Нагадуємо, що приведена власна відстань є МО похибки кластеризації ( $D_{ПВ} = \tilde{M}_E$ ) відповідно до формул (8) та (9).

На рис. 4 показано процес зміни в часі значень МО похибки на фрагментах 1с, 5с та 15с. В таблиці 1 та на рис. 5 на прикладі 4 пісень показано результати для мінімального  $\tilde{m}_{E_{\min}}$  та максимального  $\tilde{m}_{E_{\max}}$  значення МО похибки на фрагментах тривалістю 1, 5, 15с та усереднене значення МО в цілому для музичного твору  $\tilde{M}_E$ .

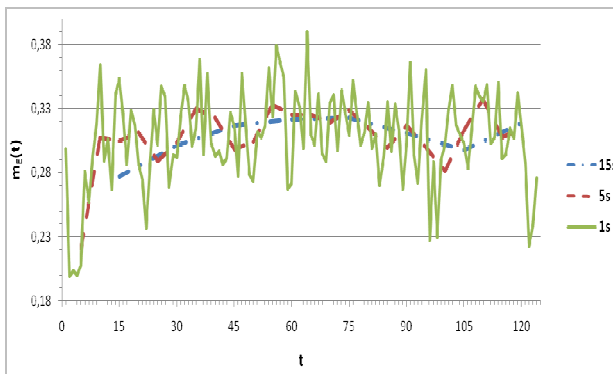


Рисунок 4 – Зміна значень МО похибки в часі для 1,5 та 15с

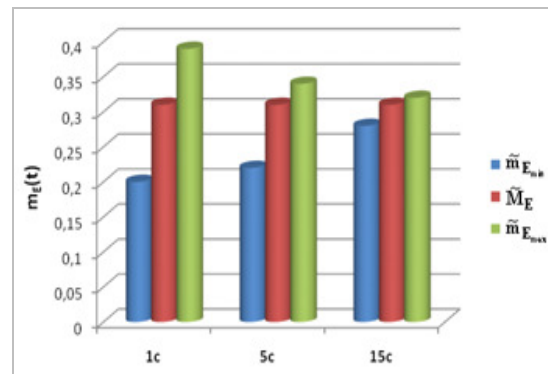


Рисунок 5 – Діапазон коливань  $\tilde{m}_E(t_i), i = \tau, 2\tau, \dots, T_{\text{song}}$  залежно від  $\tau$

Таблиця 1. Коливання МО похибки для фрагментів тривалістю 1, 5 та 15с

№	1	2	3	4
$\tau$	1с			
$\tilde{M}_E$	<b>0,31</b>	<b>0,38</b>	<b>0,44</b>	<b>0,43</b>
$\tilde{m}_{E_{\min}}$	0,20	0,31	0,34	0,32
$\tilde{m}_{E_{\max}}$	0,39	0,47	0,59	0,59
$\tau$	5с			
$\tilde{M}_E$	<b>0,31</b>	<b>0,38</b>	<b>0,45</b>	<b>0,43</b>
$\tilde{m}_{E_{\min}}$	0,22	0,34	0,37	0,35
$\tilde{m}_{E_{\max}}$	0,34	0,41	0,50	0,52
$\tau$	15с			
$\tilde{M}_E$	<b>0,31</b>	<b>0,38</b>	<b>0,45</b>	<b>0,43</b>
$\tilde{m}_{E_{\min}}$	0,28	0,36	0,40	0,37
$\tilde{m}_{E_{\max}}$	0,32	0,40	0,48	0,50

Отримані результати свідчать, що для усіх наведених тривалостей  $\tau$  значення МО від фрагменту до фрагменту  $\tilde{m}_E(t_i), i = \tau, 2\tau, \dots, T_{\text{song}}$  коливаються незначно відносно  $\tilde{M}_E$ , тобто стаціонарність зберігається.

Однак, як правило, випадковий процес починається з нестационарної стадії (перехідний процес), після чого процес можна вважати стаціонарним. Виходячи з цього небажано давати для розпізнавання фрагмент на початку або в кінці музичного твору.

З таблиці 1 та рисунку 5 також можна бачити, що зі зменшенням тривалості фрагменту  $\tau$  коливання  $\tilde{m}_E(t_i)$  стають більшими, тобто значення коефіцієнта  $k_\tau$  зростає.

Значне зростання коефіцієнта  $k_\tau$  зі зменшенням  $\tau$  фрагменту, особливо за умов великої кількості музичних творів в БД, може призвести до виникнення і швидкого зростання кількості випадків неправильного прийняття рішення, коли інший шаблон музичного твору буде прийнятий за власний. Таким чином, враховуючи отримані результати та зростання рівня коливань для фрагментів тривалістю  $\tau = 1с$ , що виражає коефіцієнт  $k_\tau$ , через недостатню кількість статистичних даних для висновків, було обрано мінімальну тривалість фрагменту –  $\tau = 5с$  для ідентифікації музичного твору за шаблоном.

#### Перевірка відповідності теоретичних припущень та експериментальних результатів

Для підтвердження теоретичних положень, наведених в попередніх розділах, було проведено експериментальне дослідження на базі 1000 музичних творів. Всі музичні твори мали формат wav (mono) з частотою дискретизації 44,1кГц. Попередньо з аудіо файлів було видалено тишу з початку та кінця записів. В процесі формування БД аудіо записи шаблонів необхідно було:

1. поділити на фрейми по 20мс з перекриттям 10мс;
2. для кожного фрейму розрахувати вектор параметрів MFCC розмірності 13.

Для отримання векторів параметрів MFCC з аудіо записів використовувався набір інструментів Матлаб – MIRtoolbox, що є відкритим ПЗ і описаний в [12].

Послідовності векторів параметрів MFCC, що описують музичні твори, було кластеризовано, використовуючи вдосконалений метод кластеризації k-середніх запропонований у одній з попередніх робіт [13]. В результаті чого кожен шаблон БД було представлено 1000 кластерів MFCC.

Одночасно в процесі кластеризації для кожного з 1000 музичних творів БД було визначено значення  $\tilde{m}_{E_{\min}}$ ,  $\tilde{m}_{E_{\max}}$  та  $D_{\text{ДВ}}$ , на основі яких має здійснюватись ідентифікація фрагмента музичного твору в БД музики згідно з формулою (12).

З 1000 еталонів БД випадковим чином було обрано фрагменти 100 пісень для їх ідентифікації. Результати, отримані в процесі порівняння 1000 шаблонів музичних творів БД з 100 фрагментів, які необхідно ідентифікувати, наведено в таблиці 2.

Таблиця 2. Результати ідентифікації фрагменту за шаблоном БД

№ фрагменту	№ шаблону	Тривалість фрагменту $\tau$								
		1с			5с			15с		
		$D(X, \tilde{X})$	$[\tilde{m}_{Emin}(\tilde{X}), \tilde{m}_{Emax}(\tilde{X})]$	$D_{min}(X, \tilde{Y})$	$D(X, \tilde{X})$	$[\tilde{m}_{Emin}(\tilde{X}), \tilde{m}_{Emax}(\tilde{X})]$	$D_{min}(X, \tilde{Y})$	$D(X, \tilde{X})$	$[\tilde{m}_{Emin}(\tilde{X}), \tilde{m}_{Emax}(\tilde{X})]$	$D_{min}(X, \tilde{Y})$
1	83	0,408	[0,219; 0,510]	0,561	0,360	[0,255; 0,474]	0,543	0,361	[0,291; 0,437]	0,588
2	96	0,512	[0,285; 0,666]	1,026	0,495	[0,333; 0,618]	1,173	0,496	[0,380; 0,571]	1,215
3	165	0,441	[0,258; 0,601]	0,554	0,439	[0,301; 0,558]	0,638	0,422	[0,344; 0,515]	0,567
4	200	0,437	[0,275; 0,641]	0,854	0,411	[0,320; 0,595]	0,772	0,466	[0,366; 0,549]	0,897
5	213	0,400	[0,237; 0,553]	0,632	0,386	[0,276; 0,513]	0,683	0,386	[0,316; 0,474]	0,692
6	268	0,465	[0,269; 0,628]	0,646	0,437	[0,314; 0,583]	0,737	0,459	[0,359; 0,538]	0,894
7	403	0,483	[0,279; 0,651]	0,587	0,454	[0,325; 0,604]	0,692	0,450	[0,372; 0,558]	0,650
8	408	0,397	[0,226; 0,527]	0,571	0,372	[0,264; 0,490]	0,621	0,372	[0,301; 0,452]	0,592
9	523	0,325	[0,179; 0,418]	0,470	0,301	[0,209; 0,388]	0,450	0,297	[0,239; 0,358]	0,466
10	525	0,480	[0,252; 0,587]	0,711	0,452	[0,294; 0,545]	0,703	0,419	[0,336; 0,503]	0,743
11	2	0,420	[0,285; 0,666]	0,650	0,453	[0,333; 0,618]	0,765	0,492	[0,380; 0,571]	0,802
12	64	0,315	[0,199; 0,465]	0,475	0,324	[0,232; 0,431]	0,477	0,318	[0,266; 0,398]	0,503
13	225	0,446	[0,283; 0,660]	0,695	0,426	[0,330; 0,613]	0,659	0,440	[0,377; 0,566]	0,746
14	600	0,240	[0,144; 0,335]	0,400	0,227	[0,168; 0,311]	0,406	0,228	[0,192; 0,287]	0,422
15	50	0,362	[0,226; 0,527]	0,735	0,392	[0,263; 0,489]	0,709	0,364	[0,301; 0,451]	0,630

В таблиці показано результати ідентифікації за шаблоном фрагментів 15 музичних творів для тривалостей 1, 5 та 15с. Зокрема в таблиці наведено відстань до власного шаблону  $D(X, \tilde{X})$  та відповідні значення  $\tilde{m}_{Emin}(\tilde{X}), \tilde{m}_{Emax}(\tilde{X})$ , а також відстань до найближчого шаблону іншого музичного твору  $D_{min}(X, \tilde{Y})$ .

Фрагмент вважається розпізнаним, тобто визначено власний шаблон  $\tilde{X}$  фрагменту, якщо для певного шаблону БД  $\tilde{Y}$  виконується  $D(X, \tilde{Y}) \in [\tilde{m}_{Emin}(\tilde{Y}); \tilde{m}_{Emax}(\tilde{Y})]$  відповідно до формули (12). Однак у випадку з фрагментом №3 таблиці тривалістю 1с було виявлено декілька шаблонів БД, для яких виконується (12): №159 –  $0,554 \in [0,254; 0,592]$ , №165 –  $0,441 \in [0,258; 0,601]$ . Це означає потенційну можливість неправильного прийняття рішення під час ідентифікації фрагменту тривалістю 1с, коли інший шаблон БД приймається за власний. Очевидно, що розходження з власним шаблоном повинно бути мінімальним, тобто власним все ж таки є шаблон під номером №165 з відстанню  $D(X, \tilde{X}) = 0,441$ , тому помилки можна уникнути.

### Висновки

В роботі теоретично обґрунтовано можливість ідентифікації музичного твору за його фрагментом на основі приведеної власної відстані (МО похибки кластеризації), значення якої не залежить від кількості фреймів (тривалості запису). Запропоновано аналітичний вираз для визначення власного шаблону музичного твору на основі його фрагменту. Визначено мінімальну тривалість фрагменту (5с), що дозволяє зменшити складність обчислень в десятки разів в процесі автоматизованої ідентифікації музичного твору. Проведені експериментальні дослідження підтвердили справедливості наведених теоретичних положень.

### Літературні джерела

1. Cano P. A review of audio fingerprinting / P. Cano, E. Batlle, T. Kalker, and J. Haitsma // Journal of VLSI Signal Processing. – 2005. – no. 41. – pp. 271–284.
2. Downie J.S. Music information retrieval / J.S. Downie // Annual Review of Information Science and Technology. – 2003. – no. 37. – pp. 295–340.
3. Hoashi K. Personalization of user profiles for content-based music retrieval based on relevance Feedback / K. Hoashi, K. Matsumoto, and N. Inoue // Proceedings of the ACM International Conference on Multimedia. – 2003. – pp. 110–119.

4. Wang Y. Multimedia content analysis using both audio and visual cues / Y. Wang, Z. Liu, and J. C. Huang // IEEE signal processing magazine. – 2000. – no. 17. – pp. 12–36.
5. Grosche P. Audio content-based music retrieval / P. Grosche, M. Müller, J. Serrà // Dagstuhl Follow-Ups Multimodal Music Processing. – V. 3. – Dagstuhl, Germany. – 2012. – pp. 157–175.
6. Ganchev T. Comparative evaluation of various mfcc implementations on the speaker verification task / T. Ganchev, N. Fakotakis, and G. Kokkinakis // Proceedings of 9th International Conference on Speech and Computer, SPECOM'05. – 2005. – pp. 191–194.
7. Logan B. A music similarity function based on signal analysis / B. Logan and A. Salomon // Proc. IEEE Int. Conf. Multimedia Expo. – 2001. – pp. 745–748.
8. Tzanetakis G. Musical genre classification of audio signals / G. Tzanetakis and P. Cook // IEEE Trans. Speech Audio Process. – No. 5. – V. 10. – 2002. – pp. 293–301.
9. Senin P. Dynamic time warping algorithm review / P. Senin – Honolulu, USA. – 2008.
10. Gersho A. Vector Quantization and Signal Compression. / A. Gersho, R. M. Gray. – Boston: Kluwer Academic. – 1992. – 760 p.
11. Jain A. K. Algorithms for Clustering Data / A. K. Jain, R. C. Dubes. – Englewood Cliffs, N.J.: Prentice Hall. – 1988. – 334 p.
12. Ткаченко О.М. Метод кластеризації на основі послідовного запуску k-середніх з удосконаленим вибором кандидата на нову позицію вставки / О. М. Ткаченко, О. Ф. Грійо Тукало, О. В. Дзісь, С. М. Лаховець // Електронний журнал «Наукові праці ВНТУ». – №2. – В.2. – Вінниця: ВНТУ. – 2012.
13. Lartillot O. A Matlab Toolbox for Musical Feature Extraction From Audio / O. Lartillot, P. Toivaiainen // International Conference on Digital Audio Effects. – Bordeaux, France. – 2007.

#### **Інформація про авторів**

**Ткаченко Олександр Миколайович** – к.т.н., доцент кафедри обчислювальної техніки, Вінницький національний технічний університет, Хмельницьке шосе, 95, м. Вінниця, alextk1960@gmail.com.

**Грійо Тукало Оксана Франсисківна** – аспірант кафедри обчислювальної техніки, Вінницький національний технічний університет, Хмельницьке шосе, 95, м. Вінниця, xxmargoxx@gmail.com.

УДК 004.9

О. М. ЯШИНА

Хмельницький національний університет, м. Хмельницький

**СТРУКТУРНО-АНАЛІТИЧНЕ ЗАБЕЗПЕЧЕННЯ ПРОЦЕСУ УПРАВЛІННЯ ДОСТУПОМ ДО ІНФОРМАЦІЙНОЇ СИСТЕМИ ТОРГІВЕЛЬНОЇ МЕРЕЖІ СУПЕРМАРКЕТУ**

**Анотація.** В даній статті обґрунтована структурна схема моделі ієрархічного ролевого управління доступом до інформаційної системи торгівельної мережі супермаркету. Запропоновано аналітичний опис процесу управління доступом до інформаційних ресурсів торгівельної мережі супермаркету.

**Ключові слова:** ієрархічне ролеве управління доступом, розподілена структура інформаційної мережі, система безпеки інформаційної мережі, атрибути керування доступом.

**Аннотация.** В статье обоснована структурная схема модели иерархического ролевого управления доступом к информационной системе торговой сети супермаркета. Предложено аналитическое описание процесса управления доступом к информационным ресурсам торговой сети супермаркета.

**Ключевые слова:** иерархическое ролевое управление доступом, распределенная структура информационной сети, система безопасности информационной сети, атрибуты управления доступом.

**Abstract.** In this article a model structural scheme of hierarchical role management access to information system of trading supermarket network has been grounded. The analytical description of the management access process to information resources of trading supermarket network has been offered.

**Keywords:** information system, hierarchichal role management access, divided structure of information network system, safety system of information network, attribute of access control.

**Вступ**

Розвиток сучасних підприємств різноманітних сфер економіки, як правило, супроводжується синтезом корпоративних інформаційно-комунікаційних мереж, функціонування яких ґрунтується на використанні сучасних інформаційних технологій. Одним із завдань щодо стабільного (безпечного з точки зору захисту інформації) функціонування інформаційно-комунікаційних мереж (ІКМ) вважається контроль і управління за доступом до баз даних та інформації, що в них міститься.

На сьогодні відома значна кількість робіт, що присвячені розробці інформаційних технологій забезпечення безпеки та управління доступом до інформаційних мереж (систем) [1,2]. Як на думку автора, управління доступом до інформаційних систем (ІС) сучасних торгівельних мереж супермаркетів доцільно організувати з використанням смарт-технологій [2]. Однак, наукових робіт, присвячених застосуванню смарт-технологій для ідентифікації користувачів інформаційних систем торгівельних мереж недостатньо. Так в роботах Фороузана Б.А., Шнайера Б., Шоріна Д.В., Шкурка М.І., Борисенка О.В., Стасенка Л., Кулікова А.Л. та інш. висвітлені окремі аспекти застосування як криптографії, так і смарт-технологій для забезпечення інформаційної безпеки інформаційних систем загального, або конкретного призначення, що ускладнює їх використання для управління доступом до ІС торговельних мереж супермаркетів [3,4]. Це пов'язано із необхідністю врахування специфіки функціонування ІС торговельної мережі супермаркету, а саме: обсягів інформації, що в них утримується, а також особливості її утримання, оновлення, циркуляції, що в свою чергу обумовлює особливості допуску до цієї інформації.

Усі ключові бізнес-процеси в торгівельній мережі, такі як фінансовий і бухгалтерський облік, управління кадрами, клієнтами, товаром і складом, документообіг, автоматизуються відповідним класом систем. Процес впровадження великої кількості інформаційних систем, при всій їх незаперечній корисності, несе в собі нові витрати і ризики для торгівельної мережі. Обов'язковою умовою високої конкурентоспроможності супермаркету стає захист інформації.

У таких умовах надання доступу до інформаційних систем тільки уповноваженим співробітникам – одна з найважливіших завдань ІТ та підрозділів служб безпеки. Основний крок у вирішенні даного завдання – побудова системи надійної аутентифікації користувачів. Класичний спосіб аутентифікації з використанням статичного пароля не забезпечує адекватного рівня захисту [2]. Необхідно використання сучасних надійних технологій. Застосування технологій строгої аутентифікації є фундаментом для побудови надійних і безпечних систем управління доступом.

Таким чином, викладене вище обумовлює необхідність синтезу інформаційної технології управління доступом до інформаційної системи торгівельної мережі супермаркету, що являє собою актуальне науково-прикладне завдання.

Відомо, що управління доступом до інформаційних ресурсів направлено на забезпечення політики інформаційної безпеки шляхом встановлення порядку доступу суб'єктів до відповідних баз даних. Для встановлення порядку та пріоритетів доступу до інформаційної мережі (ІМ) супермаркету наочна необхідність обґрунтування структури ІС торговельної мережі та відповідної структури управління

доступом до її баз даних з урахуванням ієрархічного розподілу функцій персоналу (рольового управління). Наявність таких структур відкриває можливість аналітичного опису процесу управління доступом до ІС торговельних мереж і синтезу математичної моделі доступу. Сутність математичної моделі управління доступом до ІС торговельних мереж полягає в формальному визначенні безпеки системи при певних умовах і обґрунтуванні кількісної міри безпеки. При цьому слід відмітити, що математична модель управління доступом повинна адаптивно відображувати стан всієї ІС торговельної мережі, її переходи з одного стану в інший, реакцію на зовнішні та внутрішні впливи, а також визначати безпечні стани і переходи впродовж управління доступом до інформаційних ресурсів системи.

*Метою статті* є обґрунтування структури ІС торговельної мережі та відповідної їй структурної схеми ієрархічного, рольового управління доступом. Побудова математичної моделі управління доступом до ІС торговельної мережі супермаркету у вигляді множинно-формалізованого опису процедури управління доступом до інформаційних ресурсів та обґрунтуванні рівня безпеки.

#### Результати дослідження

Зазвичай у структурі підрозділів торговельної мережі супермаркету притаманна ієрархічна побудова, що обумовлює розподілену структуру її інформаційної системи, як це показано на рис.1.

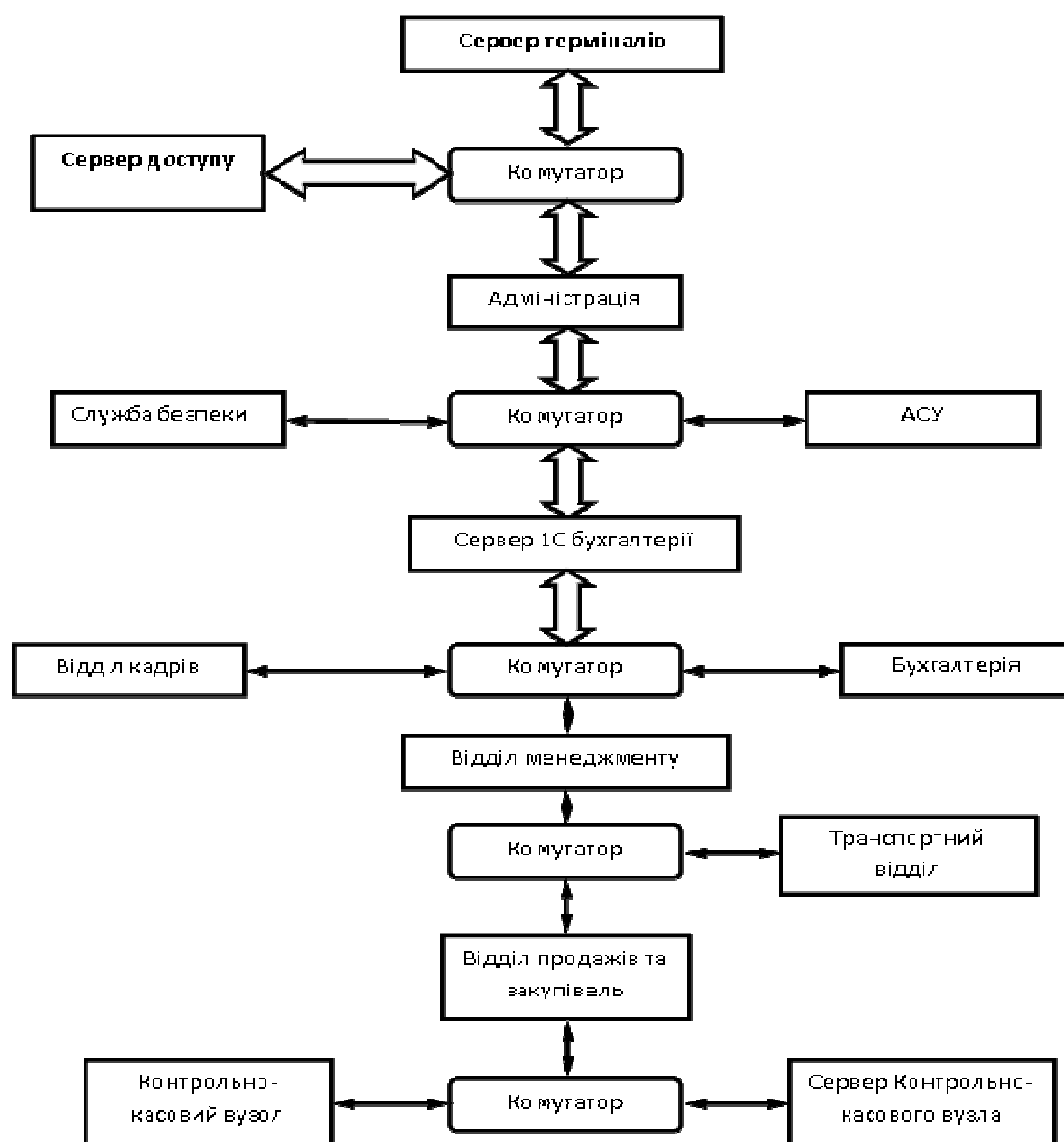


Рисунок 1 – Структурна схема розподіленої інформаційної системи торговельної мережі супермаркету

У відповідності зі встановленими організаційно-управлінськими відносинами та заданою політикою безпеки ІС торговельної мережі супермаркету (ТМСМ) користувачам надаються деякі види прав доступу

до даних свого підрозділу, що визначаються функціональними та посадовими обов’язками. При цьому співробітники вищої ланки по організаційно-штатній структурі підрозділів повинні мати доступ до даних підлеглих нижчої ланки. Разом з тим співробітники нижчої ланки не повинні отримувати права доступу до даних вищої ланки та топ-менеджерів. Отже, у відповідності із вимогами політики безпеки управління доступом до відповідних баз даних, надання прав доступу залежить від рівня сутностей в ієрархії ІС торгівельної мережі супермаркету.

Припустимо, що торгівельна мережа, а саме інформаційна система в цій мережі має деревовидну ієрархічну структуру (як це показано на рис.1) і для кожної сутності визначений певний рівень ієрархії. Така побудова дає підстави розглядати рівень ієрархії сутності якості одного з атрибутів, що використовуються при керуванні доступом. На основі цього атрибуту визначена політика управління доступом до системи: призначення всім сутностям рівнів ієрархії у відповідності зі структурою інформаційної системи та надання суб’єкту права доступу до сутності тільки в тому випадку, якщо рівень ієрархії суб’єкта не менший рівня ієрархії сутності (на множині ієрархій сутностей повинна бути задана верхня напіврешітка [5]).

Враховуючи основні вимоги управління доступом до інформаційної системи ТМСК запропоновано ієрархічну структуру рольового управління доступом у вигляді рис.2.

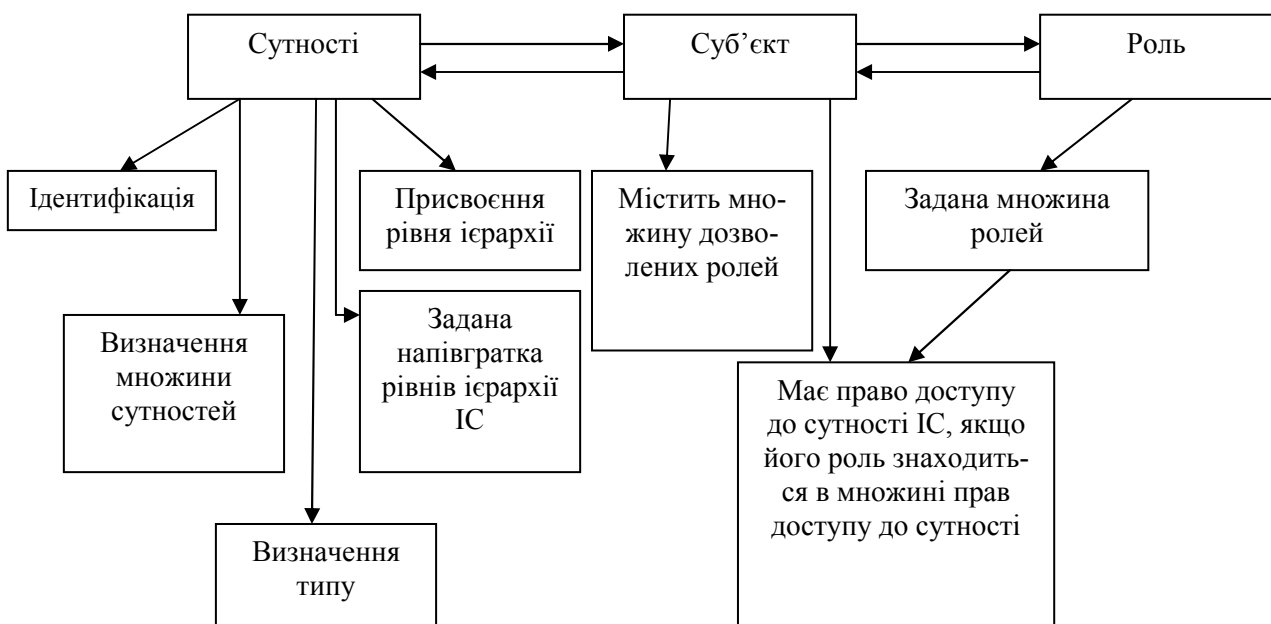


Рисунок 2 – Структура інформаційної технології ієрархічного рольового управління доступом до ІС торгівельної мережі

Для побудови формальних моделей інформаційна система, як правило, представляється у вигляді сукупності взаємодіючих сутностей – суб’єктів (S) та об’єктів (O). Суб’єкт безпеки – активна системна складова, до якої застосовується політика безпеки, а об’єкт – пасивна. Користувачі, дані, файли, системні таблиці, периферійні пристрої – це суб’єкти. З точки зору формалізованого подання інформаційної технології управління доступом (ІТУД) в даній роботі під суб’єктами слід розуміти підмножину сукупності об’єктів. На практиці реалізація політики безпеки (ПБ) інформаційної системи ТМСМ полягає у присвоєнні суб’єктам та об’єктам ідентифікаторів та фіксації набору правил, що дозволяють визначити дозвіл даного суб’єкта на авторизацію, достатню для надання до певного об’єкту вказаного типу доступу.

Модель, що описує політику ієрархічного рольового управління доступом до інформаційної системи торгівельної мережі є ієрархічною рольовою моделлю управління доступом та позначається RBAC-M. Основними елементами цієї моделі, згідно із [5,6,7], слід вважати:

$E = O \cup C$  – множина сутностей, де O – множина суб’єктів, C – множина об’єктів та  $O \cap C = \emptyset$ ;

U – множина користувачів, при цьому користувачі не є сутностями ( $U \cap E = \emptyset$ );

$S \subseteq E$  – множина суб’єкт-сесій користувачів;

T – множина типів сутностей;

L – множина рівнів ієрархії сутностей;

$R_l$  – множина видів прав доступу;

$R$  – множина ролей.

Згідно із [5] отримаємо:

$P \subseteq (R_r \times T) \cup (R_r \times E)$  – множина прав доступу до всіх сутностей одного типу та окремих сутностей;

$PA : R \rightarrow 2^P$  – функція прав доступу ролей, що задає для кожної ролі множину прав доступу до сутностей, при цьому для кожного права доступу  $p \in P$  існує роль  $r \in R$ , при якій виконується умова  $p \in PA(r)$ ;

$UA : U \rightarrow 2^R$  – функція авторизованих ролей користувачів, що задає для кожного користувача множину ролей, на які він може бути авторизований;

$type : E \rightarrow T$  – функція типів сутностей;

$f_c E \rightarrow L$  – функція, що задає рівень ієрархії кожної сутності;

$user : S \rightarrow U$  – функція приналежності суб'єкт-сесії користувачу, що задає для кожної суб'єкт-сесії користувачі, від імені якого вона авторизована;

$roles : S \rightarrow 2^R$  – дана функція задає для користувача множину ролей, на які він авторизований поточною сесією, при цьому в кожному стані комп'ютерної системи для кожної суб'єкт-сесії  $s \in S$  виконується умова  $roles(s) \subseteq U A(user(s))$ .

Введемо ще одне значення  $X$ , яке будемо називати розбиттям множини  $E$  у відповідності із заданою ієрархією сутностей, при цьому  $|X| = |L|$ . Доменом  $d$  сутностей множини  $E$  будемо називати будь-який клас із  $X$ . Ієрархією доменів назвемо задане на множині  $X$  відношення часткового порядку  $\leq$ , що задовольняє таким вимогам:

– якщо для  $d \in X$  існують  $d_1, d_2 \in X$ , такі, що  $d \leq d_1, d \leq d_2$ , то  $d_1 \leq d_2$  або  $d_2 \leq d_1$ ;

в  $X$  існує найбільший елемент.

описана ієрархія доменів відповідає ІС із ієрархічною деревовидною структурою, що відображає організаційно-управлінські відносини, та задаю верхню напіврешітку  $(X, \leq)$ .

Нехай  $L$  – задана множина рівнів ієрархії сутностей та існує бієктивне відображення  $X$  на  $L$  [8]. Визначимо на множині  $L$  відношення часткового порядку  $\leq$ , де для будь-яких  $l_1, l_2 \in L$  буде вірним вираз тоді і тільки тоді, коли для відповідних  $d_1, d_2 \in X$  тоді  $(L, \leq)$  – верхня напіврешітка рівнів ієрархії сутностей.

Будемо вважати, що множини  $U, X, T, L, P, R, R_r$  та функції не змінюються з часом.

Нехай  $\epsilon$  множини  $E, S, X, U, T, L, P, R, R_r$ , функції  $PA, UA, type, user, roles, (L, \leq)$  – напіврешітка рівнів ієрархії. Визначимо предикат  $can\_access(s, e, p)$ , що істинний тоді і тільки тоді, коли виконується такі умови:

1)  $f_e(e) \leq f_e(s)$ ;

2)  $(p, type(e)) \in PA(roles(s))$ .

Будемо вважати, що в ІС реалізовано рольове управління доступом RBAC-M, якщо будь-яка суб'єкт-сесія  $s \in S$  користувача  $user(s) \in U$  може мати право доступу  $p \in R_r$ , тоді і тільки тоді коли істинним є предикат  $can\_access(s, e, p)$ .

Одним із важливих механізмів сімейства моделей управління доступом є обмеження, що накладаються на множину ролей, на які може бути авторизований користувач або на які він авторизується протягом однієї сесії. Основні обмеження, визначені в рамках моделей рольового управління доступом сімейства RBAC можуть бути перенесені у запропоновану модель. Разом з тим для більшої відповідності ієрархічного управління доступом процедурам обробки даних, що використовуються у торгівельній мережі необхідно визначити нові види обмежень, суттєві саме для даного виду управління доступом.

Визначимо функцію  $h : E \rightarrow X$ , таку, що  $h(e) = d$ , якщо  $e \in d, e \in E$  та  $d \in X$ .

В моделі RBAC-M задано обмеження на область доступу користувача в ієрархії сутностей при виконання таких умов:

1. Визначена функція  $UD : U \rightarrow 2^X$ .

2. Для будь-яких  $s \in S, e \in E, p \in R_r$ , якщо предикат  $can\_access(s, e, p)$  та  $user(s) = u$ , то  $h(e) \in UD(u)$ .

Будемо вважати, що в моделі RBAC-M задані обмеження на область авторизації ролей в ієрархії сутностей, якщо виконуються такі умови:

визначено функцію  $RD : R \rightarrow 2^X$ ;

для будь-яких  $s \in S, e \in E, p \in R_r$ , якщо є істинним предикат  $can\_access(s, e, p)$  та  $r \in roles(s), moh(c) \in RD(r)$ .

Будемо вважати також, що в моделі RBAC-M задані обмеження на область застосування прав доступу в ієрархії сутностей, якщо виконуються такі вимоги:

визначена функція  $URD : U \times R \rightarrow 2^X$ .

для будь-яких  $s \in S, e \in E, p \in R_r$ , якщо є істинним предикат  $can\_access(s, e, p)$ ,  $user(s) = u, r \in roles(s), moh(e) \in URD(u, r)$ .

Будемо вважати, що в моделі RBAC-M задані обмеження на область застосування прав доступу в ієрархії сутностей, якщо виконуються наступні вимоги:

визначена функція  $PD : P \rightarrow 2^X$ ;

для будь-яких  $s \in S, e \in E, p \in R_r$ , якщо є істинним предикат  $can\_access(s, e, p)$ , то  $h(e) \in PD(p)$ .

Вважатимемо, що в моделі RBAC-M задані обмеження на область доступу до типу сутностей в ієрархії сутностей, якщо виконуються такі умови:

визначена функція  $TD : T \rightarrow 2^X$ ;

для будь-яких  $s \in S, e \in E, p \in R_r$ , якщо істинним є предикат  $can\_access(s, e, p)$ , то  $h(c) \in TD(type(c))$ .

Отже, загальна структура моделі ієрархічного рольового управління доступом торгівельної мережі матиме вигляд, як це показано на рисунку 3.

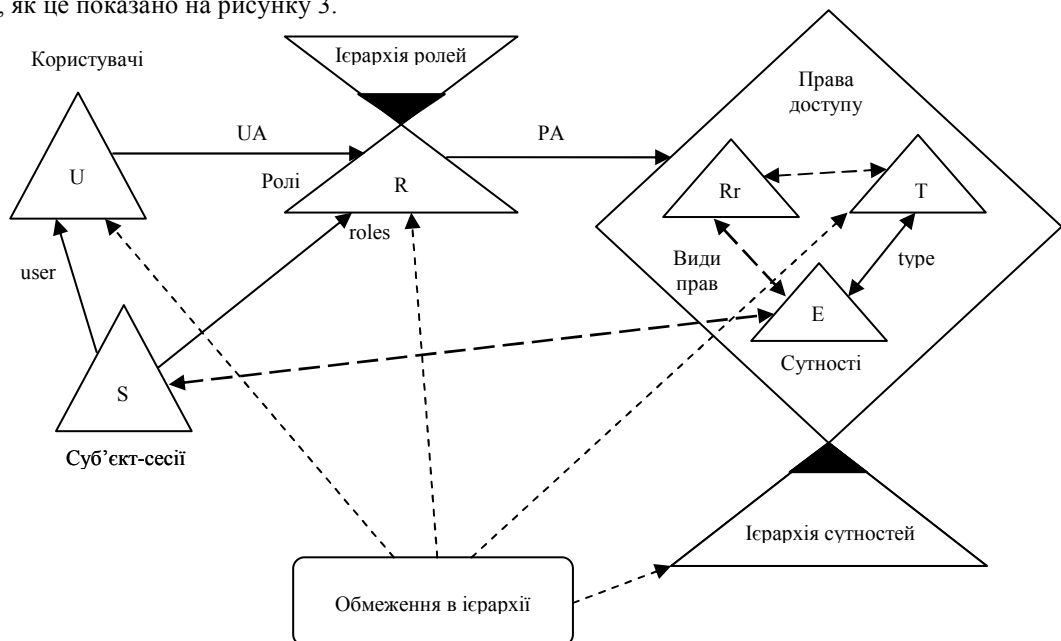


Рисунок 3 – Модель ієрархічного рольового управління доступом до ІС торгівельної мережі

### Висновки

На підставі обґрунтованих структурної схеми розподіленої інформаційної системи торгівельної мережі супермаркету та структури інформаційної технології ієрархічного рольового доступу до інформаційної системи торгівельної мережі проведено формалізований опис політики безпеки доступу до відповідної інформаційної системи.

Встановлено, що для інформаційної системи, яка характеризується ієрархією сутностей, адекватною математичною моделлю керування доступом слід вважати рольову модель керування доступом виду RBAC-M. Саме така модель здатна відобразити встановлені організаційно-управлінські відносини, що реалізуються в інформаційних системах торгівельних мереж супермаркетів. При цьому запропоновані атрибути ієрархії інформаційної

системи торгівельної мережі та типи сутностей до елементів моделей RBAC, що дозволяє адаптувати їх до умов функціонування ІС торгівельних мереж. Це в свою чергу спрощує реалізацію рольового управління доступом до ІС торгівельних мереж та відкриває можливість апаратно-програмної реалізації інформаційної технології процесу керування доступом до ІС торгівельної мережі.

#### Література

1. Куликов А. Л. Информационная система университета с управлением приоритетами доступа к ресурсам на основе СМАРТ-технологий: Дис. ... канд. техн. наук: 05.13.10: Астрахань, 2005. – 197 с.
2. Шинкарук О.М., Яшина О.М. Використання смарт-карт для ідентифікації користувачів інформаційних систем. // Вісник Хмельницького національного університету. Хмельницький: ХНУ. – №1, 2013. С.114-116.
3. Фороузан Б. А. Математика криптографии и теория шифрования информация. Режим доступа – <http://www.intuit.ru/department/security/mathcryptet/14/3.html>.
4. Шнайер Б. Прикладная криптография. – М.: Триумф, 2003. – 815 с.
5. Колегов Д.Н. Построение иерархического ролевого управления доступом // Математические основы компьютерной безопасности: ТГУ. – №3(17), 2012. С.71-76.
6. Девянин П. Н. Формирование словаря терминов теории моделирования безопасности управления доступом и информационными потоками в компьютерных системах // Прикладная дискретная математика. 2011. №2. С. 17-39.
7. Коньков А.К. Разработка и реализация моделей защищенности в рабочих группах и доменах Windows. [Электронный ресурс]: дис. ...канд. тех. наук: 05.13.18. – М.:РГБ, 2007.
8. Ершов Ю. Л., Палютин Е. А. Математическая логика: Учебное пособие. – 3-е, стереотип. изд. – СПб.: «Лань», 2004 – 336 с.

#### Інформація про автора

**Яшина Оксана Миколаївна** – ст. викладач кафедри програмної інженерії, аспірант Хмельницького національного університету.

## КОМП'ЮТЕРНІ СИСТЕМИ ТА КОМПОНЕНТИ

УДК 681.325.5

О. Д. АЗАРОВ, О. І. ЧЕРНЯК

Вінницький національний технічний університет

## ІНФОРМАЦІЙНО-СТРУКТУРНИЙ ПІДХІД ДО ВИРІШЕННЯ ПРОБЛЕМИ МІЖЗ'ЄДНАНЬ ПРИ РЕАЛІЗАЦІЇ ПОВНОФУНКЦІОНАЛЬНОЇ АРИФМЕТИКИ

**Анотація.** Запропоновано підхід до вирішення проблеми міжз'єднань за рахунок використання побітової потокової обробки у системі числення золотої 1-пропорції. Описано алгоритмічну і структурну організацію побітового додавання, як основної арифметичної операції у цій системі числення.

**Ключові слова:** проблема міжз'єднань, золота пропорція, побітова арифметика, потокова обробка.

**Аннотация.** Предложен подход к решению проблемы междоусоединений за счет использования побитовой потоковой обработки в системе счисления золотой 1-пропорции. Описаны алгоритмическая и структурная организации побитового суммирования, как основной арифметической операции в этой системе счисления.

**Ключевые слова:** проблема междоусоединений, золотая пропорция, побитовая арифметика, потоковая обработка.

**Abstract.** The approach for solution of interconnection problem by using bit-serial pipeline operation in golden 1-ratio numerical system is suggested. The algorithmic and structural organization of bit-serial addition as basic arithmetic operation in this numerical system is described.

**Key words:** interconnection problem, golden ratio, bit-serial arithmetic, pipeline operation.

## Вступ

Підвищення рівня інтеграції у цифровій техніці призводить до зменшення товщини провідників в інтегральних схемах і зменшення відстаней між ними. Для того, щоб при цьому не підвищувати опір провідників, збільшують їх висоту. Це призводить до взаємного впливу електричних сигналів у сусідніх провідниках. Підвищення тактової частоти в мікросхемах ще більш посилює цей вплив. Найбільшу проблему становлять довгі лінії інформаційного зв'язку, які мають велику площу взаємодії [1,2]. При зменшенні розмірів компонентів мікросхем виникають також й інші технологічні проблеми, складність яких постійно зростає [3]. Для подолання цих проблем провідні комп'ютерні фірми і наукові установи прикладають значні інтелектуальні й фінансові зусилля в розробки нових технологій. Проте на даний час всі розроблювані технології мають суттєві недоліки, що заважають їх широкому впровадженню. А саме:

- об'ємне монтування в кристалі (IBM) – паразитні наведення між шарами, збільшення потужності споживання [4];
- об'ємне монтування в кристалі з вбудованими каналами охолодження водою (IBM) – необхідність підвищення точності технології виробництва у 10 разів [5];
- з'єднання на вуглецевих нанотрубках (Toshiba) – не вирішені проблеми якості нанотрубок [6];
- оптичні волокна – складність пакування [7];
- оптичні хвилеводи (Sun Microsystems) – необхідність точного вирівнювання довжин з'єднань [8];
- оптичні комутатори (IBM) – складна технологія виготовлення [9];
- оптичні з'єднання через вільний простір – низька швидкість передавання, складність [10].

Перспективні дослідження у цій галузі зможуть вирішити дану проблему у тому стані, в якому вона є на даний час. Проте, вони будуть впроваджені у виробництво не раніше, ніж через 10–15 років. На той час рівень розпаралелювання, необхідний для вирішення задач обробки цифрової інформації, значно зростає (100–1000 процесорних ядер на кристалі). Це приведе до такого збільшення впливу проблеми з'єднань, що одними лише технологічними засобами її навряд чи можна буде вирішити.

## Актуальність

Одним з відомих напрямків інформаційно-структурного підходу до вирішення проблеми з'єднань є побітова обробка у класичній двійковій системі числення [11]. Вона забезпечує мінімальну кількість ліній інформаційних зв'язків між окремими обробляючими модулями, проте, значно збільшує час виконання арифметичних операцій. Тому така обробка виконується у потоковому режимі. Побітова потокова обробка широко використовується при побудові спеціалізованих обчислювальних засобів на основі FPGA [12] та у системах з можливістю реконфігурації і дозволяє отримати мінімальну кількість довгих ліній інформаційних зв'язків, які мають найбільший вплив на проблему з'єднань. Побітова обробка може бути ефективною при використанні, як мінімум, у двох електронних технологіях майбутнього. Це QSA, що належить до шести найперспективніших технологій [14], і SFQ, що вважається однією з найперспективніших [15].

Недоліком класичної двійкової системи числення є неможливість виконання всіх побітових арифметичних операцій в єдиному потоці, оскільки додавання, віднімання і множення можуть виконуватися, починаючи з молодших розрядів, а ділення – починаючи зі старших.

## Мета

У статті пропонується інформаційно-структурний підхід до вирішення проблеми з'єднань, що полягає у використанні системи числення золотої 1-пропорції для виконання всіх арифметичних операцій у потоковому режимі, починаючи зі старших розрядів. Даний підхід базується на можливості побітового потокового додавання у цій системі числення, починаючи зі старших розрядів.

### Організація побігової арифметики у системі числення золотої 1-пропорції

Особливістю системи числення золотої 1-пропорції є адитивне співвідношення між розрядами виду  $100=011$ . На основі даного співвідношення у [16] описано операції згортки і розгортки кодів у цій системі числення як логічні операції заміни значень одиниць на нулі і нулів на одиниці у розрядах при виконанні певних умов. Таке визначення згорток і розгорток формально не дозволяє використовувати їх що надає можливість виконання адитивних перетворень кодів чисел при виконанні певних умов. Адитивні перетворення (А-перетворення) – це умовні арифметичні операції, що змінюють код, не змінюючи його значення. А-перетворення є новим видом умовних арифметичних операцій, які виконують перенесення і запозичення при додаванні і відніманні. За рахунок надлишковості системи числення золотої 1-пропорції А-перетворення можуть виконуватись не тільки при переповненні розрядів, але і при досягненні у групі розрядів деякого граничного значення. А-перетворення  $n$ -розрядного коду числа з номером  $i$  ділить цей код на дві частини: молодші розряди від 0-го до  $(n-i-1)$ -го і старші розряди від  $(n-i)$ -го до  $(n-1)$ -го. За напрямком А-перетворення поділяються на перетворення з перенесенням у старші розряди (AL-перетворення) і перетворення з перенесенням у молодші розряди (AR-перетворення). За умовами виконання А-перетворення поділяються на елементарні (EA-перетворення), універсальні (UA-перетворення) та повні (FA-перетворення).

Перевірка умов виконання EA-перетворень є простою і потребує аналізу значень у кожному з трьох сусідніх розрядів. EA-перетворення кодів золотої 1-пропорції описуються такими виразами:

$$EAL_i(X_0^{n-1}): (X_{n-i}^{n-1} + \alpha^{n-i}) + (X_0^{n-i-1} - \alpha^{n-i-1} - \alpha^{n-i-2}) \text{ при } x_{n-i} \leq 0 \wedge x_{n-i-1} \geq 1 \wedge x_{n-i-2} \geq 1,$$

$$EAR_i(X_0^{n-1}): (X_{n-i}^{n-1} - \alpha^{n-i}) + (X_0^{n-i-1} + \alpha^{n-i-1} + \alpha^{n-i-2}) \text{ при } x_{n-i} \geq 1 \wedge x_{n-i-1} \leq 0 \wedge x_{n-i-2} \leq 0.$$

Результатом  $i$ -го EAL-перетворення є код 100 в розрядах з  $(n-i)$ -го по  $(n-i-2)$ -й. Результатом  $i$ -го EAR-перетворення є код 011 в розрядах з  $(n-i)$ -го по  $(n-i-2)$ -й.

У випадку, якщо в окремих розрядах умова певного А-перетворення не виконується, але в цілому значення старшої і молодшої частин коду задовольняє даній умові, виконується відповідне UA-перетворення. На відміну від EA-перетворень, перевірка умов в даному випадку є більш складною і потребує порівняння загальних значень частин коду. UA-перетворення кодів золотої 1-пропорції описуються такими виразами:

$$UAL_i(X_0^{n-1}): (X_{n-i}^{n-1} + \alpha^{n-i}) + (X_0^{n-i-1} - \alpha^{n-i-1} - \alpha^{n-i-2}) \text{ при } x_{n-i} = 0 \wedge X_0^{n-i-1} \geq \alpha^{n-i},$$

$$UAR_i(X_0^{n-1}): (X_{n-i}^{n-1} - \alpha^{n-i}) + (X_0^{n-i-1} + \alpha^{n-i-1} + \alpha^{n-i-2}) \text{ при } x_{n-i} = 1 \wedge X_0^{n-i-1} < \alpha^{n-i-2}.$$

Найпростіше UA-перетворення можна виконувати за допомогою послідовності EA-перетворень у відповідності до рекурсивних алгоритмів, граф-схеми яких наведені на рис. 1 а і б. Результатом  $i$ -го UAL-перетворення є код, у якого загальне значення розрядів з 0-го по  $(n-i-1)$ -й менше значення  $(n-i)$ -го розряду. Результатом  $i$ -го UAR-перетворення є код, у якого загальне значення розрядів з 0-го по  $(n-i-1)$ -й більше або дорівнює значенню  $(n-i)$ -го розряду.

FA-перетворення виконуються як суперпозиція UA-перетворень того ж виду:

$$FAL_i(X_0^{n-1}): UAL_{n-2}(\dots UAL_{i+1}(UAL_i(X_0^{n-1})))\dots),$$

$$FAR_i(X_0^{n-1}): UAR_1(\dots UAR_{i-1}(UAR_i(X_0^{n-1})))\dots).$$

Результатом  $i$ -го FAL-перетворення є код, в якому не виконуються умови жодного з  $i$ -го по  $(n-2)$ -го UAL-перетворень. Результатом  $i$ -го FAR-перетворення є код, в якому не виконуються умови жодного з  $i$ -го по 1-го UAR-перетворень.

FA-перетворення, використовуються для реалізації перенесення і запозичення при побітовому додаванні та відніманні кодів золотої пропорції. На кожному кроці побітового додавання кодів золотої пропорції розряди доданків додаються до групи з п'яти розрядів результату, через які можливе перенесення. Ця група називається проміжним результатом  $T$  і повинна оброблятися паралельно. Над отрима-

ним кодом виконується FAL-перетворення, що утворює черговий розряд результату додавання  $z$  і новий проміжний результат, який має інформаційний резерв для збільшення  $IRI$ . Даний резерв забезпечує обмеженість перенесення у старші розряди на наступному кроці. На рис. 2 представлено структурну організацію побітового суматора кодів золотої пропорції, що складається з регістра та комбінаційної частини, розділеної на суматор  $\Sigma$  і блок FAL-перетворення.

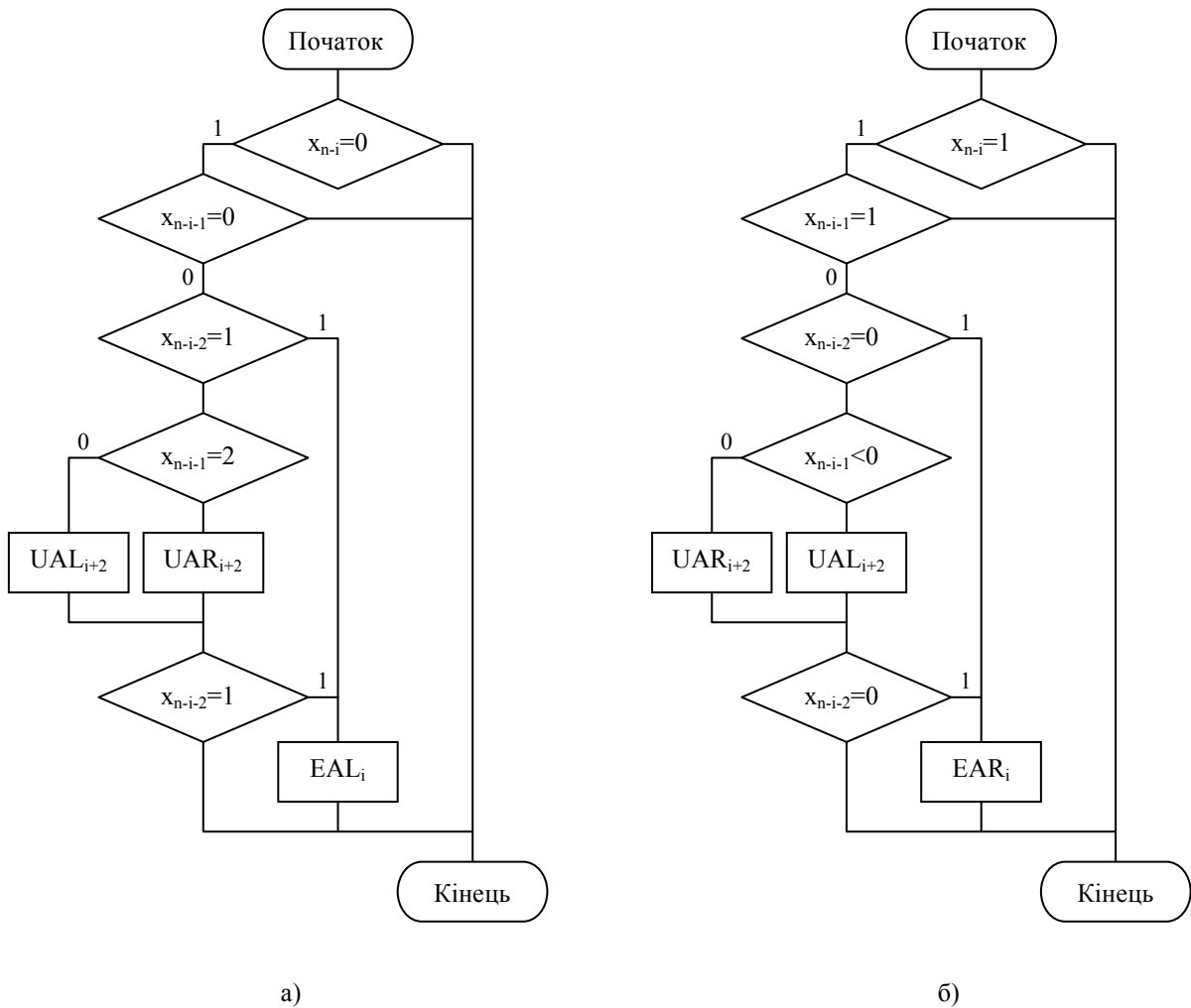


Рисунок 1 – Граф-схеми рекурсивних алгоритмів UA-перетворень кодів золотої пропорції:  
а)  $UAL_i$ -перетворення, б)  $UAR_i$ -перетворення

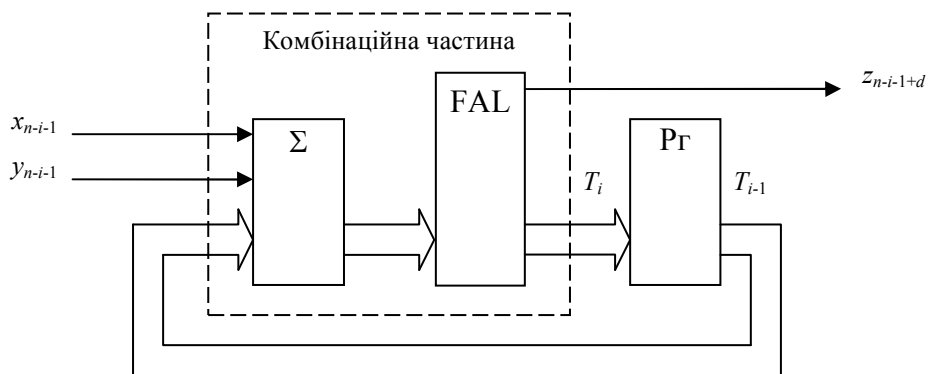


Рисунок 2 – Використання FAL-перетворення при побітовому додаванні кодів золотої 1-пропорції

Слід зазначити, що функціональне відокремлення блоку FAL-перетворення від суматора потрібно лише для побудови таблиці переходів комбінаційної частини пристрою додавання, яка формує сигнали на  $D$ -входи тригерів регістра. При створенні даного пристрою комбінаційна частина реалізується як одне ціле на основі таблиці переходів.

Для спрощення виразів, що описують дані сигнали, введемо допоміжні сигнали  $A, B$  і  $C$ .

$$\begin{aligned} A &= \overline{x + y}, \\ B &= \overline{x + y + C}, \\ C &= xy. \end{aligned}$$

Тут  $A=1$  означає, що сума чергових розрядів доданків дорівнює нулю,  $B=1$  означає, що ця сума дорівнює одиниці, а  $C=1$  означає, що вона дорівнює двом. Логічні вирази для чергового сигналу виходу та для  $D$ -входів тригерів регістра в залежності від його виходів  $Q$ , та сигналів  $A, B$  і  $C$  мають вид:

$$\begin{aligned} D_1 &= BQ_2 + C\overline{Q_1}\overline{Q_3}, \\ D_2 &= AQ_1, \\ D_3 &= AQ_2 + B\overline{Q_1}\overline{Q_2}\overline{Q_3} + CQ_3 + CQ_1Q_4, \\ D_4 &= AQ_3 + B(Q_1 + Q_2)\overline{Q_4} + D_1\overline{Q_2}\overline{Q_4}, \\ D_5 &= A + \overline{Q_3} + AQ_4 + B\overline{Q_1}Q_4 + C(Q_1 + Q_2)\overline{Q_4}, \\ z &= \overline{Q_4} + D_5. \end{aligned}$$

На вхід даного потокового пристрою побітового додавання на кожному такті поступають чергові розряди доданків, починаючи зі старших розрядів, а з виходу, починаючи зі старших розрядів, поступають розряди суми. Потокове побітове віднімання виконується аналогічним чином.

На основі додавання і віднімання можна виконувати побітове потокове множення і ділення також, починаючи зі старших розрядів [17]. Отже, використання системи числення золотої 1-пропорції дозволяє організувати виконання всіх арифметичних операцій в єдиному потоці забезпечуючи при цьому мінімальну кількість довгих ліній інформаційного зв'язку. Таким чином, запропонований авторами у даній статті інформаційно-структурний підхід вносить вклад у вирішення проблеми взаємозв'язків у цифровій техніці, доповнюючи технологічний напрямок вирішення даної проблеми.

#### Висновки

Обґрунтовано актуальність використання побітової потокової обробки як інформаційно-структурного підходу до вирішення проблеми між'єднань при реалізації засобів обчислювальної техніки у мікроелектронному виконанні.

Запропоновано використання системи числення золотої 1-пропорції для побітового виконання всіх арифметичних операцій в єдиному потоці, що дозволяє досягти мінімальної кількості довгих ліній інформаційних зв'язків в інтегральних мікросхемах.

Запропоновано нові умовні арифметичні операції адитивного перетворення, що дозволяють виконувати побітове додавання і віднімання з обмеженням довжини перенесення у старші розряди.

Запропоновано структурну організацію побітового потокового суматора кодів золотої 1-пропорції.

#### Список літератури

1. Яковлев Ю. С. Однокристалльные компьютерные системы высокой производительности. Особенности архитектурно-структурной организации и внутренних процессов : монография / Яковлев Ю. С. – Винница : ВНТУ, 2009. – 294 с.
2. SoC interconnect crisis: Path delays cancel speed increase [Електронний ресурс] / Chappell Brown, // EE Times. – Jun. 2003. – Режим доступу до мат. : <http://www.eetimes.com/story/OEG20030620S0028>.
3. Wilson R. Industry takes aim at 22nm interconnect stack [Електронний ресурс] / R.Wilson // EDN network: electronics news – Режим доступу до мат. : <http://www.edn.com/electronics-news/4314122/Industry-takes-aim-at-22-nm-interconnect-stack>.
4. Данилина Г. IBM переводит закон Гордона Мура в третье измерение [Електронний ресурс] / Г. Данилина. // Наука и образование – Режим доступу до мат. : <http://technomag.edu.ru/doc/71226.html>.

5. Будик А. IBM пронзила чипы водяными капиллярами [Електронний ресурс] / А. Будик. // 3D news: Новости Hardware – Режим доступу до мат. : [http://www.3dnews.ru/news/ibm\\_osnastila\\_chipi\\_vodyanimi\\_kapillyarami/](http://www.3dnews.ru/news/ibm_osnastila_chipi_vodyanimi_kapillyarami/).
6. Novikov A. Применение наноструктурированных материалов в технологии соединений [Електронний ресурс] / А. Novikov. // Abercade consulting: Аналитика – Режим доступу до мат. : <http://www.abercade.ru/research/analysis/1134.html>.
7. Ахманов А. С. Оптическая передача информации в суперЭВМ и микропроцессорных системах [Електронний ресурс] / А. С. Ахманов, О. Е. Наний, В. Я. Панченко. // Lightwave Russian Edition. – 2008. – № 3. – Режим доступу до мат. : [http://photonics.net.ua/files/LRE/2008\\_No\\_3.pdf](http://photonics.net.ua/files/LRE/2008_No_3.pdf).
8. G. Overton. Photonics applied: Can optical integration solve the computational bottleneck? [Електронний ресурс] – Режим доступу до мат. : <http://www.laserfocusworld.com/articles/2009/03/photonics-applied-integrated-photonics-can-optical-integration-solve-the-computational-bottleneck.html>.
9. Харьковский А. Оптический коммутатор IBM для применения в процессорах. [Електронний ресурс] / Александр Харьковский // 3DNews : Новости Hardware – Режим доступу до мат. : [http://www.3dnews.ru/news/opticheskiy\\_kommutator\\_ibm\\_dlya\\_primeneniya\\_v\\_protsessorah/](http://www.3dnews.ru/news/opticheskiy_kommutator_ibm_dlya_primeneniya_v_protsessorah/).
10. Venditti M. B. Design and test of an optoelectronic VLSI chip with 40 element receiver transmitter arrays using differential optical signaling / M. B. Venditti, E. Laprise, J. Faucher et al. // IEEE J. Select. Topics Quantum Electron. – 2003. – Vol. 9. – No. 2. – P. 361.
11. Hariyama M. A Field-Programmable VLSI Based on an Asynchronous Bit-Serial Architecture / M. Hariyama, S. Ishihara, C. Wei, M. Kameyama // In: IEEE, Asian Solid-State Circuits Conference / Jeju, Korea. – Nov. 2007. – P. 380–383.
12. Ferguson L. FPGA-based FIR Filter Using Bit-Serial Digital Signal Processing [Електронний ресурс] / Lee Ferguson // – 2000. – Режим доступу до мат. : [http://www.atmel.com/dyn/resources/prod\\_documents/DOC0529.PDF](http://www.atmel.com/dyn/resources/prod_documents/DOC0529.PDF).
13. Dittmann F. Path Concepts for a Reconfigurable Bit-Serial Synchronous Architecture [Електронний ресурс] / F. Dittmann, A. Rettberg, R. Weber // – 2005. – Режим доступу до мат. : <http://www.springerlink.com/content/y5kk543864352395/fulltext.pdf?page=1>.
14. Свидиненко Ю. Клеточные автоматы на квантовых точках : Nanotechnology News Network [Електронний ресурс] / Ю. Свидиненко // – 2006 – Режим доступу до мат. : <http://old.nanonewsnet.ru/index.php?module=Pagesetter&func=viewpub&tid=9&pid=93>.
15. V. K. Kornev, I. I. Soloviev and O. A. Mukhanov, "Possible Approach to the Driver Design Based on Series SQIF," IEEE Transaction on Applied Superconductivity. – 2005. – Vol. 15. – P. 388–391.
16. Азаров О. Д. Аналого-цифрове порозрядне перетворення на основі систем числення з ваговою надлишковістю : монографія / Азаров О. Д. – Вінниця : ВНТУ, 2010. – 232 с. – Режим доступу до мат. : <http://publish.vntu.edu.ua/txt/Azarov-AnCifrPorozrPeretvNaOsnNadlishkSistChislZVagovNadlishk354.pdf>
17. Азаров О. Д. Структурна організація побітового множення і ділення кодів золотої пропорції / О. Д. Азаров, О. І. Черняк // Проблеми інформатизації та управління. – 2007. – Вип. 3(21). – С. 5–13. – ISSN 2073-4751.

#### Відомості про авторів

**Азаров Олексій Дмитрович**, д.т.н., професор, директор інституту інформаційних технологій та комп'ютерної інженерії, завідувач кафедри обчислювальної техніки Вінницького національного технічного університету; адреса: 21021, м. Вінниця, вул. Порика, 16, кв. 13.; роб. т.: (0432) 43-90-02; дом. т.: 43-75-07; e-mail: azarov2@vstu.vinnica.ua.

**Черняк Олександр Іванович**, к. т. н., доцент кафедри обчислювальної техніки Вінницького національного технічного університету; адреса: 21021, м. Вінниця, вул. Келецька, 94, кв. 40; роб. т. (0432) 43-90-02; дом. т.: 46-37-16; e-mail: alexandr.chernyak@gmail.com.)

УДК 681.586.773

Й.І. СТЕНЦЕЛЬ, К.А. ЛІТВИНОВ, А.В. РЯБІЧЕНКО, О.І. ШАПОВАЛОВ

Східноукраїнського національного університету імені Володимира Даля, м.Сєверодонецьк

**ЗАСІБ КОНТРОЛЮ РІВНЯ РІДИННИХ СЕРЕДОВИЩ З КІЛЬЦЕВИМ РУХОМ УЛЬТРАЗВУКОВОГО СИГНАЛУ**

**Анотація.** У роботі описується ультразвуковий засіб контролю рівня, заснований на принципі кільцевого руху ультразвукового сигналу. Одним з елементів такого кільця є рідинне середовище технологічного апарату. Ультразвуковий сигнал проходить газове середовище тільки в одному напрямку – від поверхні рідинного середовища.

**Ключові слова:** ультразвук, засіб, контроль, середовище, поверхня, рідина, газ, спосіб, випромінювач, приймач, сигнал, вимірювання, діапазон, похибка.

**Аннотация.** В работе описывается ультразвуковое средство контроля уровня, основанное на принципе кольцевого движения ультразвукового сигнала. Одним из элементов такого кольца является жидкостная среда технологического аппарата. Ультразвуковой сигнал проходит газовую среду только в одном направлении - от поверхности жидкостной среды.

**Ключевые слова:** ультразвук, средство, контроль, среда, поверхность, жидкость, газ, способ, излучатель, приемник, сигнал, измерения, диапазон, погрешность.

**Abstract.** The paper describe the operation of the ultrasonic level controls, based on principle of the circular motion of the ultrasonic signal. One of the elements of such a ring is a liquid medium technological device. The ultrasonic signal passed the gaseous medium in one direction only – from the liquid medium to the receiver.

**Key words:** ultrasonic tool control medium surface, a liquid gas, the process transmitter, a receiver, a signal measurement range error.

**Вступ**

Ультразвукове випромінювання широко використовується в різних областях науки і техніки, у тому числі для контролю різноманітних технологічних параметрів, у дефектоскопії, для побудови неінвазивних засобів лікування тощо [1-5]. Ультразвуковий сигнал (УЗС) випромінюється відповідним випромінювачем (УЗВ) і направляється на досліджуваній об'єкт. Після перетворення УЗС повертається в ультразвуковий приймач (УЗП). На шляху свого проходження УЗС може декілька разів перетворюватися в різну форму. Кожний такий перехід є реологічним. В ультразвукових ЗКР використовуються, як правило, п'єзоелектричні (в основному п'єзокерамічні) елементи, які призначені для створення ультразвукових коливань (УЗК) при дії на них електричного збуджуючого імпульсу [6, 7]. Під дією електродинамічного зусилля п'єзоелектричний елемент, який має форму пластини круглої форми, деформується подібно металевій мембрані, яка жорстко закріплена за периметром. Після припинення дії електродинамічного зусилля УЗВ здійснює вільні механічні коливання, частота яких знаходиться в ультразвуковому діапазоні. З іншої сторони, механічні коливання наводять в УЗВ електрорушійну силу з частотою тих же коливань, яка створює електродинамічне зусилля, котре направлено протилежно до напрямку руху УЗВ, створюючи гальмуючу дію [10-12]. За рахунок цього загасання коливань УЗВ здійснюється швидше, що приводить до формування короткого ультразвукового імпульсу у формі інтегральної  $\delta$  – функції Дірака. Виходячи зі сказаного в УЗВ протікають електромагнітні процеси, які супроводжуються перетвореннями електричної енергії в механічну та навпаки.

Більшість відомих ультразвукових рівнемірів працює за принципом вимірювання часу проходження ультразвуковим імпульсом (УЗІ) подвійної відстані від передавача до контролюючої поверхні та визначає рівень за добутком половини вимірюного часу та відомої швидкості ультразвуку в даному середовищі. Недоліком таких рівнемірів є проходження УЗІ подвійної відстані в газовому середовищі (ГС), що суттєво послаблює сприймаючий сигнал, залежність його показань від параметрів цього середовища (тиску, температури, складу газу), в якому розповсюджується УЗІ. Ультразвуковий рівнемір «ВЗЛЕТ УР» виробництва ЗАТ «Взлет» [8], який працює за принципом вимірювання часу проходження УЗІ подвійної відстані від передавача до поверхні рідинного середовища (ПРС), але відрізняється тим, що в якості часу проходження УЗІ вимірюють час проходження від моменту зміни однієї частоти на іншу у випроміненому імпульсі до моменту відповідної зміни частоти в прийнятому передавачем імпульсі, відбитим від ПРС. Для компенсації зміни параметрів ГС у рівнемірі використовується реперна пластина, яка встановлена на фіксованій відстані від передавача. За часом проходження УЗІ подвійної відстані від передавача до реперної поверхні визначається швидкість розповсюдження УЗІ. До недоліків, які обмежують використання даного рівнеміра, відноситься наявність великої зони нечутливості (0,5 метрів) і достатньо малий діапазон вимірювання (від 0,5 до 5 метрів). В ультразвуковому рівнемірі МТМ900 виробництва ТОВ НВП «Мікротерм» [9] вимірювання рівня РС здійснюється за принципом визначення часу проходження УЗІ подвійної відстані від п'єзокерамічного передавача до ПРС, який реалізований в блоці датчика. Блок датчика складається з п'єзокерамічного передавача, який підключений до підсилювача. Управління підсилювачем здійснюється з допомогою мікроконтролера, який управляє блоком формування імпульсів збудження, підключеного до п'єзокерамічного випромінювача (ПКВ). Час проходження УЗІ визначається від моменту досягнення максимального відхилення мембрани передавача до

моменту перевищення встановленого значення підсилення відбитим сигналом. Порівняння відбитого сигналу з встановленим значенням здійснюється з допомогою мікроконтролера. Для компенсації зміни параметрів ГС використовується реперна пластина, яка встановлена на фіксованій відстані від ПКВ. До недоліків такого рівнеміра слід віднести: наявність великої зони нечутливості, яка досягає 600 мм, чутливість до перешкод і повторно відбитих УЗІ, які обумовлені великим коефіцієнтом підсилення при зміні малих значень рівня, малий діапазон вимірювального контролю, який не перевищує 5000 мм, що обмежує широке їх використання в системах господарських розрахунків особливо при контролі рівня РС у резервуарах висотою до 10 метрів.

### Мета статті

Метою статті є описання принципу роботи ультразвукового засобу контролю рівня, заснованого на кільцевому односторонньому проходженні випроміненого ультразвукового сигналу.

### Постановка задачі

Задача полягає у зменшенні зони нечутливості, розширенні діапазону вимірювального контролю та підвищенні точності й вірогідності вимірювального контролю рівня РС.

### Принцип роботи ультразвукового засобу контролю рівня з кільцевим проходженням ультразвукового сигналу

Поставлена задача вирішується за рахунок того, що вимірювання рівня здійснюється за рахунок часу проходження УЗІ тільки відстані від ПРС до приймача УЗІ, який визначається від моменту подачі одиночного електричного збуджуючого імпульсу (ЕЗІ) на ПКВ до моменту перевищення напруги, створеної п'єзокерамічним приймачем (ПКП) і підсиленої підсилювачем, встановленого значення опорної напруги. На рис. 1 приведена схема ультразвукового засобу контролю рівня рідинних середовищ з кільцевим проходженням ультразвукового сигналу.

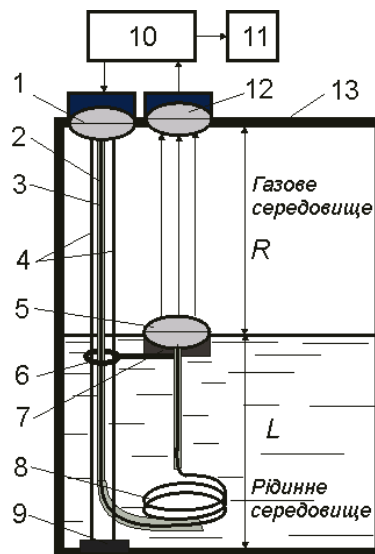


Рис. 1. Схема ультразвукового засобу контролю рівня з кільцевим проходженням УЗС

На рис.1 позначено: 1 - первинний випромінювач УЗПП1; 2 - гнучка металева линва для перенесення УЗІ; 3 - полімерне покриття гнучкої металевій линви; 4 - направляючі штанги; 5 - вторинний випромінювач УЗПП5; 6 - стабілізуючий хомут для поплавка вторинного УЗПП5; 7 - поплавок для розміщення вторинного УЗПП5; 8 - пружинна скрутка передаючої металевій линви; 9 - грузило для закріплення направляючих штанг до дна резервуару; 10 - вимірювальна схема; 11 - блок реєстрації вимірюваного значення рівня; 12 - приймач УЗІ, відбитого від ПРС; 13 - кришка резервуару для рідини. Принцип роботи такого УЗЗКР полягає в наступному. Первинний УЗПП1 розташований поруч з приймачем УЗП 12 на кришці резервуару 13, а вторинний УЗПП5 - на поплавку 7, який плаває на ПРС, і підключений до первинного УЗПП1 з допомогою металевій линви 2, яка ізольована від РС полімерним захисним покриттям 3. З рівномірного руху вторинного УЗПП5 відносно УЗП 12 встановлені штанги 4, які закріплені до дна резервуару грузилом 9.

Вимірювання рівня  $L$  рідини в резервуарі здійснюється за допомогою вторинного УЗПП5 5, який випромінює УЗІ і передає їх через ГС товщиною  $R$  до УЗП 12, котрий змонтований на кришці резервуару (13). Так як відстань від первинного УЗПП1 1 до УЗПП5 5 є сталою і визначається довжиною линви 2 зі захисним полімерним покриттям 3, по якій розповсюджуються УЗІ, а відстань від УЗПП5 до

УЗП визначається рівнем рідини  $L$  в резервуарі, то час проходження УЗІ цієї відстані буде залежним від рівня РС  $L$ . Так як сигнал первинного УЗПП1 пропорційний сигналу УЗПП5, а час проходження УЗІ по металевій линві є сталим і значно менший від проходження УЗІ через ГС, то цей час практично не впливає на результат вимірювання рівня. Час, за яким визначається рівень РС, складається з часу  $\tau_1$  подачі одиничного ЕЗІ на ПКВ, часу  $\tau_2$  формування УЗІ первинним УЗПП1, часу  $\tau_3$  проходженням УЗІ по металевій линві до вторинного УЗПП5, часу  $\tau_{\Gamma}$  проходження УЗІ газового середовища і часу  $\tau_4$  сприйняття УЗП, перетворення в напругу та підсилення останньої. Сумарний час, за який формується цикл вимірювального контролю, дорівнює:  $\tau_{\Sigma} = (\tau_1 + \tau_2 + \tau_3 + \tau_4) + \tau_{\Gamma} = \tau_0 + \tau_{\Gamma}$ , де  $\tau_0 = \tau_1 + \tau_2 + \tau_3 + \tau_4 = const$ . УЗПП5 конструктивно має приймаючу та випромінюючу металеві мембрани, між якими знаходиться п'єзокерамічний елемент (ПКЕ), який служить передавачем первинних УЗІ. Так як канали випромінювання та сприйняття УЗІ є розділеними, а УЗПП5 знаходиться в газорідному середовищі, яке характеризується тиском газу  $D$  і температурою  $\dot{O}$ , то це дозволяє створити реперний канал для тестування та налагодження нормального початкового стану вимірювального каналу з врахуванням зміни тиску ГС і температури РС. Принцип уведення поправки за реперним сигналом полягає в наступному. УЗІ, випромінений первинним УЗПП1, сприймається УЗПП5, ПКЕ котрого формує вторинний УЗІ, котрий одночасно формує реперний УЗІ. Останній повертається по металевій линві до первинного УЗВ, і формує напругу  $\dot{a}_p$  реперного сигналу, яка запам'ятовується вимірювальною схемою 10. Ця напруга порівнюється зі заданою опорною реперною напругою  $\dot{a}_{0p}$  і при необхідності вводиться поправка в результат тестування вимірювальної схеми. Час формування реперного сигналу  $\tau_p = 2(\tau_1 + \tau_2 + \tau_3)$  записується в пам'ять мікроконтролера і використовується для уведення поправки при розрахунку дійсного значення рівня РС. Випромінений УЗІ з УЗПП5, пройшовши ГС, сприймається УЗП.

За рахунок одноразового проходження УЗІ відстані  $R$  у ГС похибка вимірювання рівня РС, обумовлена зміною температури, тиску та складу цього середовища зменшується практично в 2 рази. Так як послаблення УЗІ в газовому середовищі зменшується практично в два рази (за рахунок одноразового проходження УЗІ відстані  $R$ ), то діапазон вимірювання рівня збільшується на стільки ж у порівнянні з аналогом.

### Принцип роботи вимірювальної схеми

Блок-схема вимірювального контролю ультразвукового засобу контролю рівня з кільцевим проходженням сигналу приведена на рис. 2, на якій позначено: 1 - мікроконтролер блоку обробки інформації; 2 - блок перетворення сигналу; 3 - блок пам'яті; 4 - блок індикації; 5 - блок перетворення сигналу; 6 - мікроконтролер; 7 - регульоване джерело імпульсів збудження; 8 - джерело опорної напруги; 9 - компаратор; 10 - блок визначення амплітуди сигналу; підсилювач. Ультразвуковий пристрій працює наступним чином. Після включення ультразвукового пристрою в роботу мікроконтролер 1 видає управляючий сигнал на регулююче джерело імпульсів збудження (РДІЗ), яке формує і видає одиничний ЕЗІ на первинний УЗПП1. Одночасно обнулюється і включається лічильник тактових імпульсів (ЛТІ). При цьому первинний УЗПП1 генерує УЗІ з максимальною амплітудою, який по ізольовані металевій линві направляється до вторинного УЗПП5. Останній створює УЗІ, який одночасно повертається по металевій линві до первинного УЗПП1 і створює реперний сигнал. Після проходження ГС УЗІ сприймається УЗП, в якому створюється ЕРС напругою  $\dot{a}_{\Gamma}$ . Остання після підсилення в підсилювачі напруги 11 подається в компаратор 9, де порівнюється зі заданою опорною напругою  $\dot{a}_{0\Gamma}$ , яка формується джерелом опорної напруги 7. Якщо різниця цих напруг не дорівнює нулю ( $\dot{a}_{\Gamma} - \dot{a}_{0\Gamma} \neq 0$ ), то компаратор видає сигнал на РДІЗ, який зменшує або збільшує амплітуду одиничного ЕЗІ до тих пір, поки не виконуватиметься рівність  $\dot{a}_{\Gamma} - \dot{a}_{0\Gamma} = 0$ . Коли  $\dot{a}_{\Gamma} - \dot{a}_{0\Gamma} = 0$ , то компаратор дає дозвіл мікроконтролеру 6 на включення ЛТІ і розрахунок значення рівня РС. ЛТІ працює до тих пір, поки не поступить на компаратор сприймаючий сигнал наступного такту вимірювання рівня. Цей сигнал запам'ятовується в запам'ятовуючому пристрої 3 мікроконтролера 1, видається команда на припинення роботи ЛТІ та його обнулення. Далі виконується порівняння вимірюваного значення рівня з градуовальною характеристикою, за якою визначається нове значення опорної напруги  $e_{0\Gamma}$ , котре запам'ятовується для виконання наступного такту вимірювання рівня.

При зміні рівня РС (наприклад, при його збільшенні) поплавков підніметься вгору, відстань  $R$  від вторинного УЗПП5 до УЗП зменшиться. За рахунок цього напруга  $e_{П}$  сприймаючого сигналу збільшиться, що призведе до того, що різниця між поточним значенням напруги  $e_{П}$  і  $e_{0П}$  стане більшою нуля. При цьому компаратор відпрацьовуватиме цю різницю до тих під, поки не установить таку амплітуду ЕЗІ, при котрій різниця напруг  $e_{П}$  і  $e_{0П}$  стане рівною нулю. При цьому запам'ятовується значення амплітуди ЕЗІ, сприймаючого сигналу  $e_{П}$ , подається одиничний ЕЗІ на первинний УЗПП1, включається ЛТІ і цикл роботи повторюється.

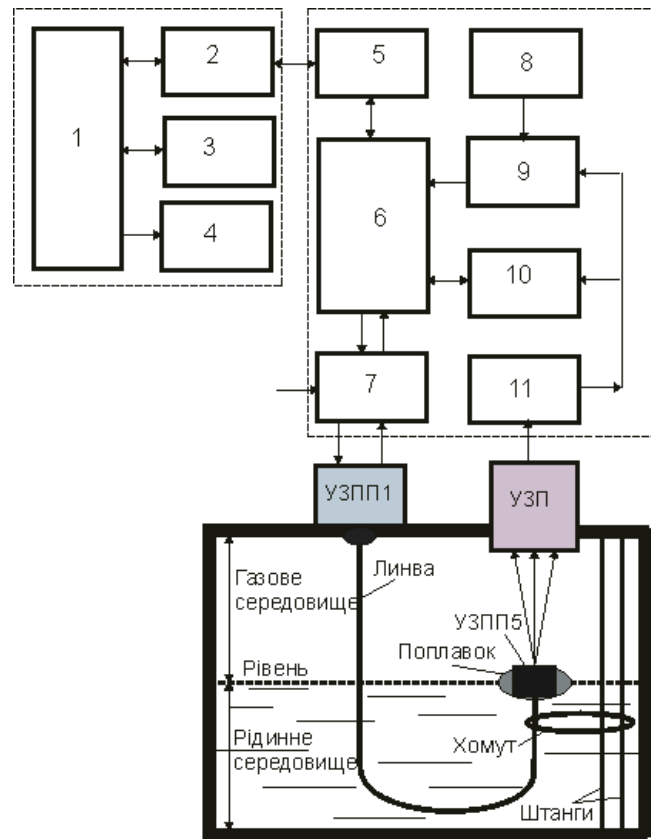


Рисунок2 – Блок-схема вимірювального контролю рівня ультразвуковим засобом з кільцевим проходженням сигналу

### Висновки

За рахунок кільцевого проходження ультразвукового сигналу засіб контролю рівня дозволяє зменшити зону нечутливості ультразвукового пристрою контролю рівня РС практично до 5 разів за рахунок відсутності спеціальної реперної пластини та розділення каналів випромінюючого й сприймаючого УЗІ, збільшити діапазон вимірювального контролю рівня РС до 2-х разів за рахунок одноразового проходження УЗІ газового середовища, підвищити точність вимірювального контролю рівня РС практично вдвічі за рахунок: зменшення відстані проходження УЗІ у ГС в 2 рази, зменшення вторинних ультразвукових ефектів всередині резервуару за рахунок відсутності реперної пластини на шляху проходження УЗІ в ГС.

### Література

1. Шкарлет Ю.М. Бесконтактные методы ультразвукового контроля. – М.: Машиностроение, 1974. – 57 с.
2. Шрайбер Д.С. Ультразвуковая дефектоскопия. М.: Металлургия, 1965. – 391 с.
3. Ермолов И.Н. Отражение ультразвука от дефектов различной формы. – Дефектоскопия, 1970, №4. – С. 17-24.
4. Бабинов О.И. Контроль уровня с помощью ультразвука. – Л.: Энергия, 1971. – 98 с.
5. Дж. Фрайден. Современные датчики. Справочник / Перевод с английского Ю.А. Заболотной под ред. Е.Л. Свинцова. – Москва: Техносфера, 2006. – 592 с.

6. А.Е. Колесников. Ультразвуковые измерения. – 2-е изд., перераб. и доп. – М.: Изд-во стандартов, 1982. – 248 с., ил.
7. Патент (А.С.) 564538, G01F23/28, 21.10.1977 «Ультразвуковой уровнемер с цифровым отсчетом».
8. Патент (А.С.) 2195635, G01F23/28? 21.02.2002 «Способ измерения уровня жидких и сыпучих сред».
9. Рябіченко А.В., Стенцель Й.І. Математичні моделі ультразвукового рівнеміра рідини з компенсуючим вхідним сигналом. Вестник национального технического университета «ХПИ», - 2008, № 48. С. – 61 – 65.
10. Рябіченко А.В., Стенцель Й.І. Математичні моделі ультразвукового рівнеміра рідини з компенсатором вихідних сигналів //Вісник національного технічного університету «ХПИ». Збірник наукових праць. Тематичний випуск: Прилади та методи неруйнівного контролю. – Харків: НТУ «ХПИ» - № 48. – 2008. – с. 61-65.
11. Стенцель Й.І., Рябіченко А.В., Літвінов К.А. Фізичні моделі перетворень в ультразвукових засобах контролю рівня рідинних середовищ. Вісник Національного технічного університету «Харківський політехнічний інститут». Збірник наукових праць. «Електроенергетика и преобразовательная техника». – Харків: НТУ «ХПИ» -№19. – 2011. – С. 3-18.
12. Стенцель Й.І., Євсюков В.В., Шаповалов О.І. Математичні моделі ультразвукового засобу контролю рівня рідинних середовищ за реперним каналом. Вісник Національного технічного університету «Харківський політехнічний інститут». Збірник наукових праць. «Електроенергетика и преобразовательная техника». – Харків: НТУ «ХПИ» -№19. – 2011. – С. 19-30.

#### **Інформація про авторів**

**Стенцель Йосип Іванович** – докт. техн. наук, професор, завідувач кафедри комп'ютерно-інтегрованих систем управління Технологічного інституту Східноукраїнського національного університету імені Володимира Даля.

**Літвінов Костянтин Анатолійович** - студент Технологічного інституту Східноукраїнського національного університету імені Володимира Даля.

**Рябіченко Антон Вікторович** – канд. техн. наук, старший науковий співробітник кафедри комп'ютерно-інтегрованих систем управління Технологічного інституту Східноукраїнського національного університету імені Володимира Даля.

**Шаповалов Олексій Іванович** – аспірант кафедри комп'ютерно-інтегрованих систем управління Технологічного інституту Східноукраїнського національного університету імені Володимира Даля.

**МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ТА ОБЧИСЛЮВАЛЬНІ МЕТОДИ**

УДК 519.863:519.812 (045)

М.В. КУКЛІНСЬКИЙ

Національний авіаційний університет, м. Київ

**ОБґРУНТУВАННЯ ВИБОРУ ПОКАЗНИКІВ ТА ЇХ АРГУМЕНТІВ В ЗАДАЧІ БАГАТОКРИТЕРІАЛЬНОГО СИНТЕЗУ АВІАЦІЙНО-КОСМІЧНОЇ СИСТЕМИ**

**Анотація.** В статті розглядаються питання багатокритеріальної оптимізації авіаційно-космічної системи з позиції досконалості її масово-траєкторних параметрів та модифікацій. Вирішення цих питань, дасть змогу суттєво зменшити час проектування системи та фінансові витрати на її реалізацію.

**Ключові слова:** авіаційно-космічна система, масово-траєкторні параметри, багатокритеріальна оптимізація, економічна ефективність, вартість.

**Аннотация.** В статье рассматриваются вопросы многокритериальной оптимизации авиационно-космической системы с позиции совершенства ее массово-траекторных параметров и модификаций. Решение этих вопросов, позволит существенно уменьшить время проектирования системы и финансовые затраты на ее реализацию.

**Ключевые слова:** авиационно-космическая система, массово-траекторные параметры, многокритериальная оптимизация, экономическая эффективность, стоимость.

**Annotation:** In the article the issues of multi-criteria optimization for aerospace system are considered from the standpoint of perfection its mass and trajectory parameters and modifications. Solution to these issues will allow significant minimization of time for system design and expenditures for its realization.

**Keywords:** aerospace system, mass and trajectory parameters, multi-criteria optimization, economic efficiency, costs.

**Вступ**

Сучасний рівень розвитку світової авіації та космонавтики свідчить про те, що даний напрямок діяльності є одним з найпріоритетніших. Його цільова складова набуває ключової ролі в забезпеченні супутникового спостереження, засобів зв'язку та передачі даних, космічної розвідки, прогнозування метеорологічних явищ та інше. Надалі значення космічних досліджень у вирішенні всього комплексу науково-дослідних, господарських, а також військових завдань тільки зростатиме.

Необхідність значного фінансування космічної галузі визначила досить вузьке коло світових держав, здатних розробляти і виробляти космічні апарати різного цільового призначення, а також засоби їх виведення на орбіту. Зважаючи на стратегічне значення, космічна галузь в цих державах оголошена національним пріоритетом.

Україна на сьогодні входить в перелік країн, що здатні створювати власні космічні багатоцільові проекти. Безумовно, одним з найважливіших завдань при цьому є забезпечення високого рівня конкурентоспроможності на світовому ринку. Проте успішне вирішення цієї задачі, насамперед, визначається правильно сформульованими цільовими показниками, для досягнення яких має бути розроблений відповідний науково-технічний підхід.

**Аналіз публікацій**

Аналіз публікацій [1, 2, 3, 4] довів, що вартість створення і експлуатації системи, яка використовує відповідні технології в інтересах країни, може досягти помітного економічного ефекту лише за умови використання авіаційно-космічних систем (АКС) як засобів доставки корисного вантажу на орбіти. Такі системи дадуть можливість:

- суттєво знизити питому вартість запуску до 900 – 1200 дол./кг, що набагато менше, ніж у одно-разових ракетноносіїв (РН);
- швидкого і оперативного розгортання супутникових угруповань;
- запуск системи з будь-якої широти;
- як часткове, так і повністю багаторазове використання системи;
- відсутність обов'язкової прив'язки до космодромів та ін.

Ці твердження вже підтверджує авіаційно-космічна система «Pegasus»/L-1011, яка на теперішній час експлуатується [5].

**Актуальність та постановка проблеми**

Серед найважливіших цільових завдань, які стоять перед розробниками і виробниками авіаційно-космічної техніки, є проблеми, які пов'язані з суттєвим скороченням часу розробки цієї техніки і зниженням її вартості. Від моменту початку розробки, коли закладаються найважливіші концептуальні характеристики авіаційно-космічної системи, до появи її в експлуатації проходить великий період часу (десятиріччя), протягом якого можуть істотно змінитися основні чинники, які спочатку висувалися до неї. Наприклад, величина попиту, значення потрібних льотно-технічних і економічних критеріїв та ін.

Україна володіє технічним і виробничим арсеналом для створення авіаційно-космічних систем в найкоротші терміни та з мінімальними витратами. Це обумовлено, перш за все, наявністю таких

транспортних літаків, як Ан-124 «Руслан» і Ан-225 «Мрія», які можуть бути використані в авіаційно-космічній системі як перші ступені, а також ракет різного класу, як засобів доставки корисного вантажу на орбіти.

Враховуючи значний прогрес в галузі обчислювальної техніки і математичного забезпечення, основним заходом щодо зниження вартості проектування і розробки авіаційно-космічних систем є впровадження методів автоматизованого проектування.

Застосування таких методів, які базуються на використанні відповідних математичних моделей, є актуальним науковим завданням, яке дозволить дослідити властивості системи з позицій досконалості компоновки, можливостей конструкції, льотних і експлуатаційних даних та масово-траєкторних параметрів (МТП) при різних варіантах її побудови.

#### **Мета і завдання дослідження**

Застосування та перехід до авіаційно-космічних систем, як до багаторазових, неминучий і об'єктивно обумовлений зовнішніми і внутрішніми причинами. До зовнішніх відносяться причини економічного та екологічного характеру, а також причини, які виникають при появі нових завдань покладених на АКС. До внутрішніх, перш за все, відносяться причини пов'язані із структурними змінами в космонавтиці. У будь-якому випадку розвиток АКС супроводжуватиметься постійною конкуренцією з одноразовими засобами вертикального старту. І цьому є просте пояснення – одноразові засоби більш дешеві, і в переважній більшості випадків, економічно ефективніші.

Хоч космічні засоби горизонтального старту мають ряд переваг, які ставлять їх попереду засобів які стартують вертикально, однак, вони поступаються останнім по вантажопідйомності і вимагають великих початкових капіталовкладень для свого створення. Тому опрацьовуючи різні концепції побудови авіаційно-космічних систем, насамперед, необхідно враховувати її економічну ефективність, як визначальний критерій, і ставити питання про необхідність проведення розрахунків його вартісних показників.

Тому метою даної роботи є знаходження таких вартісних показників визначальних критеріїв, які забезпечать мінімізацію витрат на розробку АКС за умови виконання поставленої на неї цільових програм.

Для досягнення поставленої мети в роботі вирішуються такі задачі:

- побудова математичної моделі визначення області компромісів визначальних критеріїв системи;
- пошук відповідних масово-траєкторних параметрів критеріїв з урахуванням особливостей авіаційно-космічної системи;
- пошук раціонального варіанту побудови АКС шляхом оптимізації її визначальних критеріїв.

#### **Визначення основного показника економічної ефективності**

Аналіз робіт присвячених критеріальним дослідженням перспектив розвитку авіаційно-космічної техніки [6, 7, 8] показує, що оптимізація її основних тактико-технічних характеристик (ТТХ) розглядається в аспекті широко поширеної тріади критеріїв: ефективність ( $E$ ) – вартість ( $S$ ) – час ( $T$ ). Причому частіше в якості головного критерію цієї тріади використовується вартість (економічна ефективність), а два інших критерії виступають в якості обмежень. Це викликано завданням мінімізації ресурсів під час виведення корисного навантаження в космос і проведення різних операцій в космічному просторі.

Тому основним показником критерію економічної ефективності є вартість пуску (виведення) [9, 10]. Цей показник також має особливе значення при проведенні порівняння альтернативних варіантів побудови авіаційно-космічних систем, з метою виявлення їх економічних переваг. І саме цей показник прийнято розглядати для оцінки критерію економічної ефективності в критеріальних дослідженнях.

Аналіз існуючих методик розрахунку вартості пуску [9, 10, 11] показує, що для отримання об'єктивної оцінки, крім прямих витрат, які безпосередньо пов'язані з проведенням пуску (вартість матеріалів і палива, які витрачаються в кожному пуску; витрати на виробництво елементів, які використовуються одноразово; вартість амортизації багаторазових елементів, тощо) також повинні бути враховані витрати, пов'язані з підтримкою в робочому стані всіх систем забезпечення пуску, а також, з урахуванням ймовірнісної оцінки, вартість засобів, які втрачаються в аваріях.

Таким чином, виходячи з вищесказаного, сумарна вартість пуску включає такі складові:

- вартість виробництва одноразових елементів ( $S_{o,e}$ ) – ступенів ракетноносіїв, двигунів, зовнішніх паливних баків, розгінних блоків, буксирів та інших одноразових елементів;
- вартість амортизації багаторазових засобів ( $S_{am}$ ) – літаків-носіїв, орбітальних літаків, космічних апаратів, спеціальних цільових навантажень;

- вартість транспортування засобів виведення до місця старту як від заводу-виробника, що припадає на один пуск, так і від місця попередньої посадки ( $S_{mp}$ );
- вартість витратних матеріалів ( $S_{в.м}$ ) – компонентів палива, рідин і газів з урахуванням їх втрат на всіх етапах від транспортування із заводів до старту системи;
- віднесена до одного пуску вартість експлуатації ( $S_e$ ) багаторазових засобів виведення, наземного технічного комплексу всіх елементів системи в цілому, стартового і посадкового комплексу, наземного комплексу управління; вартість залучення літаків-заправників для АКС, пошуково-рятувальної служби; вартість відчуження землі під поля падіння ступенів ракетноносіїв; вартість екологічних заходів з очищення територій в місцях падіння залишків засобів виведення з невиробленими компонентами палива;
- вартість страховки засобів виведення і корисних вантажів від ймовірних втрат в аваріях, затримок запуску корисних вантажів або їх відмов у процесі виведення ( $S_{стп}$ );
- резерв на невраховані витрати ( $S_{рез}$ ) при підготовці та проведенні пуску.

Загальний вигляд вартості пуску виглядатиме:

$$S_{пуск} = S_{о.е} + S_{ам} + S_{mp} + S_{в.м} + S_e + S_{стп} + S_{рез} \cdot \quad (1)$$

Причому аналіз цих складових показав, що аргументами кожної із них є масово-траєкторні параметри системи.

### Постановка задачі оптимізації показника економічної ефективності АКС

Загалом авіаційно-космічну програму можна розділити на три основних етапи:

- науково-дослідні та дослідно-конструкторські роботи (НДДКР), включаючи виготовлення дослідних зразків і льотні випробування;
- виробництво аерокосмічної техніки і необхідної інфраструктури;
- експлуатація парку та інфраструктури.

Виходячи з цих етапів можна побачити, що при розрахунку вартості пуску враховуються тільки другий і третій етапи, перший залишається без уваги. Крім цього, розглянутий вартісний показник враховує тільки повні витрати на створення та експлуатацію проектного зразка протягом заданого часу. Однак повний обсяг завдань може бути оцінений досить і досить наближено. Тому доцільно порівняльну оцінку економічної ефективності варіантів АКС здійснювати на тлі цілком конкретного обсягу розрахункових завдань (РЗ).

Так, наприклад, можна розробити такий номінальний варіант побудови АКС, який зміг би виконувати весь перелік розрахункових завдань. Очевидно, що він буде затребуваний довгі роки, але чи буде він по витраченим ресурсам задовольняти критерію економічної ефективності? А можна розробити варіант побудови АКС під конкретну розрахункову задачу, який, безсумнівно, буде задовольняти критерію економічної ефективності, проте чи буде він затребуваний надалі, в умовах постійно мінливих вимог до розрахункових завдань?

Тому необхідно знайти такий раціональний варіант побудови авіаційно-космічної системи, економічна ефективність якого з номінального розрахункового завдання оптимально задовольняла б конкретному варіанту розрахункового завдання. Схематично дана постановка задачі представлена на рис.1.

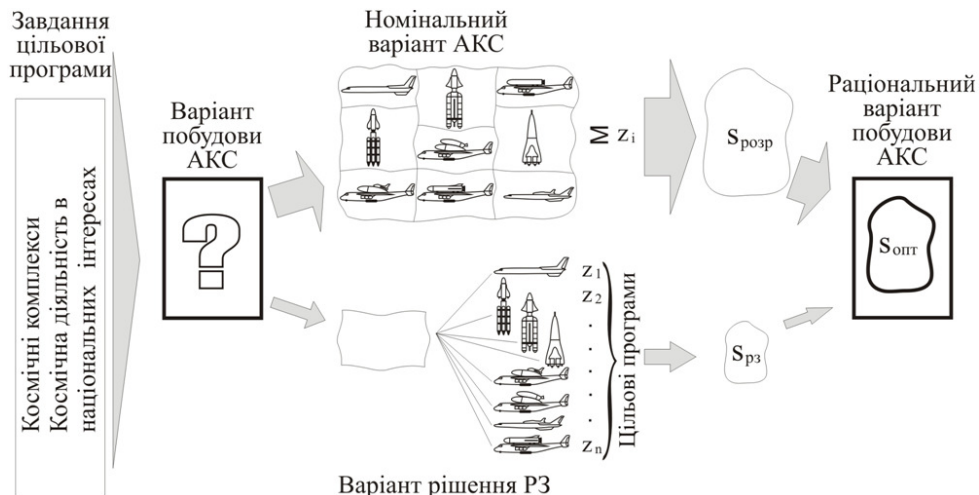


Рисунок 1 – Схематичне представлення постановки задачі

В даному випадку знаходження раціонального варіанту побудови авіаційно-космічної системи, означає обґрунтування значень її масово-траєкторних параметрів, яке виконується за допомогою оптимізації критерію економічної ефективності системи.

### Математична модель процесу обґрунтування основних критеріїв оптимізації

Як було зазначено вище, основним показником економічної ефективності АКС на сьогодні вважається вартість пуску ( $S_{пуск}$ ) (виведення) яка визначається сумою (1). Але беручи до уваги те, що цей показник враховує повні витрати, які можуть бути оцінені дуже наближено, його не можна розглядати як один цілісний показник, і відповідно застосовувати до нього методи однокритеріальної оптимізації.

Виходячи з постановки задачі оптимізації показника економічної ефективності АКС до розгляду пропонується не один показник ( $S_{пуск}$ ), а два показники які включають у себе поняття вартості: ( $S_{розр}$ ) – вартість розробки номінальної АКС і ( $S_{рз}$ ) – вартість АКС для вирішення конкретного розрахункового завдання. Розгляд саме цих двох показників дозволить надалі враховувати витрати на розробку, виготовлення дослідних зразків і льотні випробування АКС, а також дозволить робити порівняльну оцінку економічної ефективності варіантів АКС, які включають в себе повний обсяг завдань (номінальних варіантів) на тлі цілком конкретного обсягу РЗ.

Так як затрати на НДДКР вважаються пропорціональними вартості першого серійного зразка [12] то показник  $S_{розр}$  можна записати у вигляді:

$$S_{розр} = K_{НДДКР} S_I, \quad (2)$$

де  $K_{НДДКР}$  – коефіцієнт науково-дослідних та дослідно-конструкторських робіт;  $S_I$  – вартість першого серійного зразка.

Значення коефіцієнта НДДКР може бути призначено шляхом експертної оцінки, визначено шляхом перерахунку від конструкції-прототипу або за допомогою прямих економічних розрахунків.

Вартість першого зразка (в якій враховуються етапи науково-дослідних робіт, проектування, виготовлення дослідних зразків і їх випробування, розробка конструкторської документації, виготовлення матеріальної частини зразків) насамперед залежить від його складності і розмірності. Досліди показують, що від розмірності можна відволіктися, якщо за основу взяти питому вартість. Це дозволяє застосувати аналогово-порівняльний метод, на основі аналізу подібних виробів різної розмірності.

Статистичний аналіз питомих вартостей виготовлення типових виробів аерокосмічної техніки [12] показує, що цей параметр в середньому доволі сталий, тому помноживши питому вартість на суху масу виробів можна отримати вартість першого серійного зразка.

Так як значення ( $K_{НДДКР}$ ) і ( $S_I$ ) можуть бути різними для елементів однієї і тієї ж АКС виконаної в різних модифікаціях, то при розрахунку показника ( $S_{розр}$ ) необхідно враховувати всі особливості цих модифікацій.

Щодо розрахунку вартості рішення розрахункової задачі ( $S_{рз}$ ) то тут, насамперед, необхідно найбільш повно враховувати всі витрати. Вони повинні розглядатися з урахуванням складових, як безпосередньо пов'язаних з проведенням пуску, так і складових, які визначають підтримку всіх систем забезпечення пуску. Також повинні враховуватися вартість безповоротних втрат в процесі вирішення розрахункового завдання ( $S_{бв}$ ), і витрати на космічний апарат для вирішення поставленого розрахункового завдання ( $S_{ка}$ ).

Виходячи з вищесказаного в показник ( $S_{рз}$ ) будуть включатися всі доданки суми (1), з урахуванням ( $S_{бв}$ ) і ( $S_{ка}$ ), отже:

$$S_{рз} = S_{о.е} + S_{ам} + S_{тп} + S_{е.м} + S_e + S_{стп} + S_{рез} + S_{бв} + S_{ка}. \quad (3)$$

Тут ( $S_{бв}$ ) визначається за фактичними витратами пов'язаними зі зниженням потенційних можливостей АКС, наприклад, скороченням льотного ресурсу, витратами пов'язаними з коштами цільового призначення, які використовуються при вирішенні розрахункового завдання і т.д. Витрати ( $S_{ка}$ ) залежать від маси КА, від вартості виготовлення КА, від маси корисного вантажу, який виводиться

на орбіту, від параметрів орбіти виведення та сумарного числа пусків системи за весь період застосування АКС.

### Визначення основних аргументів критеріальних функцій

Як правило, всі необхідні вихідні дані, які використовуються в запропонованій методиці для пошуку ( $S_{розр}$ ) і ( $S_{рз}$ ), є в проектах космічних систем, в статистичній інформації з експлуатації застосування аналогічних систем або можуть бути отримані розрахунковим шляхом із застосуванням методів економіко-математичного моделювання процесів створення і застосування систем даного класу. До того ж всі аргументи вартісних функцій є масово-траєкторними параметрами, які беруться з переліку основних тактико-технічних характеристик системи.

Проте застосування всіх МТП в критеріях ( $S_{розр}$ ) і ( $S_{рз}$ ), привів би до невиправданого ускладнення критеріальних функцій, а також до надмірних труднощів вирішення завдань оптимізації. Тому природним є пошук найбільш інформативних МТП – координат критеріального простору, в якому буде здійснюватися екстремізація введених критеріїв.

З точки зору ефективного застосування, як частково, так і повністю багаторазових АКС, з усього переліку визначальних параметрів запропонованої вартісної моделі, до таких параметрів відносяться:

- $x_1$  – маса корисного вантажу, що виводиться на орбіту  $m_{кв}$ ;
- $x_2$  – сумарна маса одноразових елементів  $m_{\Sigma}^I$ ;
- $x_3$  – маса палива, що заправляється (з урахуванням дозаправки)  $m_{пал}$ ;
- $x_4$  – льотний ресурс системи  $R_{акс}$ ;
- $x_5$  – кут нахилу траєкторії у момент розділення ступенів ракето-носія  $\Theta_{рн}$ .

Для оцінки вартості можуть використовуватися і більш детальні моделі, однак це доцільно лише на останніх стадіях проектування.

Вибір саме цих МТП, насамперед, пояснюється основними вимогами які ставляться до математичних моделей оптимізації, а саме:

- всі вибрані МТП враховують можливість використання найбільш простих методів їх аналізу;
- всі вони можуть бути нормовані і виражені у вартісному вигляді;
- всі вони мають виражену область існування і для них однозначно можуть бути визначені обмеження;
- вони здатні забезпечити заданий рівень якості системи і ефективність виконання розв’язуваних нею завдань;
- всі вибрані МТП з достатньою повнотою і достовірністю характеризують АКС, а також дозволяють понизити розмірність розв’язуваної задачі, так як більшість параметрів, які визначають вартісну математичну модель, можуть бути однозначно визначені через запропонований набір МТП. Наприклад:
  - $m_{кв}$  та  $\Theta_{рн}$  визначають висоту орбіти, на яку виводиться КА;
  - $m_{пал}$  є сумою маси компонентів витрачених матеріалів і маси компонентів які заправляються з літака-заправщика, з урахуванням втрат на всіх етапах від заправки в літаку-заправщику до заправки в літаку-носії;
  - $m_{\Sigma}^I$  є сумарною масою всіх одноразових елементів;
  - через  $R_{акс}$  визначається період застосування засобів АКС і ресурс елементів системи в польотах, та інше. Решта визначальних параметрів моделі є або статистичними коефіцієнтами, або задаються експертно.
- якість системи може бути оцінена за сукупністю приватних критеріїв ( $S_{розр}$ ) і ( $S_{рз}$ ), що представляють собою функції від обраних МТП:

$$\begin{aligned} S_{розр} &= S_{розр}(m_{кв}, m_{\Sigma}^I, m_{пал}, R_{акс}, \Theta_{рн}) \Rightarrow f_1(x_1, x_2, x_3, x_4, x_5), \\ S_{рз} &= S_{рз}(m_{кв}, m_{\Sigma}^I, m_{пал}, R_{акс}, \Theta_{рн}) \Rightarrow f_2(x_1, x_2, x_3, x_4, x_5). \end{aligned} \quad (4)$$

Причому ці критерії знаходяться в тісному зв’язку і суперечності один з одним.

### Суперечливість критеріальних функцій

Так як авіаційно-космічна система розробляється для виконання широкого спектру розрахункових завдань, а реалізується під конкретне завдання що входить в цей спектр (рис. 1), критерії ( $S_{розр}$ ) і ( $S_{рз}$ ) не можна підсумовувати, а тому вони є суперечливі. Критерії ( $S_{розр}$ ) і ( $S_{рз}$ ) можуть підсумовуватися лише у тому випадку, коли авіаційно-космічна система розробляється і реалізується тільки під один спектр розрахункових завдань (рис. 2). У решті випадків ( $S_{розр}$ ) і ( $S_{рз}$ ) будуть суперечливими (рис. 3).

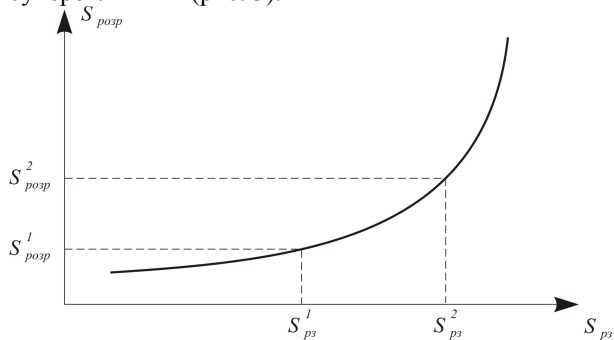


Рисунок 2 – Залежність  $S_{розр}$  від  $S_{рз}$  у разі несуперечливості (однокритеріальна задача)

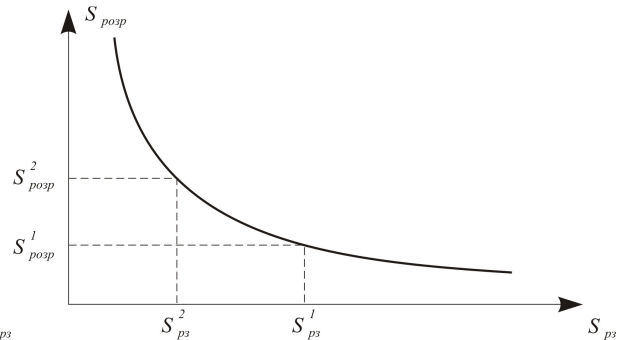


Рисунок 3 – Залежність  $S_{розр}$  від  $S_{рз}$  у разі суперечливості (багатокритеріальна задача)

Суперечливість критеріїв ( $S_{розр}$ ) і ( $S_{рз}$ ) можна простежити ще й на прикладі того, що ( $S_{розр}$ ) буде зменшуватися з часом і удосконаленням системи в межах одного класу. Так як в кожний новий варіант реалізації АКС вже будуть включатися дані з попередніх варіантів, відповідно вартісні ресурси витратяться на них не будуть. І навпаки ( $S_{рз}$ ) буде збільшуватися з удосконаленням системи. Тому оптимізуючи систему в межах цих двох критеріїв одночасно, неможливо буде зменшити один, без збільшення іншого.

Також у разі суперечливості (рис. 3) спостерігається ще одна ситуація, в якій із зменшенням витрат на розробку, витрати на реалізацію ( $S_{рз}$ ) будуть збільшуватися.

Це не говорить про те, що спектр розв'язуваних розрахункових завдань збільшиться, витрати збільшаться у зв'язку з дефектами і поломками які виникатимуть при реалізації, і для усунення яких будуть витратитися додаткові кошти. Відповідно при збільшенні витрат на розробку ( $S_{розр}$ ) багатьох дефектів під час реалізації можна буде уникнути, за рахунок чого знизяться втрати ( $S_{рз}$ ). У цьому випадку можна говорити про більш «гнучкі» і функціональні варіанти побудови АКС, так як АКС може розроблятися під великий спектр вирішуваних завдань, а реалізовуватися під певну РЗ з цього спектра.

Враховуючи суперечливість критеріїв ( $S_{розр}$ ) і ( $S_{рз}$ ) знаходження оптимального варіанта побудови АКС потребує застосування до них методів багатокритеріальної оптимізації. Необхідно також відзначити, що всі вищевикладені рекомендації та пояснення про несуперечливість і суперечливість критеріїв ( $S_{розр}$ ) і ( $S_{рз}$ ) та графічні пояснення до них відносяться і розглядаються на області оптимальних значень цих критеріїв.

### Висновок

Таким чином, можна сказати, що запропонований набір МТП визначає генеральну ідею створення АКС, яка може бути виражена через деякі критерії ефективності, визначені на множині цих МТП. Це дозволяє сформулювати парето-оптимальну множину варіантів побудови системи [13].

Запропонований підхід обґрунтування масово-траєкторних параметрів авіаційно-космічної системи також дозволяє сформулювати рекомендації при проектуванні і розробці авіаційно-космічних систем, оцінка якості яких відбувається по декількох критеріях.

Використання інформації про відносну важливість критеріїв [14] з запропонованим підходом дозволяє звузити область рішень, що шукаються, у декілька разів, а це дає можливість істотно мінімізувати витрати на розробку та доведення системи з урахуванням її цільових програм.

### Література

1. Авдуевский В. С. Народнохозяйственные и научные космические комплексы / В. С. Авдуевский, Г. Р. Успенский. – М. : Машиностроение, 1985. – 416 с.

2. Ануреев И. И. Ракеты многократного использования / И. И. Ануреев. – М. : Воениздат, 1975. – 214 с.
3. Гуцин В. Н. Основы устройства космических аппаратов : учебник для вузов / В. Н. Гуцин. – М. : Машиностроение, 2003. – 272 с.
4. Дмитриев В. Г. Проблемы создания перспективной авиационно-космической техники / В. Г. Дмитриев. – М. : Физматлит, 2005. – 648 с.
5. Pegasus User's Guide : Release 4.0 / Orbital Sciences Corporation, 1998. – 175 с.
6. Лебедев А. А. Основы синтеза систем летательных аппаратов : учебное пособие для студентов вузов / [Лебедев А. А., Баранов В. Н., Бобронников В. Т. и др.] ; под ред. А. А. Лебедева. – М. : Машиностроение, 1987. – 224 с.
7. Проектирование самолетов / [Егер С. М., Мишин В. Ф., Лисейцев Н. К. и др.] ; под ред. С. М. Егера. – [3-е изд.]. – М. : Машиностроение, 1982. – 616 с.
8. Егер С. М. Основы автоматизированного проектирования самолетов : учебное пособие для студентов авиационных специальностей вузов / С. М. Егер, Н. К. Лисейцев, О. С. Самойлович. – М. : Машиностроение, 1986. – 232 с.
9. Методика оценки и анализ стоимости пуска и стоимости транспортно-технического обслуживания производства в космосе / НИИ Экономики, 1990.
10. Методическая схема исследования технико-экономической эффективности одноразовых и многоразовых средств выведения / ВИМИ, 1991.
11. Отчет о научно-исследовательской работе «Технико-экономические исследования в обоснование создания авиационно-космической системы» : Методика оценки стоимости пуска (этап 02) / НИР «Вега-91», 1991.
12. Модель TRANSCOST для оценки стоимости ракеты-носителя и ее применение для анализа будущих систем : Экспресс-информация ВИНТИ // Астронавтика и ракетодинамика. – 1986. – № 12.
13. Куклинский М. В. Формирование парето-оптимального множества вариантов построения сложной технической системы / Максим Куклинский // Научно-виробничий збірник «Наукові записки УНДІЗ». – 2013. – №3 (27). С. 59–63.
14. Куклінський М. В. Проблеми звуження парето-оптимальної множини варіантів побудови авіаційно-космічної системи / Максим Куклінський // Збірник наукових праць «Проблеми інформатизації та управління». – 2013. – № 2 (42). С. 56–60.

#### Відомості про авторів

**Куклінський Максим Володимирович** – к.т.н., доцент кафедри Комп'ютерних інформаційних технологій, Національний авіаційний університет, корпус №6 кімн.206.

УДК 519.711

Б. І. МОКІН, А. В. ПИСКЛЯРОВА, О. Б. МОКІН

### МАТЕМАТИЧНА МОДЕЛЬ ПРОЦЕСУ САМОСТІЙНОГО ЗАСВОЄННЯ СТУДЕНТОМ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ У МІЖЛЕКЦІЙНИЙ ПЕРІОД

**Анотація.** Досліджено процес самостійного засвоєння студентом навчальної дисципліни у міжлекційний період і побудована математична модель цього процесу.

**Ключові слова:** студент, навчальна дисципліна, синергетичний ефект, математична модель.

**Аннотация.** Исследован процесс самостоятельного усвоения студентом учебной дисциплины в межлекционный период и построена математическая модель этого процесса.

**Ключевые слова:** студент, учебная дисциплина, синергетический эффект, математическая модель.

**Abstract.** Investigated a process of the independent mastering of educational discipline between lectures by the student and built mathematical model of this process.

**Key words:** student, educational course, synergetic effect, mathematical model.

#### Постановка задачі і вихідні передумови

В роботі [1] з використанням тижневої «кривої забування» Г. Еббінгауза [2, 3], яка у вигляді таблиці може бути задана і так нами досліджено вплив синергетичної складової математичної моделі процесу засвоєння студентом навчальної дисципліни.

Таблиця 1

t (години)	0	2	12	24	72	144	156	166	168
$x_1$ (%)	100	85	40	34	25	20	19	18,2	18

Дослідження здійснено з використанням результатів, отриманих в роботі [4], у якій ця математична модель синтезована у вигляді трьох пар нелінійних диференціальних рівнянь, із яких для дослідження впливу синергетичної складової було використане лише одне диференціальне рівняння, а саме:

$$\frac{dx_1}{dt} = -\alpha_{11}x_1 + \alpha_{12}x_1x_2, \quad (1)$$

де  $\alpha_{11}$  – параметр, що характеризує ступінь забування студентом матеріалу навчальної дисципліни, вивченого в аудиторії з викладачем,  $\alpha_{12}$  – параметр, що характеризує синергетичний вплив одна на одну складових процесу засвоєння студентом навчального матеріалу в аудиторії з викладачем та самостійно, а  $x_1, x_2$  – фазові координати, що задають у відносних одиницях ступінь засвоєння студентом програми навчальної дисципліни відповідно на заняттях в аудиторії з викладачем та самостійно, для яких виконуються умови:

$$x_1 = \frac{X_1}{X}, x_2 = \frac{X_2}{X}, \quad (2)$$

$$\begin{cases} x_1 \leq 1, \\ x_2 \leq 1, \end{cases} \quad (3)$$

$$x_1 + x_2 \leq 1 \quad (4)$$

де  $X$  – це та кількість знань, яку може мати студент, засвоївши протягом визначеного терміну часу  $T$  усі розділи програми певної навчальної дисципліни,  $X_1$  – це така кількість знань з даної навчальної дисципліни, яку студент отримує від викладача під час аудиторних занять, а  $X_2$  – це та кількість знань з даної дисципліни, яку студент засвоює, самостійно вивчаючи певні розділи програми.

В результаті співставлення розв'язків диференціального рівняння (1) та похідного від нього диференціального рівняння без синергетичної складової  $\alpha_{12}x_1x_2$  у роботі [1] показано, що процес засвоєння студентом програми навчальної дисципліни у міжлекційний період 10-тижневого триместру з однією лекцією на кожному тижні не затухає до нуля у кінці першого періоду після першої лекції, а у

пам'яті студента згідно з «кривою забування» Г. Еббінгауза залишається через тиждень після прослуховування лекції (тобто, через 168 годин) 18 % від тієї кількості знань, які він отримав на цій лекції. І тому цей процес може бути описаний функцією, що має вигляд:

$$x_1(t) = 10e^{(-\alpha_1 + \alpha_1 x_2)t}, \quad (5)$$

у якій

$$\begin{aligned} \alpha_{11} &= 0,08013, \\ \alpha_{12}x_2 &= 0,06980 \end{aligned} \quad (6)$$

Таблиця 1 та співвідношення (1) – (6) задають першу вихідну передумову, необхідну для розв'язання задачі синтезу математичної моделі процесу самостійного засвоєння студентом навчальної дисципліни у міжлекційний період, яку ми ставимо перед собою у даній роботі.

В якості другої вихідної передумови використаємо побудовані нами у роботі [1] графіки можливих траєкторій процесу засвоєння студентом навчальної дисципліни протягом 10-тижневого триместру, приведені на рис. 1, на якому цифрою 1 позначено графік процесу засвоєння ним цієї навчальної дисципліни у вигляді «сходинок кривої» для випадку, якби студент в міжлекційний період нічого не забував, цифрою 2 позначено графік процесу засвоєння ним цієї навчальної дисципліни у вигляді «пилкоподібної кривої», яка мала б місце в разі «чистого забування», а цифрою 3 позначено графік процесу засвоєння студентом цієї навчальної дисципліни у вигляді іншої «пилкоподібної кривої», яка має місце згідно з «кривою забування» Г. Еббінгауза.

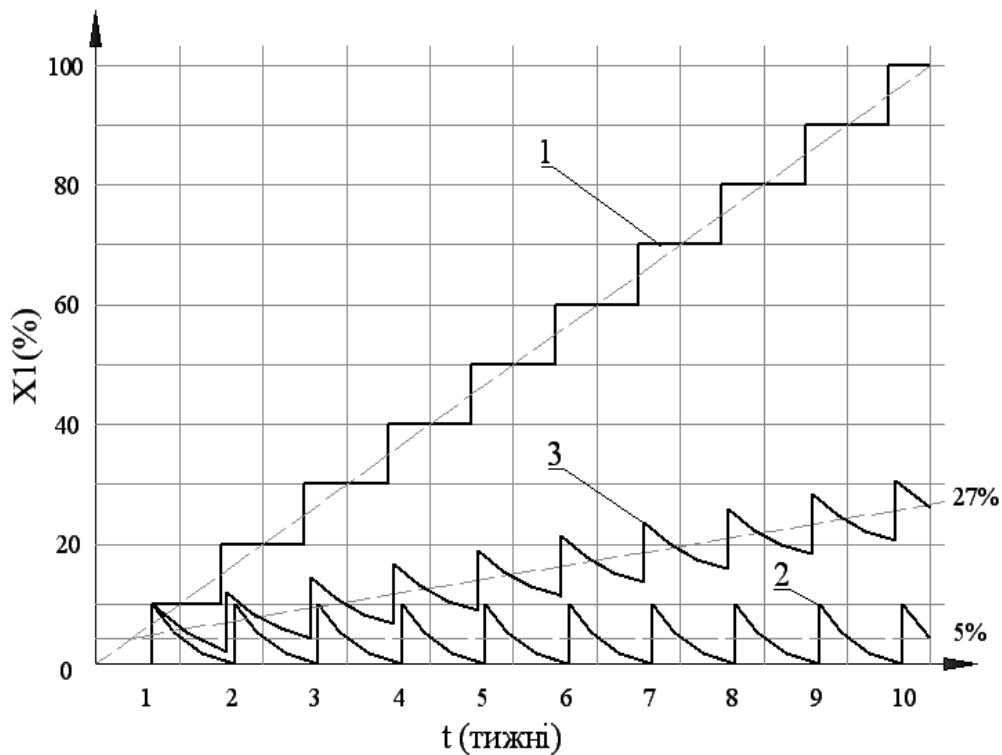


Рисунок 1 – Графіки можливих траєкторій процесу засвоєння студентом навчальної дисципліни протягом 10-тижневого триместру

А в якості третьої вихідної передумови використаємо ще одне із синтезованих у роботі [4] трьох пар нелінійних диференціальних рівнянь, а саме:

$$\frac{dx_2}{dt} = -\alpha_{22}x_2 + \alpha_{21}x_1x_2 + \beta_{22}x_2, \quad (7)$$

де  $\alpha_{22}$  – параметр, що характеризує ступінь забування студентом матеріалу навчальної дисципліни, вивченого самостійно,  $\alpha_{21}$  – параметр, що характеризує синергетичний вплив одна на одну складових процесу засвоєння студентом навчального матеріалу в аудиторії з викладачем та самостійно, а параметр  $\beta_{22}$  характеризує ступінь засвоєння студентом нових знань під час самостійної роботи у міжлекційний період.

#### Розв'язання задачі

Дослідимо, за яких умов графік процесу засвоєння студентом навчальної дисципліни протягом триместру буде наближатись до «сходінкової кривої», позначеної на рис.1 цифрою 1 – адже лише у цьому випадку у кінці триместру студентом буде повністю засвоєна програма цієї навчальної дисципліни.

Оскільки, як показано в роботі [1], у перший міжлекційний період цей процес описується експоненціальною функцією (5), для якої справедливими є умови (6), то легко бачити, що із цієї експоненти наблизитись до «сходінки», для якої повинна виконуватись рівність

$$x_1(t) + x_2(t) = 10(\%) \quad (8)$$

без додаткової генерації складової  $x_2(t)$ , що характеризує процес самостійного вивчення студентом навчальної дисципліни у цей міжлекційний період, не можна. Але, перш ніж перейти до синтезу закону формування цієї складової і його математичної моделі, звернемо спочатку увагу на те, що, як витікає із умови (6), справедливим є наступний вираз –

$$\alpha_{12} = \frac{0,06980}{x_2} \quad (9)$$

А тепер знайдемо значення  $x_2$  та  $\alpha_{12}$  для усіх значень  $x_1$  із таблиці 1, починаючи з другого і закінчуючи шостим, скориставшись тим, що для усіх, визначених у таблиці 1, точок першої міжлекційної «сходінки» на графіку 1 (рис.1) справедливою є рівність (8), а  $x_2$  та  $\alpha_{12}$  зв'язані між собою рівнянням (9). Знайдені значення  $x_2, \alpha_{12}$  приведені у таблиці 2.

Таблиця 2

t (години)	2	12	24	72	144	156	166
$x_1$ (%)	8,5	4,0	3,4	2,5	2,0	1,9	1,82
$x_2$ (%)	1,5	6,0	6,6	7,5	8,0	8,1	8,18
$\alpha_{12}$	0,04653	0,01163	0,01057	0,00931	0,00872	0,00861	0,00853

А на рис. 2 зображено графік залежності  $\alpha_{12} = f(x_2)$ , заданої таблицею 2.

Із цього графіку видно, що з ростом кількості знань  $x_2$ , засвоєних студентом самостійно, значення коефіцієнта синергетики  $\alpha_{12}$  зменшуються, що повністю співпадає з відомим з мікроекономіки [5] законом спадної граничної корисності у разі, якщо цю кількість знань розглядати як «благо», яке студент отримує в процесі вивчення навчальної дисципліни.

Хоча усі викладки ми здійснювали для першого міжлекційного періоду, очевидно, що вони будуть справедливими і для усіх інших

Отже, ми встановили, яку кількість знань у міжлекційний період студент повинен засвоїти самостійно для того, щоб перетворити експоненту у «сходінку». Тож тепер перейдемо до розгляду питання, якою повинна бути модель цього самостійного засвоєння. Для цього спочатку, вважаючи  $x_1$  параметром, проінтегруємо (в межах від моменту часу  $t_n$  – початку самостійного вивчення навчальної дисципліни до  $t$ ) диференціальне рівняння (7), представивши його у вигляді:

$$\frac{dx_2}{x_2} = (-\alpha_{22} + \alpha_{21}x_1 + \beta_{22})dt \quad (10)$$

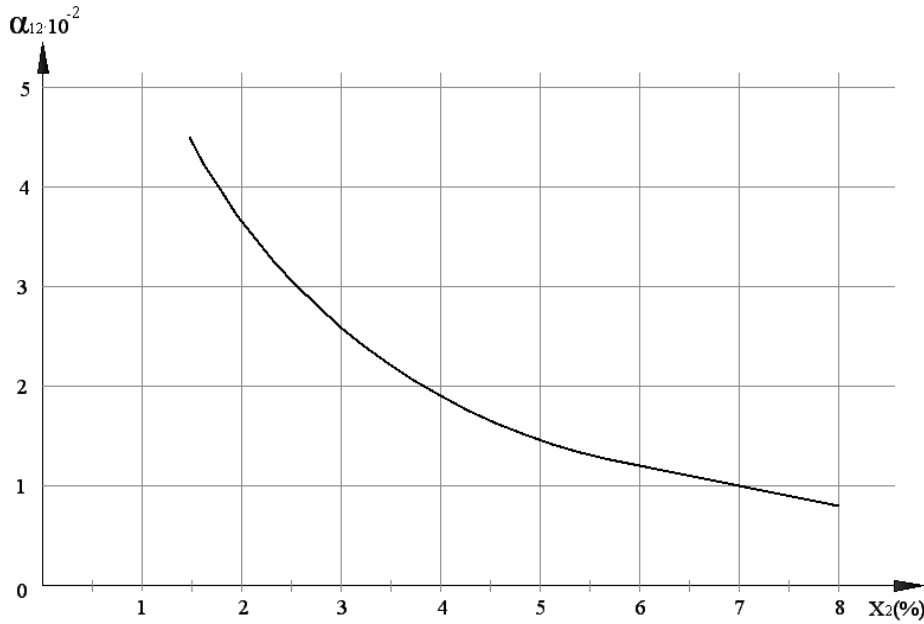


Рисунок 2 – Графік залежності коефіцієнта синергетики  $\alpha_{12}$  від відносної кількості знань  $X_2$ , засвоєних студентом самостійно, побудований згідно з «кривою забування» Г. Еббінгауза

В результаті інтегрування отримаємо:

$$x_2(t) = x_2(t_n) e^{(-\alpha_{22} + \alpha_{21}x_1 + \beta_{22})(t-t_n)} \quad (11)$$

Очевидно, що ця експонента у момент часу  $t_K$  дістанеться до «сходинки» лише за умови, що

$$x_1(t_n) + x_2(t_K) = 10(\%) \quad (12)$$

Далі розглянемо ідеалізований варіант, коли студент в однаковій мірі забуває і те, що сказав викладач, і те, що він прочитав у підручнику, а на процес уповільнення забування знань, отриманих від викладача, знання, отримані самостійно, впливають у тій же мірі, у якій на уповільнення забування знань, отриманих самостійно, впливають знання, отримані від викладача, тобто, покладемо, що

$$\alpha_{11} = \alpha_{22}, \quad (13)$$

$$\alpha_{12} = \alpha_{21}, \quad (14)$$

Для цього випадку, як витікає із виразів (11) – (14), математична модель (11) процесу засвоєння студентом знань самостійно перетворюється в математичну модель

$$x_2(t) = x_2(t_n) e^{(-\alpha_{11} + \alpha_{12}x_1 + \beta_{22})(t-t_n)}, \quad (15)$$

числове значення невідомого параметру  $\beta_{22}$  у якій знаходиться із правої граничної умови, заданої у момент закінчення  $t_K$  самостійного вивчення навчальної дисципліни.

Очевидно, що не буде входити у протиріччя з логікою наступне твердження 1: у пам'яті студента із прослуханої лекції залишається лише те, що він з якоїсь причини згадує.

А із цього твердження витікає, що процес згадування – це теж елемент самостійного вивчення навчальної дисципліни, а тому відносну кількість  $x_1^E(t_n)$  згаданого у момент часу  $t_n$ , взяту із «кривої

забування» Г. Еббінгауза для цього моменту часу, ми маємо право використати і в якості лівої граничної умови для математичної моделі (15), тобто, для моменту часу  $t_n$  маємо умову -

$$x_2(t_n) = x_1^E(t_n). \quad (16)$$

А оскільки шляхом самостійного вивчення навчальної дисципліни у міжлекційний період ми намагаємось у момент часу  $t_K$  домогтися виконання рівності (8), то в якості правої граничної умови для математичної моделі (15) маємо право взяти умову -

$$x_2(t_K) = 10 - x_1^E(t_K), \quad (17)$$

яка витікає із рівності (12).

Далі сформулюємо твердження 2: для самостійного засвоєння розділів навчальної дисципліни кожному студенту в залежності від його індивідуальних здібностей потрібна різна кількість часу.

Практика вивчення авторами цієї статті досвіду самостійного засвоєння окремих розділів кількох навчальних дисциплін студентами Вінницького національного технічного університету свідчить, що забутий матеріал, прослуханий протягом однієї двогодинної лекції, при наявності конспекту самостійно відновлюється і осмислено усвідомлюється здібними студентами протягом відрізка часу

$$t_B = t_K - t_n, \quad (18)$$

не більшого 20 хвилин, студентами середніх здібностей – протягом відрізка часу, не більшого 40 хвилин, і навіть студентами слабких здібностей – на протязі відрізка часу, не більшого однієї години. Це дає нам право при визначенні правої граничної умови задавати значення  $t_K$  у розмірі

$$t_K^{63} = t_n + \frac{1}{3}, \quad (19)$$

для студентів з високими здібностями, у розмірі

$$t_K^{c3} = t_n + \frac{2}{3}, \quad (20)$$

для студентів середніх здібностей та у розмірі

$$t_K^{n3} = t_n + 1, \quad (21)$$

для не здібних студентів.

Що ж до лівої граничної умови, то вона залежатиме ще й від того, коли студент почне самостійно вивчати матеріал, який він почув на лекції, – через 2 години, через 12 годин (тобто, через півдня), через 24 години (тобто, через день), через 48 годин (тобто, через два дні), через 72 години (тобто, через три дні), через 144 години (тобто, за день до наступної лекції), через 156 годин (тобто, увечері напередодні дня заслуховування наступної лекції), чи аж через 166 годин (тобто, за дві години до початку наступної лекції).

Оскільки такий важливий показник як успішність академічної групи визначається відносно усіх студентів, які склали іспит з навчальної дисципліни з оцінкою, не нижчою від «задовільно», характерною для не здібних студентів, то в подальших викладках спиратимемося на характеристики саме не здібних студентів, адже отримані нами результати можна буде легко узагальнити і на аналіз процесу самостійного засвоєння навчальної дисципліни студентами середніх здібностей та студентами з високими здібностями.

В якості значень  $t_n$  виберемо моменти часу через 24 години після прослухування лекції (тобто, через день), через 72 години (тобто, через три дні), через 144 години (тобто, за день до наступної лекції), через 156 годин (увечері напередодні наступної лекції) та через 166 годин (за дві години до наступної лекції). Множина цих значень  $t_n$  матиме вигляд:

$$\{t_n^{(24)}, t_n^{(72)}, t_n^{(144)}, t_n^{(156)}, t_n^{(166)}\}. \quad (22)$$

Згідно з виразом (16) та «кривою забування» Г. Еббінгауза множині початкових моментів часу (22) буде відповідати наступна множина лівих граничних умов  $x_2(t_n^{(*)})$ , виражених у процентах:

$$\{3, 4; 2, 5; 2, 0; 1, 9; 1, 82\}. \quad (23)$$

А множина значень  $t_K$ , сумісних згідно з виразом (21) з відповідними значеннями  $t_n$ , заданими множиною (22), матиме вигляд:

$$\{t_K^{(25)}, t_K^{(73)}, t_K^{(145)}, t_K^{(157)}, t_K^{(167)}\}, \quad (24)$$

Згідно з виразом (17) та «кривою забування» Г. Еббінгауза множині (24) моментів часу  $t_K^{(*)}$ , буде відповідати наступна множина правих граничних умов  $x_2(t_K^{(*)})$ , виражених у процентах:

$$\{6, 61; 7, 51; 8, 01; 8, 11; 8, 18\} \quad (25)$$

Підставляючи у вираз (15) замість поточного  $t$  конкретне значення  $t_K^{(*)}$ , а замість  $t_n$  відповідно конкретизоване значення  $t_n^{(*)}$  і враховуючи вираз (16), отримаємо:

$$x_2(t_K^{(*)}) = x_2(t_n^{(*)}) e^{(-\alpha_{11} + \alpha_{12} x_2(t_n^{(*)}) + \beta_{22}^{(*)})(t_K^{(*)} - t_n^{(*)})} \quad (26)$$

Шляхом логарифмування і простих перетворень із виразу (26) знайдемо, що

$$\beta_{22}^{(*)} = \alpha_{11} - \alpha_{12} x_2(t_n^{(*)}) + \frac{1}{t_K^{(*)} - t_n^{(*)}} \ln \frac{x_2(t_K^{(*)})}{x_2(t_n^{(*)})}, \quad (27)$$

або (з врахуванням виразів (5), (6)) і (22)) –

$$\beta_{22}^{(*)} = 0,01033 + \ln \frac{x_2(t_K^{(*)})}{x_2(t_n^{(*)})}, \quad (28)$$

У таблиці 3 наведені результати розрахунків, виконаних за виразом (28).

Таблиця 3

$t_n$ (години)	24	72	144	156	166
$t_K$ (години)	25	73	145	157	167
$x_2(t_n)$ (%)	3,4	2,5	2,0	1,9	1,82
$x_2(t_K)$ (%)	6,6	7,5	8,0	8,1	8,18
$\beta_{22}$	0,6730	1,1089	1,3966	1,4596	1,5123

На рис. 3 представлені графіки процесу засвоєння студентом навчальної дисципліни у перший між-лекційний період в разі самостійного вивчення матеріалу, викладеного на першій лекції, через три дні після неї (суцільна крива 1) та за дві години до наступної лекції (штрихова крива 2).

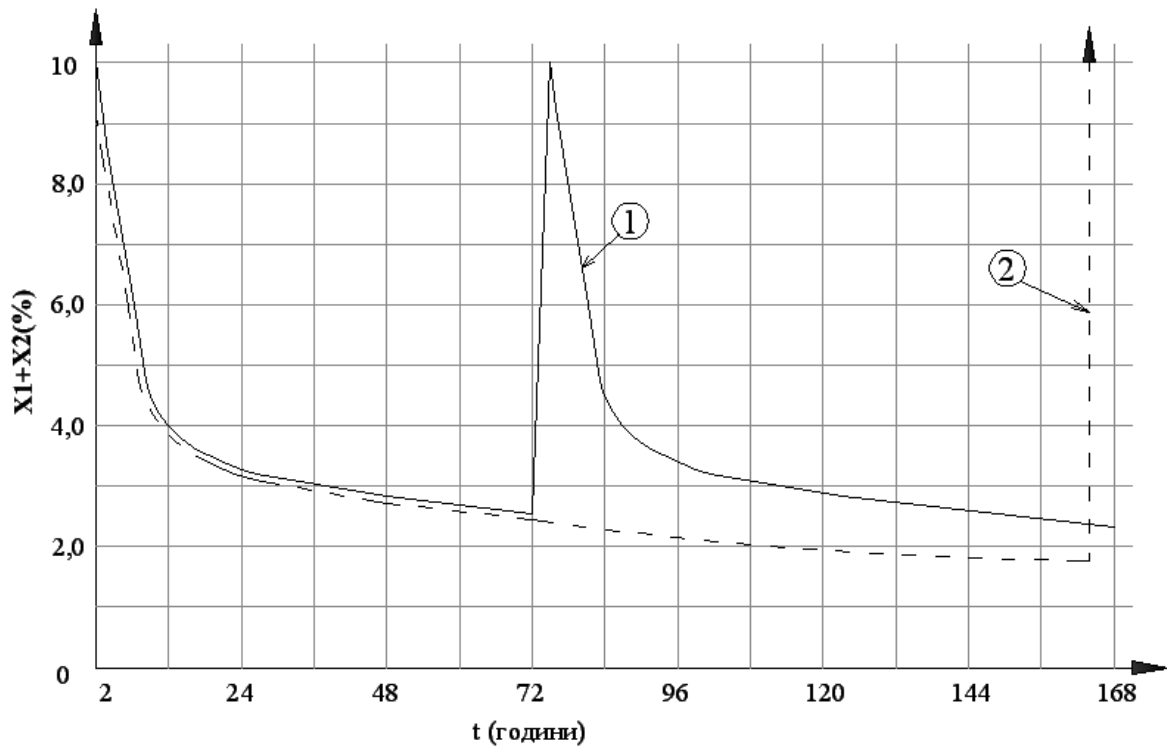


Рисунок 3 – Графіки процесу засвоєння студентом навчальної дисципліни у перший міжлекційний період в разі самостійного вивчення матеріалу, викладеного на першій лекції, через три дні після неї (сувільна крива 1) та за дві години до наступної лекції (штрихова крива 2)

Цілком очевидно, що графіки процесу засвоєння студентом навчальної дисципліни після наступних лекцій у наступних міжлекційних періодах будуть мати аналогічний вигляд. А тому графіки процесу засвоєння цієї навчальної дисципліни протягом триместру виглядатимуть так, як показано на рис.4, на якому сувільною лінією 1 зображено графік процесу засвоєння навчальної дисципліни у разі, якщо самостійно студент вивчатиме матеріал кожної лекції через три дні після її прослуховування, а штриховою лінією 2 зображено графік процесу засвоєння студентом навчальної дисципліни у разі, якщо самостійно студент вивчатиме матеріал кожної лекції за дві години до наступної лекції.

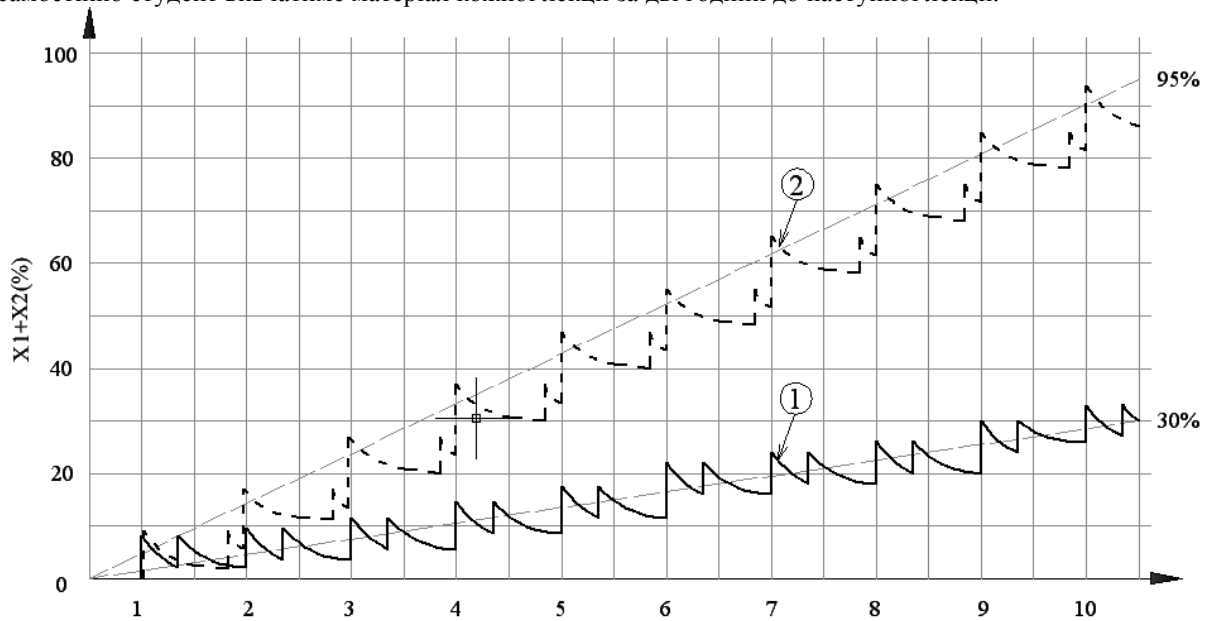


Рисунок 4 – Графіки процесу засвоєння студентом навчальної дисципліни для двох вищевказаних варіантів її вивчення ним самостійно у міжлекційні періоди

### Висновки

1. Встановлено, що параметр, який характеризує синергетичну взаємодію знань, отриманих студентом на лекції від викладача, та засвоєних ним самостійно, зі зростанням кількості синергуючих знань нелінійно зменшується, що повністю співпадає з відомим з мікроекономіки законом спадної граничної корисності у разі, якщо цю кількість знань розглядати як «благо», яке студент отримує в процесі вивчення навчальної дисципліни

2. Синтезована математична модель процесу самостійного засвоєння студентом знань у міжлекційний період та розроблено методику її ідентифікації.

3. Запропоновано спосіб побудови графіка процесу засвоєння студентом знань з навчальної дисципліни упродовж триместру.

4. Показано, що чим ближче до наступної лекції студент самостійно відновлює знання, отримані на попередній лекції, тим більшу кількість знань з даної навчальної дисципліни він матиме у кінці триместру.

### Список використаних джерел

1. Мокін Б. І. Дослідження впливу синергетичної складової у математичній моделі процесу засвоєння студентом навчальної дисципліни / Б. І. Мокін, А. В. Писклярова, О. Б. Мокін // Інформаційні технології та комп'ютерна інженерія. – 2013. – № 2. – С.9–14 .
2. М'ясоїд П. А. Загальна психологія. Навчальний посібник / П. А. М'ясоїд. – К.: Вища школа. – 1998. – 479 с.
3. Гиппенрейтер Ю. Б. Хрестоматия по общей психологии. Психология памяти / Ю. Б. Гиппенрейтер, В. Я. Романова. – М.: Издательство Московского университета. – 1979. – 272 с.
4. Мокін Б. І. Математичні моделі процесу засвоєння студентом навчальної дисципліни на фазовій площині / Б. І. Мокін, А. В. Писклярова, Ю. В. Фокіна // Вісник Вінницького політехнічного інституту. – 2010. – № 5. – С. 109–112.
5. Зянько В. В. Основи мікроекономіки. Навчальний посібник / В. В. Зянько – К.: Видавничий Дім «Сяйво». – 2009. – 344 с.

### Відомості про авторів:

**Мокін Борис Іванович** – д.т.н., професор, академік НАПНУ, професор кафедри відновлювальної енергетики та транспортних електричних систем і комплексів;

**Писклярова Анна Валеріївна** – к.т.н., проректор з науково-педагогічної роботи по організації виховного процесу;

**Мокін Олександр Борисович** – д.т.н., доцент, завідувач кафедри відновлювальної енергетики та транспортних електричних систем і комплексів.

Вінницький національний технічний університет

УДК 681.5.015:007

Г.Б. РАКИТЯНСЬКА

Вінницький національний технічний університет, м. Вінниця

## НЕЙРО-МЕРЕЖЕВИЙ ПІДХІД ДО ГЕНЕРУВАННЯ СПОЛУЧЕНИХ НЕЧІТКИХ БАЗ ЗНАНЬ НА ПРАВИЛАХ І ВІДНОШЕННЯХ

**Анотація.** Пропонується підхід до генерування сполучених правил ЯКЦО-ТО на основі генетико-нейронного алгоритму розв'язання рівнянь нечітких відношень, що дозволяє уникнути селекції правил і виключити перекриття між класами. Суть підходу полягає у побудові та навчанні *min-max* нейро-нечіткої мережі, ізоморфної лінгвістичним розв'язкам системи рівнянь нечітких відношень, яка дозволяє адаптувати структуру набору правил до змінення границь класів виходу. Розв'язання рівнянь нечітких відношень забезпечує оптимальну кількість нечітких правил для кожного вихідного терму і оптимальну геометрію вхідних термів для кожного лінгвістичного розв'язку.

**Ключові слова:** нечіткі правила і відношення, рівняння нечітких відношень, *min-max* нейронна мережа.

**Анотация.** Предлагается подход к генерированию составных правил ЕСЛИ-ТО на основе генетико-нейронного алгоритма решения уравнений нечетких отношений, что позволяет избежать селекции правил и исключить перекрытие между классами. Суть подхода состоит в построении и обучении *min-max* нейро-нечеткой сети, изоморфной лингвистическим решениям системы уравнений нечетких отношений, которая позволяет адаптировать структуру набора правил к изменению границ классов выхода. Решение уравнений нечетких отношений обеспечивает оптимальное число нечетких правил для каждого выходного термина и оптимальную геометрию входных термов для каждого лингвистического решения.

**Ключевые слова:** нечеткие правила и отношения, уравнения нечетких отношений, *min-max* нейронная сеть

**Abstract.** The adaptive approach to generating composite IF-THEN rules based on the genetic and neural algorithm of solving fuzzy relational equations is proposed. It allows us to avoid rules selection and eliminate overlaps between classes. The essence of the approach is in constructing and training the specific *min-max* neuro-fuzzy network isomorphic to linguistic solutions of fuzzy relational equations, which allows adaptation of the rules set structure while the output classes' bounds are changing. Resolution of fuzzy relational equations guarantees the optimal number of fuzzy rules for each output fuzzy term and the optimal geometry of input fuzzy terms for each linguistic solution.

**Key words:** fuzzy rules and relations, fuzzy relational equations, *min-max* neural network.

### Вступ

Нейро-мережовий підхід до здобування нечітких правил із експериментальних даних базується на генеруванні гіпербоксів [1]. Генерування правил у нейронних мережах на основі радіальних базисних функцій (RBF) [2], поєднане з навчанням машин опорних векторів (SVM), дозволяє визначити геометрію гіпербоксів [3]. Режим навчання у таких *min-max* нейронних мережах полягає у розширенні/стисненні гіпербоксів [1, 3]. Небажаним ефектом розширення гіпербоксів є їх перекриття, коли один образ повністю належить до двох або більше класів. Зменшення зон перекриття потребує більшої кількості гіпербоксів. Налаштування структури мережі полягає у відборі гіпербоксів шляхом їх об'єднання/розбиття, тобто селекції правил із множини правил-кандидатів.

### Актуальність

На сьогодні проблема селекції і оптимізації структури правил не має єдиного методичного стандарту вирішення. В цій статті пропонується підхід до генерування правил, виражений математично в термінах рівнянь нечітких відношень [4 – 6]. Система правил ЯКЦО-ТО може бути перетворена до множини лінгвістичних розв'язків рівнянь нечітких відношень шляхом переходу до сполучених нечітких термів [7, 8], де міра значимості терму I рівня (*підвищення, падіння*) описується термом II рівня (*значне підвищення, суттєве падіння*). Такий перехід дозволяє з'єднати терми I рівня нечіткими відношеннями, а терми II рівня – нечіткими правилами, які є якісними розв'язками рівнянь нечітких відношень для заданих класів виходу [7]. В цьому випадку задача здобування правил зводиться до розв'язання рівнянь нечітких відношень, що дозволяє уникнути селекції правил і виключити перекриття між класами. Розв'язання рівнянь нечітких відношень забезпечує оптимальну кількість нечітких правил для кожного вихідного терму і оптимальну геометрію вхідних термів для кожного лінгвістичного розв'язку.

### Мета

Мета роботи полягає у побудові та навчанні *min-max* нейронної мережі, ізоморфної сполученій базі знань, яка дозволяє адаптувати структуру набору правил до змінення границь класів виходу. Для випадку генерування правил узагальнений генетико-нейронний алгоритм розв'язання рівнянь нечітких відношень [9]. Початкова структура мережі встановлюється за допомогою генетичного алгоритму [10].

### Апроксимація нечіткими правилами і відношеннями

Розглядається об'єкт виду  $y = f(\mathbf{X})$  з  $n$  входами  $\mathbf{X} = (x_1, \dots, x_n)$  і одним виходом  $y$ , для якого є відомими:

- інтервал змінення входів і виходу  $x_i \in [\underline{x}_i, \bar{x}_i]$ ,  $i = \overline{1, n}$ ;  $y \in [\underline{y}, \bar{y}]$ ;

- система матриць нечітких відношень  $\mathbf{R}_i \subseteq c_{il} \times E_J = [r_{il}^J]$ ,  $i = \overline{1, n}$ ,  $l = \overline{1, k_i}$ ,  $J = \overline{1, M}$ , яка зв'язує терми I рівня  $c_{il}$  і  $E_J$  змінних  $x_i$  і  $y$ , відповідно;

- класи рішень  $d_j$  для правил,  $d_j \in [\underline{y}_{j-1}^r, \underline{y}_j^r]$ ,  $j = \overline{1, m}$ ;

- навчальна вибірка у вигляді  $L$  пар «входи – вихід»  $\langle \hat{\mathbf{X}}_s, \hat{y}_s \rangle$ ,  $s = \overline{1, L}$ , де  $\hat{\mathbf{X}}_s = (\hat{x}_1^s, \dots, \hat{x}_n^s)$  і  $\hat{y}_s$  – вектор значень вхідних і значення вихідної змінної в експерименті з номером  $s$ .

Необхідно синтезувати знання про об'єкт у вигляді сполученої нечіткої бази знань:

$$i. \quad \bigcup_{p=1, z_j} \left[ \bigcap_{i=1, n} \left\{ \bigcup_{l=1, k_i} (x_i = c_{il} \cap \mu^{c_{il}} = \alpha_{il}^{jp}) \right\} \right] \rightarrow y = d_j, \quad j = \overline{1, m}, \quad (1)$$

де  $\mu^{c_{il}}$  – міра значимості терму  $c_{il}$ ;  $\alpha_{il}^{jp}$  – терм II рівня, який описує міру значимості  $\mu^{c_{il}}$  в правилі з номером  $p = \overline{1, z_j}$ ;  $z_j$  – кількість правил у класі  $d_j$ ,  $j = \overline{1, m}$ .

При наявності матриць  $\mathbf{R}_i$ ,  $i = \overline{1, n}$ , залежність «входи – вихід» описується за допомогою розширеного композиційного правила виведення [4]

$$\mu^E(y) = \mu^{A_1}(x_1) \circ \mathbf{R}_1 \cap \dots \cap \mu^{A_n}(x_n) \circ \mathbf{R}_n, \quad (2)$$

де  $\mu^{A_i}(x_i) = (\mu^{c_{i1}}, \dots, \mu^{c_{ik_i}})$  – вектор мір значимостей нечітких термів  $c_{il}$ ,  $i = \overline{1, n}$ ,  $l = \overline{1, k_i}$ ;  $\mu^E(y) = (\mu^{E_1}, \dots, \mu^{E_M})$  – вектор мір значимостей нечітких термів  $E_J$ ,  $J = \overline{1, M}$ .

Із (2) випливає система рівнянь нечітких відношень, яка зв'язує функції належності нечітких термів I рівня вхідних і вихідної змінної:

$$\mu^{E_J}(y) = \min_{i=1, n} \left\{ \max_{l=1, k_i} \left[ \min(\mu^{c_{il}}(x_i), r_{il}^J) \right] \right\}, \quad J = \overline{1, M}. \quad (3)$$

Із сполученої бази знань (1) випливає система розв'язків рівнянь нечітких відношень (3) [10]:

$$\bigcup_{p=1, z_j} \left[ \bigcap_{i=1, n} \left\{ \bigcup_{l=1, k_i} (\mu^{c_{il}}(x_i) = \alpha_{il}^{jp}) \right\} \right] \rightarrow \mu^{D_j}(y) = \delta_j, \quad j = \overline{1, m}, \quad (4)$$

де  $D_j$  – терм I рівня, який описує змінну  $y$  в класі  $d_j$ ,  $D_j \in \{E_1, \dots, E_M\}$ ;  $\mu^{D_j}$  – міра значимості терму  $D_j$ ;  $\delta_j$  – терм II рівня, який описує міру значимості  $\mu^{D_j}$ .

Перехід від термів  $\alpha_{il}^{jp}$ , що описують міри значимостей  $\mu^{c_{il}}$ , до термів  $a_{il}^{jp}$ , що описують змінні  $x_i$ , дозволив отримати лінгвістичні розв'язки в системі нечітких відношень:

$$\bigcup_{p=1, z_j} \left[ \bigcap_{i=1, n} \left\{ \bigcup_{l=1, k_i} (x_i = a_{il}^{jp}) \right\} \right] \rightarrow y = d_j, \quad j = \overline{1, m}, \quad (5)$$

де  $a_{il}^{jp} = (c_{il}, \alpha_{il}^{jp})$  – сполучений терм, що описує змінну  $x_i$ ,  $i = \overline{1, n}$ ,  $l = \overline{1, k_i}$ , в правилі з номером  $jp$ .

ii. Нечіткій базі знань (5) відповідають нечіткі логічні рівняння, які зв'язують функції належності сполучених термів у розв'язках системи (3):

$$\mu^{d_j}(y) = \max_{p=1, z_j} \left[ w_{jp} \cdot \min_{i=1, n} \left\{ \max_{l=1, k_i} (v_{il}^{jp} \cdot \mu^{a_{il}^{jp}}(x_i)) \right\} \right], \quad j = \overline{1, m}, \quad (6)$$

де  $\mu^{d_j}(y)$  і  $\mu^{a_{il}^{jp}}(x_i)$  – функції належності змінної  $y$  і  $x_i$  до сполучених термів  $d_j = (D_j, \delta_j)$  і  $a_{il}^{jp} = (c_{il}, \alpha_{il}^{jp})$ ;  $w_{jp}$  – вага правила з номером  $jp$ ;  $v_{il}^{jp}$  – вага терму у розв'язку з номером  $jp$ .

У нечітких логічних рівняннях використовується така функція належності нечіткого терму  $T$ :

$$\mu^T(u) = 1 / (1 + ((u - \beta) / \sigma)^2), \quad (7)$$

де  $\beta$  – координата максимуму функції,  $\mu^T(\beta) = 1$ ;  $\sigma$  – параметр концентрації [6].

Для правил операція дефазифікації виконується за формулою [6]:

$$y = \frac{\sum_{j=1}^m y_j^r \cdot \mu^{d_j}(y)}{\sum_{j=1}^m \mu^{d_j}(y)}. \quad (8)$$

Якщо правила (5) є розв'язками системи рівнянь нечітких відношень (3), то для якісних значень входів  $x_i = a_{il}^{jp}$  і виходу  $y = d_j$  у розв'язку з номером  $jp$  виконується співвідношення:

$$\mu^{E_j}(d_j) = \min_{i=1, n} \left\{ \max_{l=1, k_i} \left[ \min(\mu^{c_{il}}(a_{il}^{jp}), r_{il}^j) \right] \right\},$$

де  $\mu^{E_j}(d_j)$  і  $\mu^{c_{il}}(a_{il}^{jp})$  – степені належності значень  $x_i = a_{il}^{jp}$  і  $y = d_j$  до нечітких термів  $E_j$  і  $c_{il}$ .

Тоді виникає обернена задача, яка ставиться таким чином: для класів виходу  $y = d_j$ ,  $j = \overline{1, m}$ , знайти кількість правил  $z_j$  і відновити форми функцій належності входів  $x_i = a_{il}^{jp}$  у кожному правилі.

#### Задача здобування нечітких правил

Співвідношення (3) – (8) визначають загальний вид нечіткої моделі об'єкта в системі правил і відношень таким чином:

$$\mu^E(y) = f_R(\mathbf{X}, \mathbf{R}, \mathbf{B}_C, \mathbf{\Omega}_C, \mathbf{B}_E, \mathbf{\Omega}_E), \quad (9)$$

$$y = f_r(\mathbf{X}, f_R, Z, q, \mathbf{V}, \mathbf{W}, \mathbf{B}_a, \mathbf{\Omega}_a, \mathbf{B}_d, \mathbf{\Omega}_d), \quad (10)$$

де  $\mathbf{B}_C = (\beta^{c_{i1}}, \dots, \beta^{c_{i k_i}})$ ,  $\mathbf{\Omega}_C = (\sigma^{c_{i1}}, \dots, \sigma^{c_{i k_i}})$  – вектори параметрів функцій належності нечітких термів  $c_{il}$ ,  $i = \overline{1, n}$ ,  $l = \overline{1, k_i}$ ;  $\mathbf{B}_E = (\beta^{E_1}, \dots, \beta^{E_M})$ ,  $\mathbf{\Omega}_E = (\sigma^{E_1}, \dots, \sigma^{E_M})$  – вектори параметрів функцій належності нечітких термів  $E_j$ ,  $J = \overline{1, M}$ ;  $\mathbf{W} = (w_1, \dots, w_Z)$  – вектор ваг правил;  $Z$  – число правил;  $\mathbf{V} = (v_1, \dots, v_q)$  – вектор ваг термів;  $q$  – число термів;  $\mathbf{B}_a = (\beta^{a_1}, \dots, \beta^{a_q})$ ,  $\mathbf{\Omega}_a = (\sigma^{a_1}, \dots, \sigma^{a_q})$  – вектори параметрів функцій належності нечітких термів  $a_k$ ,  $k = \overline{1, q}$ ;  $\mathbf{B}_d = (\beta^{d_1}, \dots, \beta^{d_m})$ ,  $\mathbf{\Omega}_d = (\sigma^{d_1}, \dots, \sigma^{d_m})$  – вектори параметрів функцій належності нечітких термів  $d_j$ ,  $j = \overline{1, m}$ ;  $F_R$  і  $f_r$  – оператори зв'язку «входи-вихід», що відповідають формулам (3), (7) і (6) – (8), відповідно.

Для заданої системи нечітких відношень, задача здобування нечітких правил формулюється так. Необхідно знайти таку кількість правил  $Z$  і термів  $q$ , а також такі вектори ваг правил  $\mathbf{W}$ , ваг термів  $\mathbf{V}$  та вектори параметрів функцій належності входів і виходу  $\mathbf{B}_a, \mathbf{\Omega}_a, \mathbf{B}_d, \mathbf{\Omega}_d$ , які забезпечують мінімальну відстань між модельним і експериментальним виходами об'єкта:

$$\sum_{s=1}^L [f_r(\hat{\mathbf{X}}_s, f_R, Z, q, \mathbf{W}, \mathbf{V}, \mathbf{B}_a, \mathbf{\Omega}_a, \mathbf{B}_d, \mathbf{\Omega}_d) - \hat{y}_s]^2 = \min_{Z, q, \mathbf{W}, \mathbf{V}, \mathbf{B}_a, \mathbf{\Omega}_a, \mathbf{B}_d, \mathbf{\Omega}_d}. \quad (11)$$

**Здобування лінгвістичних розв'язків системи рівнянь нечітких відношень**

Елементами розв'язків (4) рівнянь нечітких відношень (3) є значення вхідних змінних  $x_i, i = \overline{1, n}$ , для яких  $\mu^{C_{il}}(x_i) = \alpha_{il}^{jp}$ ,  $p = \overline{1, z_j}$ . Ці значення інтерпретуються як координати максимуму функцій належності нечітких термів  $a_{il}^{jp}$ , що описують змінну  $x_i$  в правилі  $jp, p = \overline{1, z_j}$ , бази знань (5), де значенню виходу  $y = d_j, j = \overline{1, m}$ , відповідає  $z_j$  лінгвістичних розв'язків системи (3).

Нехай  $\mathbf{B}_j = (\beta_1^j, \dots, \beta_N^j) = (\beta_{11}^j, \dots, \beta_{1k_1}^j, \dots, \beta_{n1}^j, \dots, \beta_{nk_n}^j)$  – вектор координат максимуму функцій належності нечітких термів у правилі в класі  $y = d_j, j = \overline{1, m}$ . Слідуючи [6, 9, 10], задача розв'язання рівнянь нечітких відношень (3) формулюється так. Для кожного класу виходу  $y = d_j, j = \overline{1, m}$ , знайти вектор координат максимуму  $\mathbf{B}_j = (\beta_1^j, \dots, \beta_N^j), \beta_{il}^j \in [\underline{x}_i, \overline{x}_i], i = \overline{1, n}$ , який забезпечує мінімальну відстань між лівою і правою частиною кожного рівняння системи (3):

$$\sum_{j=1}^m \sum_{J=1}^M \left[ \mu^{E_j}(d_j) - \min_{i=1, n} \left[ \max_{l=1, k_i} \left( \min(\mu^{C_{il}}(\beta_{il}^j), r_{il}^j) \right) \right] \right]^2 = \min_{\mu^C(\mathbf{B}_j)} . \quad (12)$$

Для кожного класу  $d_j$  система рівнянь (3) має множину розв'язків  $S_j(\mathbf{R}, \mu^E(d_j))$ , яка визначається множиною максимальних розв'язків  $\overline{S}_j^* = \{\underline{\mu}^C(\overline{\mathbf{B}}_{jh}), h = \overline{1, z_j}\}$  і множиною мінімальних розв'язків  $\underline{S}_j^* = \{\underline{\mu}^C(\underline{\mathbf{B}}_{js}), s = \overline{1, z_j}\}$ . При цьому кожному максимальному розв'язку  $\underline{\mu}^C(\overline{\mathbf{B}}_{jh}) \in \overline{S}_j^*$ , який визначає верхній опорний вектор  $\overline{\mathbf{B}}_{jh}$ , відповідає множина мінімальних розв'язків  $\underline{S}_j^* = \{\underline{\mu}^C(\underline{\mathbf{B}}_{js}), s = \overline{1, z_j}\}$ , яка визначає множину нижніх опорних векторів  $\underline{\mathbf{B}}_{js}$  [10]:

$$S_j(\mathbf{R}, \mu^E(d_j)) = \bigcup_{\underline{\mu}^C(\overline{\mathbf{B}}_{jh}) \in \overline{S}_j^*} \bigcup_{\underline{\mu}^C(\underline{\mathbf{B}}_{js}) \in \underline{S}_j^*} \left[ \underline{\mu}^C(\underline{\mathbf{B}}_{js}), \underline{\mu}^C(\overline{\mathbf{B}}_{jh}) \right], j = \overline{1, m}. \quad (13)$$

Тут  $\overline{\mathbf{B}}_{jh} = (\overline{\beta}_1^{jh}, \dots, \overline{\beta}_N^{jh})$  і  $\underline{\mathbf{B}}_{js} = (\underline{\beta}_1^{js}, \dots, \underline{\beta}_N^{js})$  – вектори верхніх і нижніх границь координат максимуму  $\beta_I^{jp}$ , де операція об'єднання виконується над усіма  $\underline{\mu}^C(\overline{\mathbf{B}}_{jh}) \in \overline{S}_j^*$  і  $\underline{\mu}^C(\underline{\mathbf{B}}_{js}) \in \underline{S}_j^*$ .

Слідуючи [10], формування інтервалів (13) здійснюється шляхом багаторазового розв'язання задачі оптимізації (12) і починається з пошуку її нульових розв'язків  $\mathbf{B}_{j0} = (\beta_1^{j0}, \dots, \beta_N^{j0}), j = \overline{1, m}$ .

Верхня границя  $(\overline{\beta}_I^{jh})$  для  $h=1$  знаходиться в діапазоні  $[\beta_I^{j0}, 1]$ , а для  $h>1$  – в діапазоні  $[\max(\underline{\beta}_I^{jp}), 1], p < s$ , причому максимальні розв'язки  $\overline{\beta}_I^{jp}, p < h$ , вилучаються із області пошуку. Нижня границя  $(\underline{\beta}_I^{js})$  для  $s=1$  знаходиться в діапазоні  $[0, \beta_I^{j0}]$ , а для  $s>1$  – в діапазоні  $[0, \min(\overline{\beta}_I^{jp})], p < h$ , причому мінімальні розв'язки  $\underline{\beta}_I^{jp}, p < s$ , вилучаються із області пошуку.

Нехай  $\mathbf{B}_j(t) = (\beta_1^j(t), \dots, \beta_N^j(t))$  – розв'язок задачі оптимізації (12) на  $t$ -ому кроці формування інтервалів, тобто  $F(\mathbf{B}_j(t)) = F(\mathbf{B}_{j0})$ , оскільки для всіх  $\mu^C(\mathbf{B}_j) \in S(\mathbf{R}, \mu^E(d_j))$  значення критерію (12) однакове. При пошуку верхніх границь передбачається, що  $\beta_I^j(t) \geq \beta_I^j(t-1)$ , а при пошуку нижніх границь передбачається, що  $\beta_I^j(t) \leq \beta_I^j(t-1)$ . Встановлення верхніх (нижніх) границь здійснюється за правилом: якщо  $\mathbf{B}_j(t) \neq \mathbf{B}_j(t-1)$ , то  $\overline{\beta}_I^{jh}(\underline{\beta}_I^{js}) = \beta_I^j(t)$ . Якщо  $\mathbf{B}_j(t) = \mathbf{B}_j(t-1)$ ,

то формування розв'язку  $[\underline{\mathbf{V}}_{js}, \overline{\mathbf{V}}_{jh}]$  припиняється. Пошук інтервалів (13) продовжується, поки виконується умова  $\overline{\mathbf{V}}_{jh} \neq \overline{\mathbf{V}}_{jp}$ ,  $p < h$ , для верхніх границь і  $\underline{\mathbf{V}}_{js} \neq \underline{\mathbf{V}}_{jp}$ ,  $p < s$ , для нижніх границь.

**Адаптивне здобування нечітких правил**

В цьому розділі пропонується спосіб представлення лінгвістичної інформації про об'єкт у вигляді спеціальної нейро-нечіткої мережі, ізоморфної сполученій базі знань (1). Структура такої мережі представлена на рис.1. Нейро-нечітка модель отримана шляхом імплантації розв'язків системи рівнянь нечітких відношень в нейронну мережу таким чином, що вагами дуг, які підлягають навчанню, є параметри функцій належності і ваги правил. Мережа на рис. 1 навчається в двох режимах. Перший режим забезпечує налаштування структури правил шляхом розв'язання системи рівнянь нечітких відношень (3). Другий режим забезпечує налаштування параметрів правил - розв'язків.

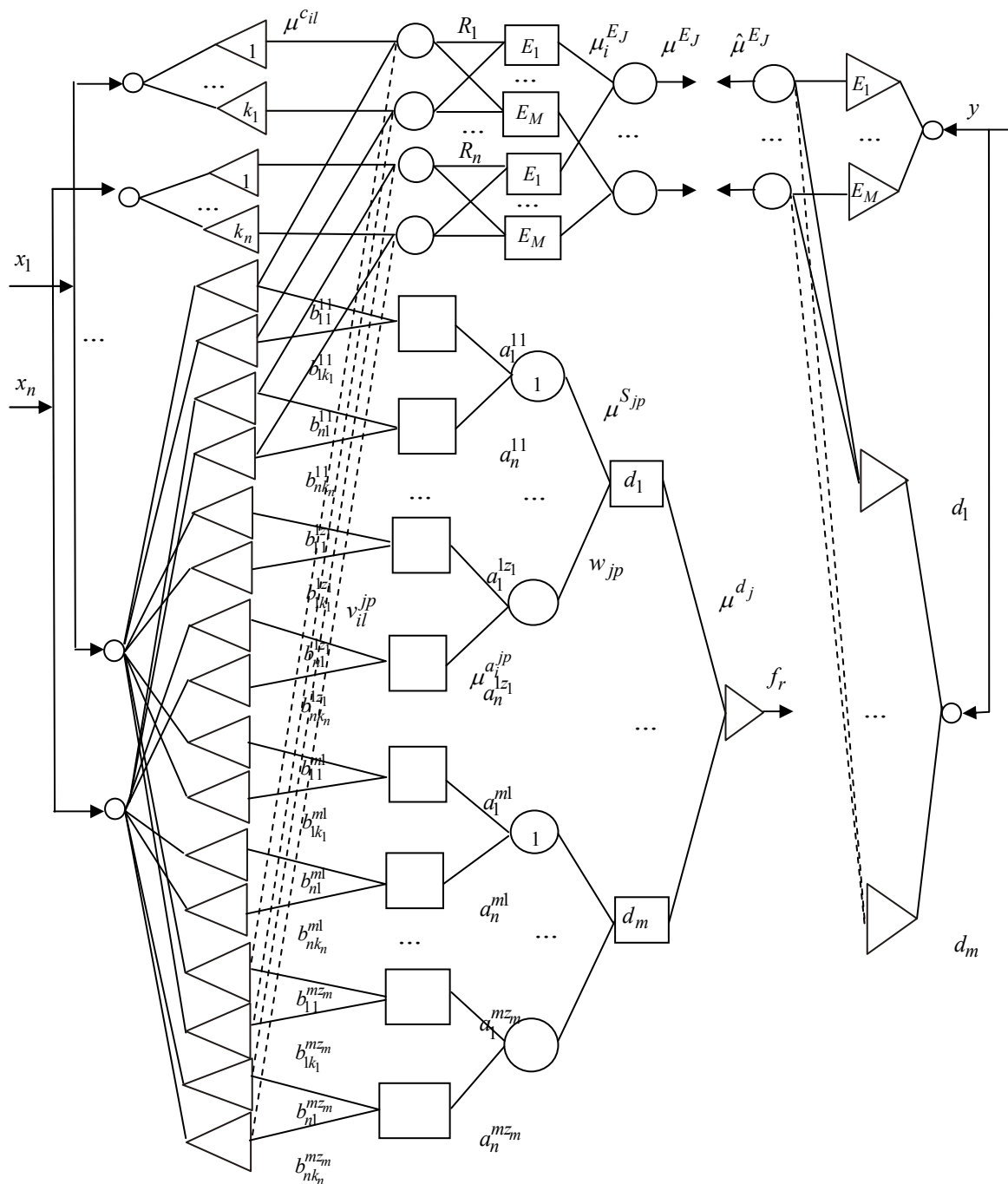


Рисунок 1 – Нейро-нечітка модель об'єкта в системі відношень і правил

Нейро-нечітка мережа для настройки структури правил має п'ять шарів: входи об'єкта (шар 1); операція фазифікації для термів I рівня у відношеннях і термів II рівня у правилах (шар 2); операція *min* для підстановки сполучених термів у рівняння нечітких відношень (шар 3); операція *max* над функціями належності виходу для відношень  $R_i$  (шар 4); операція *min* над функціями належності виходу (шар 5).

Кількість вузлів в кожному шарі нейро-нечіткої мережі визначається так: по кількості входів об'єкта  $x_i$ ,  $i = \overline{1, n}$  (шар 1); по кількості термів I рівня,  $k_1 + \dots + k_n$ , і кількості термів II рівня,  $N \cdot (z_1 + \dots + z_m)$  (шар 2); по кількості термів I рівня,  $k_1 + \dots + k_n$  (шар 3); по кількості класів виходу для відношень  $R_i$ ,  $n \cdot M$  (шар 4); по кількості класів виходу  $E_J$ ,  $J = \overline{1, M}$  (шар 5).

Для настройки структури нечітких правил використовуються рекурентні співвідношення:

$$\beta_{il}^{jp}(t+1) = \beta_{il}^{jp}(t) - \eta \frac{\partial \varepsilon_t^s}{\partial \beta_{il}^{jp}(t)}, \quad (14)$$

які мінімізують критерій

$$\varepsilon_t^s = (\hat{\mu}^E(t) - \mu^E(t))^2 / 2,$$

де  $\hat{\mu}^E(t)$  ( $\mu^E(t)$ ) – експериментальний (теоретичний) вектор мір значимостей вихідних термів системи рівнянь (3) на  $t$ -му кроці навчання;  $\beta_{il}^{jp}(t)$  – координати максимуму функцій належності нечітких термів вхідних змінних на  $t$ -му кроці навчання.

Частинні похідні, що входять у співвідношення (14), характеризують чутливість похибки  $\varepsilon_t^s$  до змінення параметрів нейро-нечіткої мережі і обчислюються таким чином:

$$\frac{\partial \varepsilon_t^s}{\partial \beta_{il}^{jp}} = \sum_{J=1}^M \left[ \frac{\partial \varepsilon_t^s}{\partial \mu^{E_J}} \cdot \frac{\partial \mu^{E_J}}{\partial \mu_i^{E_J}} \cdot \frac{\partial \mu_i^{E_J}}{\partial \mu^{c_{il}}} \cdot \frac{\partial \mu^{c_{il}}}{\partial \beta_{il}^{jp}} \right].$$

Оскільки при визначенні елемента «нечіткий вихід» присутні нечітко-логічні операції *min* і *max*, то співвідношення для навчання отримані за допомогою кінцево-різницевого схем [9].

Нейро-нечітка мережа для настройки параметрів правил має шість шарів: входи об'єкта (шар 1); фазифікація для нечітких термів II рівня у лінгвістичних розв'язках (шар 2); операція *max* для злиття термів у лінгвістичних розв'язках (шар 3); операція *min* для лінгвістичних правил-розв'язків (шар 4); операція *max* для класів розв'язків (шар 5); операція дефазифікації (шар 6).

Кількість вузлів в кожному шарі нейро-нечіткої мережі визначається так: по кількості входів об'єкта  $x_i$ ,  $i = \overline{1, n}$  (шар 1); по кількості термів у лінгвістичних розв'язках,  $N \cdot (z_1 + \dots + z_m)$  (шар 2); по кількості об'єднаних термів у базі знань,  $n \cdot (z_1 + \dots + z_m)$  (шар 3); по кількості розв'язків  $S_{jp}$  в кожному класі виходу,  $z_1 + \dots + z_m$  (шар 4); по кількості класів виходу  $d_j$ ,  $j = \overline{1, m}$  (шар 5).

Для настройки параметрів нечітких правил використовуються рекурентні співвідношення:

$$w_{jp}(t+1) = w_{jp}(t) - \eta \frac{\partial \varepsilon_t^r}{\partial w_{jp}(t)}, \quad v_{il}^{jp}(t+1) = v_{il}^{jp}(t) - \eta \frac{\partial \varepsilon_t^r}{\partial v_{il}^{jp}(t)}, \quad \sigma_{il}^{jp}(t+1) = \sigma_{il}^{jp}(t) - \eta \frac{\partial \varepsilon_t^r}{\partial \sigma_{il}^{jp}(t)}, \quad (15)$$

які мінімізують критерій

$$\varepsilon_t^r = (\hat{y}_t^r - y_t^r)^2 / 2,$$

де  $\hat{y}_t^r$  ( $y_t^r$ ) – експериментальний (теоретичний) вихід об'єкта для нечітких правил на  $t$ -му кроці навчання;  $w_{jp}(t)$ ,  $v_{il}^{jp}(t)$  – ваги нечітких правил і термів на  $t$ -му кроці навчання;  $\sigma_{il}^{jp}(t)$  – параметри концентрації функцій належності нечітких термів вхідних змінних на  $t$ -му кроці навчання.

Частинні похідні, що входять в співвідношення (15), характеризують чутливість похибки  $\varepsilon_t^r$  до змінення параметрів нейро-нечіткої мережі і обчислюються таким чином:

$$\frac{\partial \varepsilon_t^r}{\partial w_{jp}} = \varepsilon_1 \varepsilon_2 \varepsilon_3 \cdot \frac{\partial \mu^{d_j}(\mathbf{X})}{\partial w_{jp}}; \quad \frac{\partial \varepsilon_t^r}{\partial v_{il}^{jp}} = \varepsilon_1 \varepsilon_2 \varepsilon_3 \varepsilon_4 \cdot \frac{\partial \mu^{a_i^{jp}}(x_i)}{\partial v_{il}^{jp}}; \quad \frac{\partial \varepsilon_t^r}{\partial \sigma^{b_{il}^{jp}}} = \varepsilon_1 \varepsilon_2 \varepsilon_3 \varepsilon_4 \varepsilon_5 \cdot \frac{\partial \mu^{b_{il}^{jp}}(x_i)}{\partial \sigma^{b_{il}^{jp}}};$$

$$\text{де } \varepsilon_1 = \frac{\partial \varepsilon_t^r}{\partial y}; \quad \varepsilon_2 = \frac{\partial y}{\partial \mu^{d_j}(\mathbf{X})}; \quad \varepsilon_3 = \frac{\partial \mu^{d_j}(\mathbf{X})}{\partial \mu^{S_{jp}}(\mathbf{X})}; \quad \varepsilon_4 = \frac{\partial \mu^{S_{jp}}(\mathbf{X})}{\partial \mu^{a_i^{jp}}(x_i)}; \quad \varepsilon_5 = \frac{\partial \mu^{a_i^{jp}}(x_i)}{\partial \mu^{b_{il}^{jp}}(x_i)}.$$

### Комп'ютерний експеримент

Експериментальні дані про об'єкт генерувались моделлю «два входи – один вихід» (рис. 2):

$$y = ((2z - 0.9)(7z - 1)(17z - 19)(15z - 2))/10, \quad z = ((x_1 - 3.0)^2 + (x_2 - 2.5)^2)/40.$$

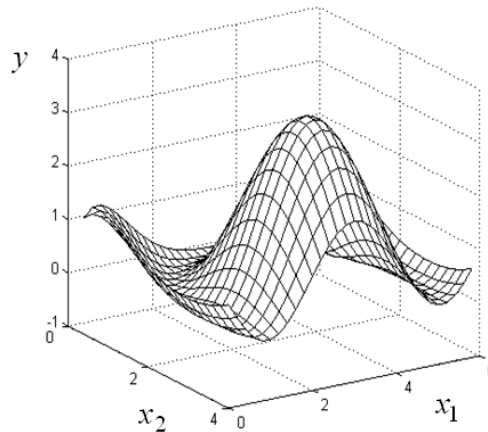


Рисунок 2 – Модель - генератор

Нечіткі відношення, здобуті із даних [6], представлені в табл. 1 разом з параметрами функцій належності нечітких термів І рівня:  $c_{11}$  ( $c_{21}$ ) зниження до 0,  $c_{12}$  ( $c_{22}$ ) наближення до 3.0,  $c_{13}$  підвищення до 6.0 для  $x_1$  і  $x_2$ ;  $E_1$  зниження до 0,  $E_2$  наближення до 1.0,  $E_3$  підвищення до 3.5 для  $y$ .

Необхідно синтезувати правила, які описують об'єкт для  $m = 5$ . Границі класів  $d_1 \div d_5$  були встановлені таким чином:  $d_1 \in [0, 0.4)$ ;  $d_2 \in [0.4, 0.9)$ ;  $d_3 \in [0.9, 1.3)$ ;  $d_4 \in [1.3, 2.5)$ ;  $d_5 \in [2.5, 3.4]$ .

Система рівнянь нечітких відношень для генерування правил - розв'язків має вигляд:

$$\begin{aligned} \mu^{D_1} &= [(\mu^{c_{11}} \wedge 0.98) \vee (\mu^{c_{12}} \wedge 0.22) \vee (\mu^{c_{13}} \wedge 0.98)] \wedge [(\mu^{c_{21}} \wedge 0.29) \vee (\mu^{c_{22}} \wedge 0.96)] \\ \mu^{D_2} &= [(\mu^{c_{11}} \wedge 0.92) \vee (\mu^{c_{12}} \wedge 0.44) \vee (\mu^{c_{13}} \wedge 0.92)] \wedge [(\mu^{c_{21}} \wedge 0.89) \vee (\mu^{c_{22}} \wedge 0.30)] \\ \mu^{D_3} &= [(\mu^{c_{11}} \wedge 0.11) \vee (\mu^{c_{12}} \wedge 0.91) \vee (\mu^{c_{13}} \wedge 0.11)] \wedge [(\mu^{c_{21}} \wedge 0.12) \vee (\mu^{c_{22}} \wedge 0.94)]. \end{aligned} \quad (16)$$

Для кожного класу міри значимостей  $\mu^D(d_j)$  визначались за допомогою функцій належності нечітких термів  $E_1 \div E_3$  і  $d_1 \div d_5$ , параметри яких представлені в Табл. 1 і 2:

$$\begin{aligned} \mu^D(d_1) &= (0.67, 0.30, 0.09); \quad \mu^D(d_2) = (0.40, 0.38, 0.10); \quad \mu^D(d_3) = (0.45, 0.80, 0.39); \\ \mu^D(d_4) &= (0.11, 0.30, 0.67); \quad \mu^D(d_5) = (0.14, 0.30, 0.90). \end{aligned}$$

За допомогою генетичного алгоритму були отримані множини розв’язків для  $\beta$  – параметрів правил, які представлені в табл. 2. Набір правил в табл. 2 відповідає множині розв’язків рівнянь нечітких відношень (16), де терми II рівня асоціюються з інтервалами мір значимостей  $\mu^{c_{ij}}(\beta_{ij}^j)$ , для яких значення критерію оптимізації (12) становить  $F = 0.0193$ . Інтервали значень  $\beta$  – параметрів були визначені за допомогою функцій належності нечітких термів  $c_{11} \div c_{13}, c_{21}, c_{22}$ .

Таблиця 1 – Матриця нечітких відношень

	ЯКЦО входи	ГО вихід $y$		
		$E_1, (0.02, 0.27)$	$E_2, (1.10, 0.29)$	$E_3, (3.40, 0.71)$
$x_1$	$c_{11}, (0, 0.71)$	0.98	0.92	0.11
	$c_{12}, (3.0, 0.92)$	0.22	0.44	0.89
	$c_{13}, (6.0, 0.70)$	0.98	0.92	0.11
$x_2$	$c_{21}, (0, 0.59)$	0.29	0.89	0.12
	$c_{22}, (3.0, 0.81)$	0.96	0.38	0.94

Таблиця 2 – Множина значень  $\beta$  – параметрів правил після генетичної настройки

ЯКЦО										ГО
$x_1$					$x_2$					$y$
$\mu^{c_{11}}$	$\beta_{11}^j$	$\mu^{c_{12}}$	$\beta_{12}^j$	$\mu^{c_{13}}$	$\beta_{13}^j$	$\mu^{c_{21}}$	$\beta_{21}^j$	$\mu^{c_{22}}$	$\beta_{22}^j$	$(\beta, \sigma)$
0.67 [0, 0.67] [0.67, 1] [0, 1]	[0, 0.50] [0.50, 6.0]	[0, 0.11] [0, 0.11] [0, 0.11] [0, 0.11]	[0, 0.40] [5.60, 6.0]	[0, 0.67] 0.67 [0, 1] [0.67, 1]	[5.50, 6.0] [0, 5.50]	[0, 0.30] [0, 0.30] [0, 0.30] [0, 0.30]	0.90	[0.67, 1] [0.67, 1] 0.67 0.67	[2.43, 3.57]	$d_1$ , (0.35, 0.21)
0.40 [0, 0.40] [0.40, 1] [0, 1]	[0, 0.86] [0.86, 6.0]	0.12 0.12 0.12 0.12	0.50 5.50	[0, 0.40] 0.40 [0, 1] [0.40, 1]	[5.15, 6.0] [0, 5.15]	0.38 0.38 0.38 0.38	0.75	[0.40, 1] [0.40, 1] 0.40 0.40	2.00 4.00	$d_2$ , (0.52, 0.15)
0.80 [0, 0.80] [0.80, 1] [0, 1.0]	[0, 0.35]	0.39 0.39 0.39 0.39	1.85 4.15	[0, 0.80] 0.80 [0, 1] [0.80, 1]	[5.65, 6.0]	[0.80, 1] [0.80, 1] 0.80 0.80	[0, 0.30]	0.45 0.45 0.45 0.45	2.10 3.90	$d_3$ , (1.27, 0.90)
[0, 0.22] [0, 0.22]		0.67 [0.67, 1]	[2.36, 3.64]	[0, 0.22] [0, 0.22]		[0, 0.30] [0, 0.30]		[0.67, 1] 0.67	[2.43, 3.57]	$d_4$ , (2.46, 0.64)
[0, 0.22] [0, 0.22]		0.90 [0.90, 1]	[2.70, 3.30]	[0, 0.22] [0, 0.22]		[0, 0.30] [0, 0.30]		[0.90, 1] 0.90	[2.73, 3.27]	$d_5$ , (2.82, 0.87)

Нехай кількість класів для правил збільшена до  $m = 6$  за рахунок нових класів  $d_{21}$  і  $d_{22}$ , на які був поділений клас  $d_2$ . Уточнені границі класів становлять:  $d_1 \in [0, 0.3]$ ;  $d_{21} \in [0.3, 0.6]$ ;  $d_{22} \in [0.6, 0.9]$ . Для решти класів границі не змінились.

Для нових класів міри значимостей  $\mu^D(d_j)$  визначались за допомогою функцій належності нечітких термів  $E_1 \div E_3$  і  $d_1, d_{21}, d_{22}$ , параметри яких представлені в табл. 1 і 3:

$$\mu^D(d_1) = (0.76, 0.28, 0.15); \mu^D(d_{21}) = (0.51, 0.45, 0.11); \mu^D(d_{22}) = (0.20, 0.38, 0.10).$$

Шляхом нейронної підстройки границь розв’язків в табл. 2\_були отримані множини розв’язків для  $\beta$  – параметрів правил, представлені в табл. 3. Уточнений набір правил для нових класів  $d_1, d_{21}$  і  $d_{22}$  в табл. 3 відповідає множині розв’язків рівнянь нечітких відношень (16), для яких значення критерію оптимізації (12) становить  $F = 0.0271$ .

Лінгвістична інтерпретація інтервалів  $\beta$  – параметрів представлена в табл. 4. Результати генетико-нейронної настройки параметрів отриманого набору правил представлені в табл. 5 і табл. 6.

Здобуті відношення і правила забезпечують апроксимацію об'єкта, яка показана на рис. 3. Точність виведення для відношень становить на рівні  $RMSE = 0.6012$ . Точність виведення для правил після генетичної настройки становить на рівні  $RMSE = 0.3805$  при  $Z = 16$ , а після нейронної настройки становить на рівні  $RMSE = 0.2402$  при  $Z = 21$ .

Таблиця 3 – Множина значень  $\beta$  – параметрів правил після нейронної настройки

ЯКЦО										ТО
$x_1$					$x_2$					$y$
$\mu^{c_{11}}$	$\beta_{11}^j$	$\mu^{c_{12}}$	$\beta_{12}^j$	$\mu^{c_{13}}$	$\beta_{13}^j$	$\mu^{c_{21}}$	$\beta_{21}^j$	$\mu^{c_{22}}$	$\beta_{22}^j$	$(\beta, \sigma)$
0.76 [0, 0.76] [0.76, 1] [0, 1]	[0, 0.40] [0.40, 6.0]	0.15 0.15 0.15 0.15	0.85 5.15	[0,0.76] 0.76 [0, 1] [0.76, 1]	[5.61, 6.0] [0, 5.61]	[0, 0.30] [0, 0.30] [0, 0.30] [0, 0.30]	0.90	[0.76, 1] [0.76, 1] 0.76 0.76	[2.55,3.45]	$d_1$ , (0.29, 0.21)
0.51 [0, 0.51] [0.51, 1] [0, 1]	[0, 0.72] [0.72, 6.0]	[0, 0.11] [0, 0.11] [0, 0.11] [0, 0.11]	[0, 0.40] [5.60,6.0]	[0,0.51] 0.51 [0, 1] [0.51, 1]	[5.31, 6.0] [0, 5.31]	0.45 0.45 0.45 0.45	0.65	[0.51, 1] [0.51, 1] 0.51 0.51	[2.20,3.80]	$d_{21}$ , (0.53, 0.25)
0.38 [0, 0.38] [0.38, 1] [0, 1]	[0, 0.90]	[0, 0.11] [0, 0.11] [0, 0.11] [0, 0.11]		[0, 0.38] 0.38 [0, 1] [0.38, 1]	[5.10, 6.0]	[0.38, 1] [0.38, 1] 0.38 0.38	[0, 0.76]	[0, 0.29] [0, 0.29] [0, 0.29] [0, 0.29]		$d_{22}$ , (0.68, 0.05)

Таблиця 4 – Лінгвістична інтерпретація розв'язків

ЯКЦО входи				ТО вихід	
Генетичний алгоритм (ГА)		Нейронна мережа (НМ)		ГА	НМ
$x_1$	$x_2$	$x_1$	$x_2$	$y$	
$нС - вС$ $Н$ або $В$	$нС$ $вС$	$вН - нВ$ $вН$ або $нВ$	$нС$ $С - В$	$d_1$	$d_1$
$нС - вС$ $Н$ або $В$	$Н$ $нС$ або $В$	$нС - вС$ $Н$ або $В$	$Н$ $С - В$	$d_2$	$d_{21}$
		$вН$ або $нВ$ $Н$ або $В$	$Н$ $нС$		$d_{22}$
$Н$ або $В$ $нС$ або $вС$	$Н$ $В$	$Н$ або $В$ $нС$ або $вС$	$Н$ $С$ або $В$	$d_3$	$d_3$
$нС$ або $вС$ $С$	$вС$ $В$	$нС$ або $вС$ $С$	$вС$ $С$ або $В$	$d_4$	$d_4$
$С$	$вС$	$С$	$вС$	$d_5$	$d_5$

Таблиця 5 – Параметри функцій належності нечітких термів і ваги правил після генетичної настройки

$x_2$	$x_1$				
	$Н, (0.30, 0.81)$	$нС, (2.43, 0.50)$	$С, (3.00, 0.62)$	$вС, (3.64, 0.55)$	$В, (5.70, 0.86)$
$В, (3.60, 0.49)$	$d_2 (0.87)$	$d_3 (0.90)$	$d_4 (0.96)$	$d_3 (0.90)$	$d_2 (0.87)$
$вС, (3.05, 0.62)$	$d_1 (0.84)$	$d_4 (0.96)$	$d_5 (1.00)$	$d_4 (0.96)$	$d_1 (0.84)$
$нС, (0.88, 0.54)$	$d_2 (0.87)$	$d_1 (0.89)$			$d_2 (0.87)$
$Н, (0.30, 0.78)$	$d_3 (0.95)$	$d_2 (0.90)$			$d_3 (0.95)$

Таблиця 6 – Параметри функцій належності нечітких термів і ваги правил після нейронної настройки

$x_2$	$x_1$						
	$H, (0.30, 0.72)$	$\bar{v}H, (0.91, 0.80)$	$HC, (2.41, 0.48)$	$C, (3.00, 0.61)$	$\bar{v}C, (3.60, 0.52)$	$HB, (5.11, 0.84)$	$B, (5.70, 0.71)$
$B, (3.60, 0.53)$	$d_{21} (0.85)$	$d_1 (0.86)$	$d_3 (0.93)$	$d_4 (0.96)$	$d_3 (0.93)$	$d_1 (0.86)$	$d_{21} (0.85)$
$\bar{v}C, (3.02, 0.61)$			$d_4 (0.96)$	$d_5 (1.00)$	$d_4 (0.96)$		
$C, (2.41, 0.54)$			$d_3 (0.93)$	$d_4 (0.96)$	$d_3 (0.93)$		
$HC, (0.91, 0.83)$	$d_{22} (0.92)$	$d_1 (0.88)$					$d_{22} (0.92)$
$H, (0.30, 0.71)$	$d_3 (0.96)$	$d_{22} (0.92)$	$d_{21} (0.86)$			$d_{22} (0.92)$	$d_3 (0.96)$

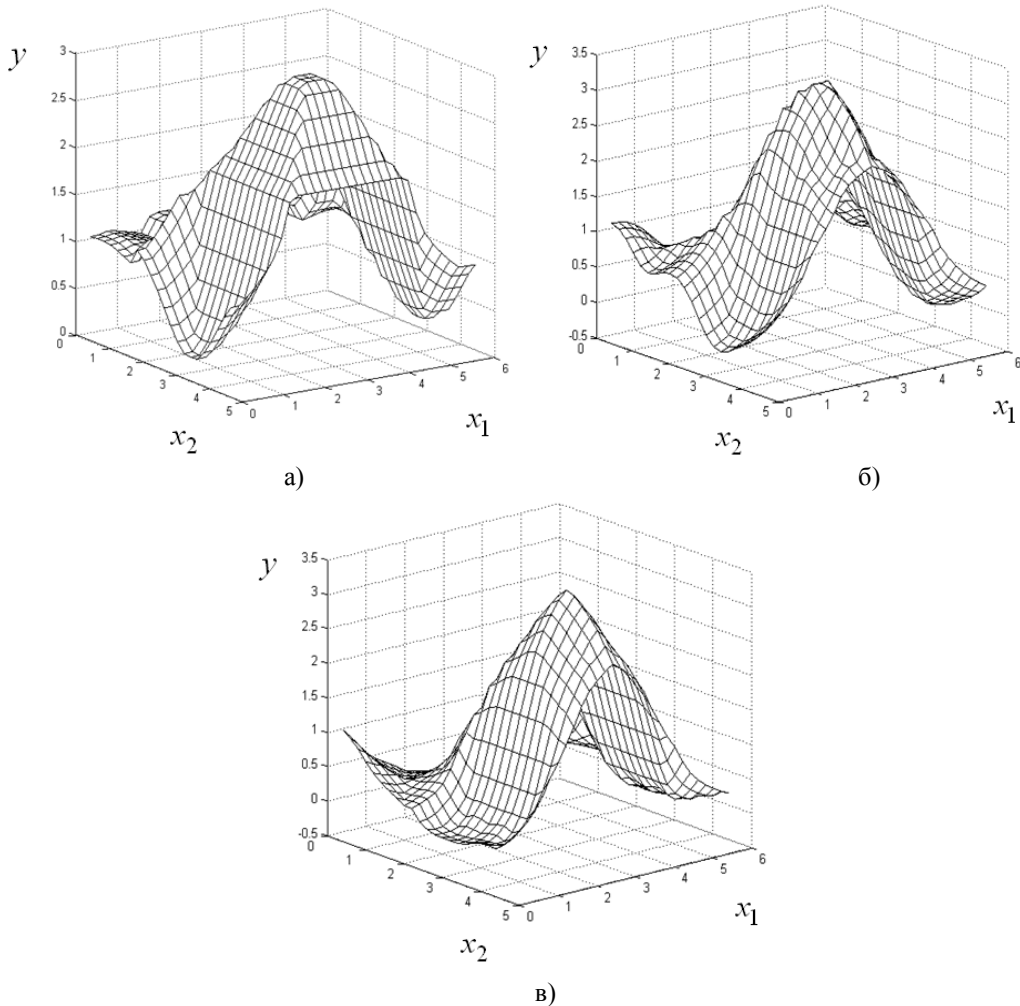


Рисунок 3 – Апроксимація нечіткими відношеннями (а); правилами для  $m = 5$  (б); правилами для  $m = 6$  (в)

**Висновки**

Запропоновано нейро-мережевий підхід, який дозволяє генерувати правила ЯКЦО-ГО шляхом розв’язання рівнянь нечітких відношень для заданих класів виходу, що є альтернативою селекції і налаштуванню структури гіпербоксів. При цьому кількість правил в класі дорівнює кількості розв’язків, а геометрія термів у правилі визначається інтервалами значень вхідних змінних.

Запропонований підхід забезпечує адаптацію структури правил до змінення границь класів виходу для заданої системи нечітких відношень. Структура початкового набору правил визначається за допомогою генетичного алгоритму. Поетапне розв'язання задач оптимізації здійснюється за допомогою рекурентних співвідношень, які відповідають навчанню *min-max* нейро-нечіткої мережі, ізоморфної лінгвістичним розв'язкам системи рівнянь нечітких відношень.

#### Список літератури

1. Gabrys B. General fuzzy min-max neural network for clustering and classification / Gabrys B., Bargiela A. // IEEE Transactions on Neural Networks. – 2000 – Vol. 11 (3). – pp. 769 – 783. – ISSN: 1045-9227.
2. Fu X.J. Linguistic rule extraction from a simplified RBF neural network / Fu X.J., Wang L.P. // Computational Statistics. – 2001. – Vol. 16(3). – pp. 361 – 372. – ISSN: 0943-4062.
3. A new approach to division of attribute space for SVR based classification rule extraction / Zhang D., Duan A., Fan Y. et al. // Advances in Neural Networks. – 2008. – Vol. 5263. – pp. 691 – 700. – ISBN 978-3-540-87731-8.
4. Yager R. Essentials of fuzzy modeling and control / Yager R., Filev D. – New York: John Willey & Sons, 1994. – 408 p. – ISBN 0-471-01761-2.
5. Peeva K. Fuzzy relational calculus. Theory, applications and software / Peeva K., Kyosev Y. – New York: World Scientific, 2004. – 304 p. – ISBN: 978-981-256-076-6.
6. Rotshtein A. Fuzzy evidence in identification, forecasting and diagnosis / A. Rotshtein, H. Rakytyanska. – Heidelberg: Springer, 2012. – 314 p. – ISBN 978-3-642-25785-8.
7. Ракитянська Г.Б. Ідентифікація нелінійних залежностей нечіткими правилами і відношеннями / Ракитянська Г.Б. // Контроль і управління в складних системах КУСС – 2012: XI Міжн. наук. конф., 9 – 11 жовтня 2012 р.: тези доп. – Вінниця: ВНТУ, 2012. – С. 255. – ISBN 966-641-187-3.
8. Rotshtein A. Expert rules refinement by solving fuzzy relational equations / Rotshtein A., Rakytyanska H. // Human System Interaction HSI – 2013: VI IEEE Conference, 6 – 8 June, 2013: Proceedings. – Sopot, Poland, 2013. – pp. 257 – 264. – ISBN 978-1-4673-5636-7.
9. Ротштейн А.П. Адаптивная система диагностики на основе нечетких отношений / Ротштейн А.П., Ракитянская А.Б. // Кибернетика и системный анализ. – 2009. – № 4. – С.135 – 150. – ISSN 0023-1274.
10. Rotshtein A. Fuzzy logic and the least squares method in diagnosis problem solving / Rotshtein A., Rakytyanska H. // In: Sarma R.D. (ed) Genetic diagnoses. – New York: Nova Science Publishers, 2011. – pp. 53 – 97. – ISBN 978-1-61324-866-9.

#### Відомості про авторів

**Ракитянська Ганна Борисівна** – к.т.н., доцент, докторант кафедри програмного забезпечення, Вінницький національний технічний університет, м. Вінниця.

УДК 665.9

Т. Б. ШАТОВСКАЯ, И. В. КАМЕНЕВА

Харьковский Национальный университет радиоэлектроники

**ПОСТРОЕНИЕ ГРАФА СВЯЗНОСТИ В АЛГОРИТМЕ КЛАСТЕРИЗАЦИИ СЛОЖНЫХ ОБЪЕКТОВ**

**Аннотация.** В статье представлена модификация алгоритма Хамелеон. Алгоритм Хамелеон состоит из следующих этапов: построение графа, огрубление, разделение и восстановление. На каждом из этапов могут быть использованы различные подходы и алгоритмы. Рассмотрено 2 вида графов: симметричный k-nn граф и асимметричный k-nn граф.

**Ключевые слова:** кластеризация, алгоритм Хамелеон, построение графа, связность, k-ближайших соседей.

**Анотація.** У роботі представлений модифікований алгоритм Хамелеон. Алгоритм Хамелеон побудований з таких етапів: побудова графа, огрубіння, поділ та відновлення. На кожному з цих етапів можуть бути використані різні підходи та алгоритми. Головною метою роботи є дослідження з покращення етапу побудови через оптимізацію алгоритму вибору k під час побудови графа k найближчих сусідів. Розглянуто 2 види графів: симетричний k-nn граф та асиметричний k-nn граф.

**Ключові слова:** кластеризація, алгоритм Хамелеон, побудова графа, зв'язність, k-найближчих сусідів.

**Annotation.** In the article, modification of Chameleon algorithm is presented. Chameleon algorithm consists of the following stages: graph construction, coarsening, partitioning and uncoarsening. At each of these steps, different algorithms and approaches can be used. The main goal of this work is investigation and improvement of graph construction stage. This can be done by modification of k-selection algorithm during k-nn graph construction. It is considered two kinds of graphs: symmetric and asymmetric.

**Key words:** clustering, Chameleon algorithm, graph construction, connectivity, k-nearest neighbors.

**Введение**

На данный момент весьма активно исследуются различные методы кластеризации. Каждым из целого множества имеющихся методов можно получить различные разбиения исходного множества. Выбор определенного метода зависит от типа желаемого результата. Производительность метода с определенными типами данных зависит от характеристик сервера и технических возможностей программного обеспечения, размера множества. Модификация алгоритма построения графа в алгоритме Хамелеон

В последнее время ведутся активные разработки новых алгоритмов кластеризации, способных обрабатывать сверхбольшие базы данных. В них основное внимание уделяется масштабируемости. Разработаны алгоритмы, в которых методы иерархической кластеризации интегрированы с другими методами. К наиболее актуальным алгоритмам относятся: BIRCH, CURE, CHAMELEON, ROCK [1]. Главной целью работы является исследование и улучшение этапа построения графа посредством оптимизации алгоритма выбора k при построении графа k ближайших соседей.

**1. Модифицированный алгоритм Хамелеон**

Хамелеон – это новый иерархический алгоритм, который преодолевает ограничения существующих алгоритмов кластеризации. Данный алгоритм рассматривает динамическое моделирование в иерархической кластеризации. В нем можно выделить следующие стадии:

1. Построение графа. Граф может быть построен симметричный или асимметричный. Различные виды расстояний могут быть применены при построении графа: Euclidian, Manhattan, Minkowski, SquEuclidian.

2. Огрубление графа (Coarsening). Огрубление графа может быть выполнено следующими методами: Random Matching(RM), Heavy Edge Matching(HEM), Light Edge Matching(LEM).

3. Начальное разделение графа (Initial Partitioning). Существует несколько подходов к разделению графов: графические методы, комбинаторные методы и спектральные методы. Также алгоритмы могут быть выполнены в рамках рекурсивной бисекции, так как большинство методов выполняет деление графа пополам.

4. Восстановление графа (Uncoarsening) и усовершенствование разделения графа (Refinement). Для улучшения разделения графа применяются следующие алгоритмы: Kernighan–Lin (KL), Boundary KL, Fiduccia–Mattheyses (FM), BoundaryFM. Эти же алгоритмы могут быть применены на этапе разделения, взяв за начальное случайное разделение огрубленного графа.

5. Объединение схожих классов для получения финального разбиения.

Целью построения графа является соединение точек локальных соседей. Точки соединяются в зависимости от типа графа.

- Граф эпсилон-окрестности (Epsilon-neighborhood graph). Две вершины графа соединены, если расстояние между рассматриваемыми объектами меньше эпсилон. Данный граф может быть взвешенным и не взвешенным. В случае взвешенного графа вес ребра равняется значению схожести соседних точек (не расстоянию). Параметр данного графа – эпсилон – устанавливается пользователем.

- Полностью связный граф (completely connected graph). Граф может быть получен из графа эпсилон-окрестности установкой эпсилон в максимальное значение.

- Симметричный граф  $k$  ближайших соседей (symmetric  $k$ -nearest neighbor graph( $k$ -nn)): две вершины  $x, y$  соединены, если  $x$  находится среди  $k$  ближайших соседей  $y$  и наоборот.
- Ассиметричный граф  $k$  ближайших соседей (mutual  $k$ -nearest neighbor graph): две вершины  $x, y$  соединены, если  $x$  находится среди  $k$  ближайших соседей  $y$  или наоборот [2].

## 2. Введение в $k$ -nn граф

Задача графа  $k$  ближайших соседей определена следующим образом: дано множество точек  $P$  из  $n$  точек в  $R^d$  и положительное целое число  $k \leq n-1$ , рассчитать  $k$  ближайших соседей для каждой точки  $P$ . Более формально задача может быть представлена следующим образом: пусть  $P = \{p_1, p_2, \dots, p_n\}$  множество точек в пространстве  $R^d$  где  $d \leq 3$ . Для каждой вершины  $p_i \in P$  пусть  $N_i^k$   $k$  точек из  $P$  ближайших к  $p_i$ . Граф  $k$  ближайших соседей ( $k$ -nearest neighbor graph ( $k$ -NNG)) - это граф где множество вершин  $\{p_1, p_2, \dots, p_n\}$  и множество ребер  $E = \{(p_i, p_j) : p_i \in N_i^k \text{ или } p_j \in N_j^k\}$  [3]. Следует отметить, что это ассиметричный граф  $k$  ближайших соседей, так как отношения близости ассиметричны.  $p_i$  может быть среди ближайших соседей  $p_j$ , но  $p_j$  нет. В симметричном графе  $p_i$  и  $p_j$  будут соединены ребром только в том случае, если каждая из них находится среди  $k$  ближайших соседей другой вершины.

В данной работе рассмотрено 2 вида графов: симметричный  $k$ -nn граф и ассиметричный  $k$ -nn граф. При построении графа для каждой пары объектов измеряется «расстояние» между ними — степень похожести. В данном случае, чем больше сходство между двумя объектами - тем тяжелее будет ребро между ними.

Еще одним важным параметром при построении графа является  $k$  – количество соседей, с которыми будет связана каждая из вершин. Граф называется *связным*, если в нем для любых двух вершин имеется маршрут, соединяющий эти вершины. При решении поставленной задачи для построения графа  $k$  должно быть выбрано таким образом, чтобы соблюдалось условие связности построенного графа. Но слишком большое значение  $k$  очень сильно увеличивает вычислительную дороговизну метода и время выполнения не только этапа построения графа, а и всех последующих этапов. Самым простым подходом для выбора  $k$  является (2.1), но и данный метод имеет вышеперечисленные недостатки.

$$k = \sqrt{n}$$

На практике применяется два принципиально различных порядка обхода, основанных на поиске в глубину и поиске в ширину соответственно.

Поиск в ширину. Вначале все вершины помечаются как новые. Первой посещается вершина  $a$ , она становится единственной открытой вершиной. В дальнейшем каждый очередной шаг начинается с выбора некоторой открытой вершины  $x$ . Эта вершина становится активной. Далее исследуются ребра, инцидентные активной вершине. Если такое ребро соединяет вершину  $x$  с новой вершиной  $y$ , то вершина  $y$  посещается и превращается в открытую. Когда все ребра, инцидентные активной вершине, исследованы, она перестает быть активной, и становится закрытой. Если на данном этапе остались незакрытые вершины - то граф несвязный.

Поиск в глубину. Главное отличие от поиска в ширину состоит в том, что при поиске в глубину в качестве активной выбирается та из открытых вершин, которая была посещена последней. Основной алгоритм тот же, что и в случае поиска в ширину, только нужно очередь заменить стеком, а процедуру BFS - процедурой DFS.

Общая оценка трудоемкости для алгоритмов одинаковая -  $O(m + n)$ .

## 3. Оптимизация выбора $k$ для построения $k$ -nn графа

Для оптимизации выбора начального параметра  $k$  при построении  $k$ -nn графа необходимо построить математическую модель зависимости  $k$  от характеристик обрабатываемой выборки. Построение математической модели выполнялось на наборе экспериментальных выборок. Набор выборок состоит из 132 выборок, среди них 33 уникальных выборки и 3 вариаций каждой из них полученной путем добавления 20%, 40% и 60% шума. Эксперимент так же проводился на наборах экспериментальных и реальных выборок полученных с ресурсов обмена наборами данных.

Целью данных экспериментов был выбор управляемых параметров данной модели зависимости, способных отобразить необходимые характеристики выборки данных. В рамках работы было проведено 3 эксперимента для выбора управляемых параметров.

- В первом эксперименте анализировались такие характеристики как: количество объектов в выборке, минимальные и максимальные значения математического ожидания, дисперсии и разброса. Зависимости между данными параметрами и значением  $k$  не выявлено.

- Во втором эксперименте в качестве управляемого параметра были выбраны длина наибольшего остова ребра полносвязного графа и среднее значение длины всех остальных ребер остова.

Данные характеристики показывают зависимость, но использование данного подхода не является целесообразным в связи с трудоемкостью построения остова полносвязного графа.

- В третьем эксперименте в качестве характеристики использовались количество компонент связности, максимальное расстояние между компонентами связности и количество элементов в компоненте связности.

В результате исследования была построена математическая модель для оптимизации выбора начального значения k при построении ассиметричного k-пн графа. Модель для ассиметричного k-пн графа имеет следующий вид и представлена на (рис. 1):

$$k = a + b \cdot x_1 + c \cdot x_2 + d \cdot x_1^2 + e \cdot x_2^2 + f \cdot x_1 \cdot x_2 + g \cdot x_1^3 + h \cdot x_2^3 + i \cdot x_1 \cdot x_2^2 + j \cdot x_1^2 \cdot x_2,$$

где  $x_1$ - коэффициент расстояния;  $x_2$  – количество компонент связности. Значения коэффициентов представлены в табл. 1.

Таблица 1 – Значения коэффициентов модели для определения k в k-пн графе.

$a$	4,963024	$f$	4,18E-04
$b$	2,33E-02	$g$	1,05E-08
$c$	0,42939	$h$	1,14E-05
$d$	-4,45E-05	$i$	1,19E-05
$e$	-3,86E-03	$j$	-4,73E-07

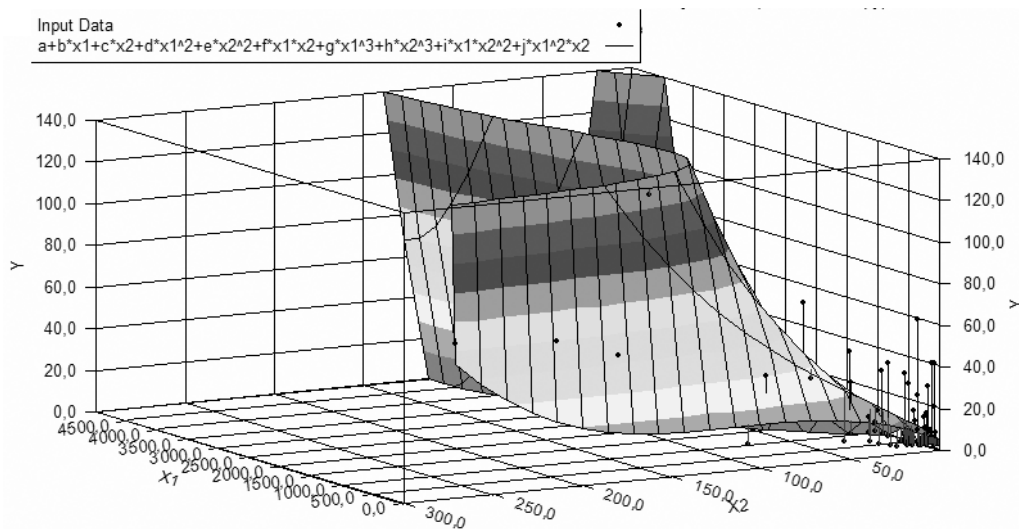


Рисунок 1 – Графическое представление описания данных математической моделью

О качестве построенной модели можно судить, исходя из следующих характеристик: стандартная ошибка оценки равна 11,2986020522291, коэффициент множественной детерминации равен 0,6452864929, статистика Дублина-Ватсона составляет 1,24157318003058. Остатки при построении данной модели представлены на рис. 3.2.

Оценки и статистики качества данной модели не являются остаточными показателями эффективности применения полученной модели, так как модель является лишь одним из этапов выбора k. Применение подхода исследовалось на 285 выборках. Применение данной модели улучшили время выполнения этапа построения графа в 62,45% случаев. В 37,55% случаев время выполнения ухудшилось.

Время выполнения ухудшилось лишь в тех случаях, когда k было меньше или равно 3 и время выполнения мало, следовательно, ухудшение временного показателя несущественно сказывается на производительности метода в целом. Отрицательный результат применения модели получен в 7,71% случаев. В среднем время выполнения улучшилось на 161%. Отрицательным результатом считается при получении k существенно большем минимально необходимого для соблюдения условия связности, даже если время построения графа уменьшилось.

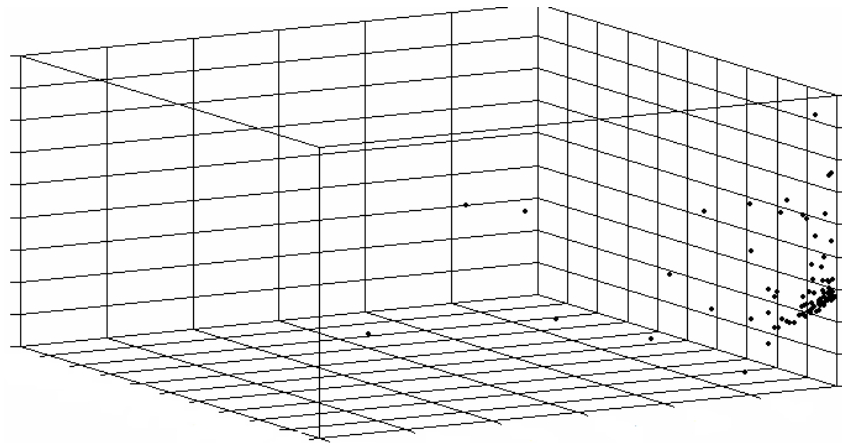


Рисунок 2 – Графическое представление остатков

Сравнение времени выполнения до и после применения модели в зависимости от количества элементов в обрабатываемой выборке представлено на рис 3, и в зависимости от полученного значения k представлено на рис. 4.

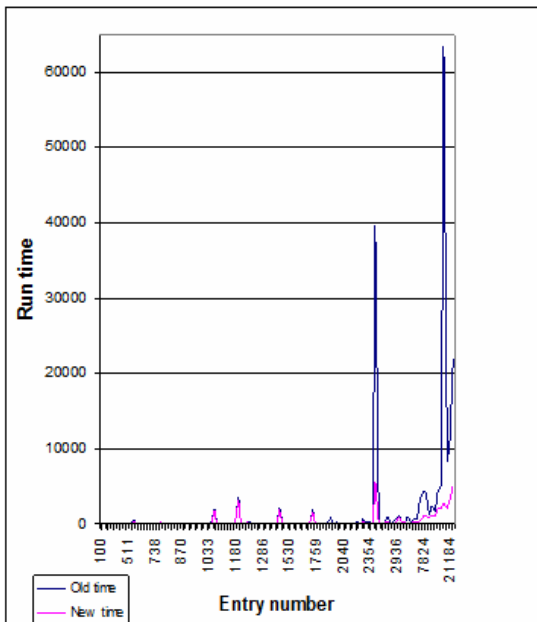


Рисунок 3 – Зависимость времени построения асимметричного графа в зависимости от количества элементов выборки для модифицированного и не модифицированного вариантов алгоритма

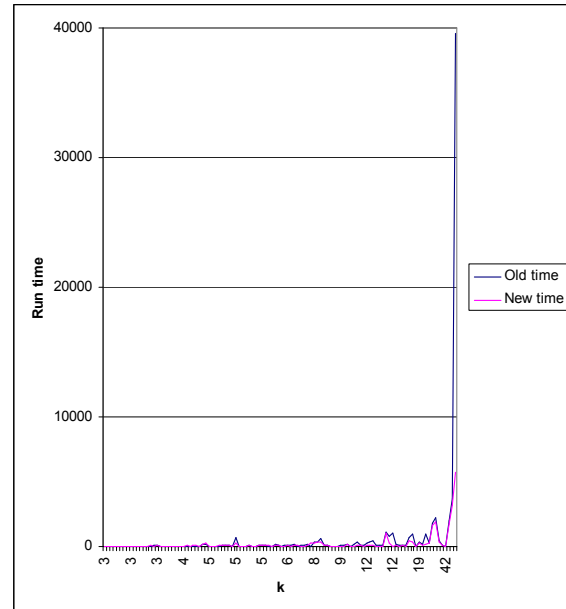


Рисунок 4 – Зависимость времени построения асимметричного графа в зависимости от полученного значения k для модифицированного и не модифицированного вариантов алгоритма

Так же в результате исследования была построена математическая модель для оптимизации выбора начального значения k при построении симметричного k-пн графа. Модель для асимметричного k-пн графа имеет следующий вид и представлена на (рис. 5):

$$k = a + b \cdot x_1 + c \cdot x_1^2 + d \cdot x_1^3 + e \cdot x_2 + f \cdot x_2^2 + g \cdot x_2^3 + h \cdot x_2^4 + i \cdot x_2^5,$$

где  $x_1$ - коэффициент расстояния;  $x_2$  – количество компонентов связности. Значения коэффициентов представлены в табл. 2.

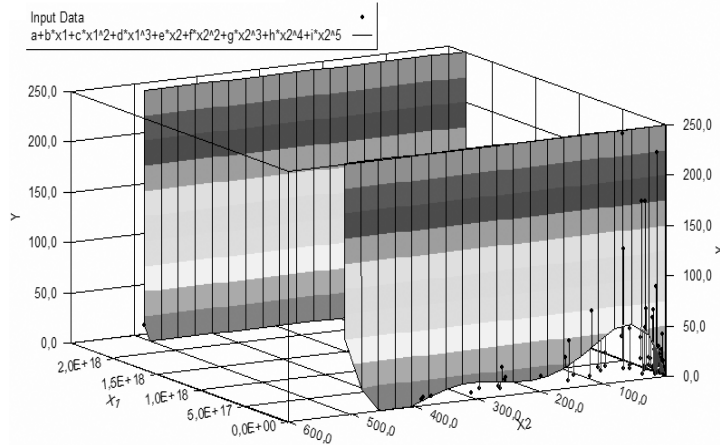


Рисунок 5 – Графическое представление описания данных математической моделью

Таблица 2 – Значения коэффициентов модели для определения k в k-nn графе.

$\alpha$	-0,547360564	f	-3,09E-02
b	-7,46E-14	g	1,55E-04
c	1,51E-29	h	-3,34E-07
d	-6,56E-48	i	2,61E-10
e	2,323285358		

О качестве построенной модели можно судить, исходя из следующих характеристик: стандартная ошибка оценки равна 42,8805641130193, коэффициент множественной детерминации равен 0,15118817, статистика Дублина-Ватсона составляет 1,26055939255469. Остатки при построении данной модели представлены на рис. 6.

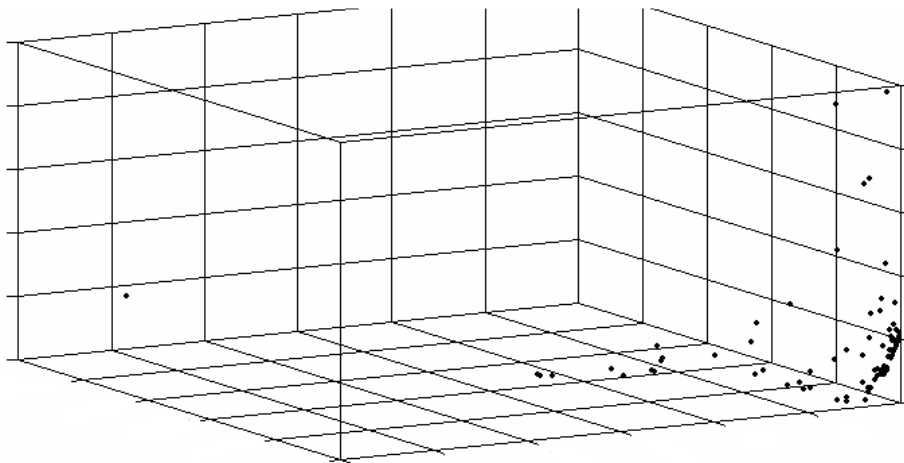


Рисунок 3.6 – Графическое представление остатков

Применение данной модели улучшило время выполнения этапа построения графа в 69,23% случаев. В 20,51% случаев время выполнения ухудшилось. Отрицательный результат применения модели получен в 5,12% случаев. В среднем время выполнения улучшилось на 169%.

Сравнение времени выполнения до и после применения модели в зависимости от количества элементов в обрабатываемой выборке представлено на рис 7, и в зависимости от полученного значения k представлено на рис. 8.

Использование модели особенно критично для больших выборок. Полученные результаты будут использованы для дальнейших исследований и модификаций алгоритма Хамелеон.

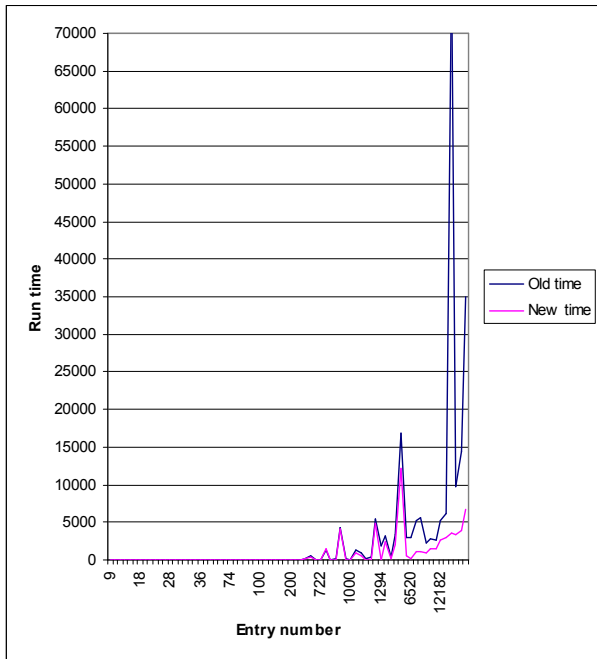


Рисунок 7 – Зависимость времени построения симметричного графа в зависимости от количества элементов выборки для модифицированного и не модифицированного вариантов алгоритма

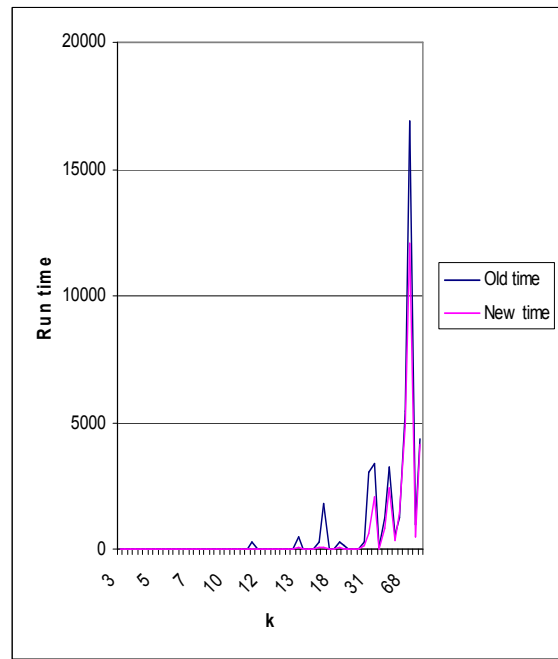


Рисунок 8 – Зависимость времени построения симметричного графа в зависимости от полученного значения k для модифицированного и не модифицированного вариантов алгоритма

#### 4. Математическая модель выбора алгоритмов

Управляемые параметры – характеристики выборки (максимальные и минимальные значения математического ожидания, дисперсии, разброса и вычисляемый параметр).

Составляющими целевого параметра являются:

- алгоритм построения графа;
- мера расстояния;
- алгоритм огрубления графа;
- алгоритм начального разделения графа;
- мера сходства классов;
- алгоритм восстановления графа.

На основании имеющихся алгоритмов составлено 14784 комбинаций для анализа.

Математическая модель, полученная на основании этих данных имеет вид:

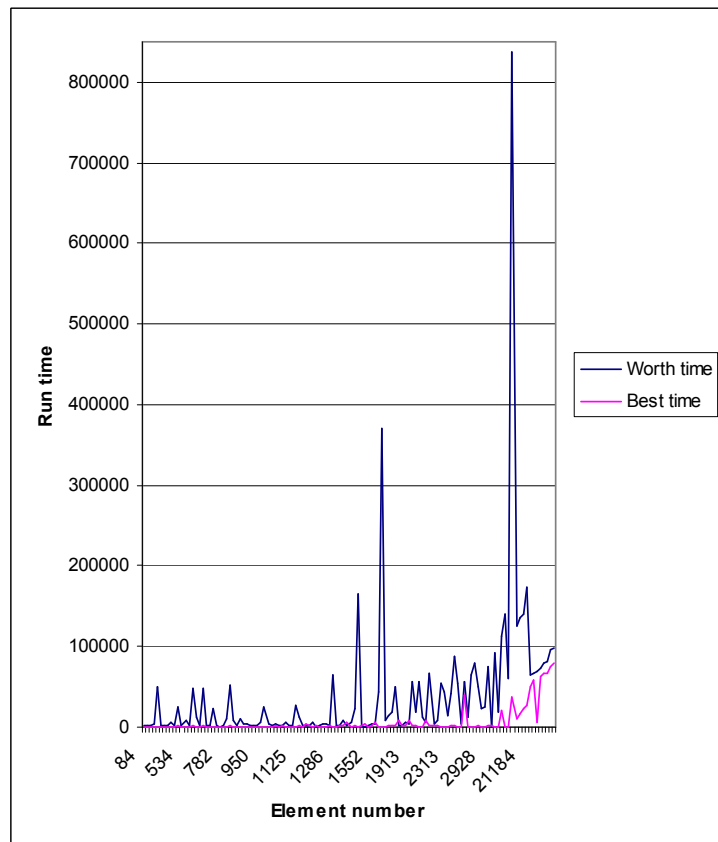


Рисунок 9 – Зависимость времени построения симметричного графа в зависимости от количества элементов выборки для лучшего и худшего вариантов алгоритма.

$$Y = a*x_1 + b*x_2 + c*x_3 + d*x_4 + e*x_5 + f*x_6 + g*x_7 + h*x_8 + i*x_9 + j,$$

где  $x_1 - x_9$  соответствуют характеристикам выборок, на основании которых построена модель.

Следует отметить, что в ходе исследования было установлено нецелесообразность использования симметричного алгоритма построения графа и в дальнейших экспериментах будет использован асимметричный алгоритм. Построение математической модели проводилось на основании результатов экспериментов разделения экспериментальных выборок с помощью различных комбинаций алгоритмов в рамках модифицированного алгоритма Хамелеон.

#### Выводы

В данной статье разработана математическая модель для выбора алгоритмов в рамках модифицированного алгоритма Хамелеон, построена математическая модель для выбора  $k$  при построении  $k$ -nn графа в рамках модифицированного алгоритма Хамелеон, приведены результаты применения разработанных методов на реальных данных.

Была разработана математическая модель для выбора  $k$  при построении  $k$ -nn графа в рамках модифицированного алгоритма Хамелеон. Приведены результаты, полученные как на экспериментальных, так и на реальных данных.

Результаты работы позволили усовершенствовать этап построения графа путем модификации алгоритма Хамелеон с целью улучшения процессов кластеризации, ориентированных на работу с очень большими базами данных.

#### Список использованной литературы

1. Чубукова И.А. Data Mining БИНОМ / А.И. Чубукова // Лаборатория знаний. Интернет-университет информационных технологий. – ИНТУИТ.ру. – 2008.
2. Brian Read Advances in Databases : 18th British National Conference on Databases, BNCOD 18 Chilton, UK, July 9 – 11. – 2001.
3. Karypis G. Multilevel  $k$ -way Partitioning Scheme for Irregular Graphs / G. Karypis, V. Kumar // Journal of parallel and distributed computing. – 1998. – № 48. – P. 96-129
4. Karypis G. A fast and highly quality multilevel scheme for partitioning irregular graphs / G. Karypis, V. Kumar // SIAM J. Sci. Comput., to appear. [Also available on WWW at URL <http://www.cs.umn.edu/~karypis>]. – 1995.
5. Karypis G. Multilevel  $k$ -way Partitioning Scheme for Irregular Graphs / G. Karypis, V. Kumar // Society of Industrial and Applied mathematics. – 1999.
6. Karypis G. Chameleon: Hierarchical Clustering Using Dynamic Modeling / G. Karypis, E.-H. (Sam) Han, V. Kumar // Computer. – 1999. – Vol. 32, № 8. – P. 68-75.

#### Информация об авторах

**Шатовская Татьяна Борисовна** – к.т.н. доцент кафедры Программной инженерии, Харьковский Национальный университет радиоэлектроники.

**Каменева Ирина Витальевна** – к.т.н. старший преподаватель кафедры Программной инженерии, Харьковский Национальный университет радиоэлектроники.