

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

УДК 681.518.25:004.056

В.М. ДУБОВОЙ, Г.Ю. ДЕРМАН

Вінницький національний технічний університет, Вінниця

ПІДХІД ДО ВИЗНАЧЕННЯ РІВНЯ БЕЗПЕКИ СИСТЕМИ, ЩО РОЗВИВАЄТЬСЯ

Анотація. Розрахунок ризику базується на прогнозі розвитку структури інформаційної системи. Наведені деякі фактори і характеристики небезпеки неузгодженого розвитку. Для розрахунку ймовірності небезпеки використана експоненційно-мультиплікативна апроксимація. Для прогнозування розвитку використана модифікована логістична крива.

Ключові слова: інформаційна система (ІС), безпека інформаційної системи, розвиток.

Аннотация. Расчет риска базируется на прогнозе развития структуры информационной системы. Приведены некоторые факторы и характеристики опасности несогласованного развития. Для расчета вероятности опасности использована экспоненциально-мультипликативная аппроксимация. Для прогнозирования развития использована модифицированная логистическая кривая. Текст аннотации.

Ключевые слова: информационная система (ИС), безопасность информационной системы, развитие.

Abstract. The calculation of risk based on forecasts of the structure of information system. The following are some factors and characteristics of risk inconsistent development. To calculate the probability of danger used exponentially-multiplicative approximation. To predict the development used a modified logistic curve.

Key words: information system, security of information system, development.

Вступ

Інформаційні системи, забезпечуючи постійно зростаючі потреби суспільства у отриманні, обробці, збереженні, пошуку та поданні інформації, знаходяться у стані відповідного постійного розвитку. При цьому розвиваються усі компоненти ІС: програмні і апаратні, призначені для передавання, зберігання і обробки інформації тощо. Розвиток має як параметричний характер (заміна окремих елементів ІС на інші з кращими параметрами), так і структурний (зміна функцій, кількості, якості елементів та зв'язків між ними). Будь-які зміни в ІС можуть приводити до погіршення рівня її безпеки. Таким чином, забезпечення такої стратегії розвитку ІС, при якому принаймні не погіршується рівень її безпеки, є важливою проблемою.

Наразі є низка міжнародних і національних стандартів та наукових досліджень з безпеки інформаційних систем (інформаційних технологій) [1, 2, 3, 4]. Відповідно до документа [4], який є основою організації системи безпеки інформаційних технологій у Росії і відповідає міжнародним стандартам безпеки [1, 2, 3], «Под безопасностью информационной технологии понимается состояние ИТ, определяющее защищенность информации и ресурсов ИТ от действия объективных и субъективных, внешних и внутренних, случайных и преднамеренных угроз, а также способность ИТ выполнять предписанные функции без нанесения неприемлемого ущерба субъектам информационных отношений». Хоча це означення охоплює усі можливі випадки і наслідки порушення безпеки, проте на практиці розглядаються переважно випадки появи у складі ІС програмно-апаратного забезпечення, яке може містити джерело загрози або само по собі, або у разі зовнішніх злочинних дій. Проте важливим фактором зменшення рівня безпеки ІС може бути неузгоджений розвиток окремих компонентів ІС, при тому, що кожен компонент сам по собі не несе загрози безпеці.

Актуальність

На даний час існує багато моделей комплексного захисту інформаційних систем, а також великий вибір технічних та програмних засобів захисту інформації, методик інформаційної безпеки [5, 6], але відсутня методика визначення рівня безпеки саме систем, що розвиваються.

Рівень безпеки прийнято оцінювати величиною ризику R , а забезпечення безпеки – це процес управління ризиками. Для визначення рівня безпеки потрібно виконати оцінку ризиків функціонування кожної окремої складової, що буде розвиватися і вдосконалюватися в процесі розвитку системи.

За узагальнений показник рівня безпеки ІС можна взяти комплексну оцінку ризику. Для обчислення ризику використовують статистичну оцінку [7]:

$$R = \sum_j^N P_j g_j, \quad (1)$$

де P_j – ймовірність j -го варіанту подій, $(0 \leq P_j \leq 1)$, $\sum_j P_j = 1$;

g_j – втрати від реалізації j -го варіанту подій.

Визначення рівня безпеки ґрунтується на прогнозуванні процесу розвитку ІС та появи небезпек. Низка досліджень показала, що S-подібна логістична крива з високою ймовірністю описує розвиток різних систем [8, 9, 10, 11]. Кожен із параметрів системи, що розвивається, можна описати виразом:

$$S(t) = \frac{K \cdot S_0 e^{rt}}{K + S_0 (e^{rt} - 1)}, \quad (2)$$

де Γ – параметр, що характеризує швидкість росту (розвитку), K –максимальна границя параметра, S_0 – початкове значення, $S(t)$ – значення параметра на момент часу t .

Події, які можуть стати чинниками небезпеки інформаційних систем, перераховані в [12], але вони кардинально відрізняються від тих, що можуть стати причиною порушень безпеки ІС, що розвивається:

- зниження рівня надійності при збільшенні складності системи, якщо тільки це ускладнення не пов'язане з введенням елементів, спеціально призначених для підвищення надійності (резервних потужностей);
- впровадження нового програмного забезпечення, що розширює функції системи, може приводити до боротьби програм за ресурси системи, а ці ресурси завжди є обмеженими;
- зниження захищеності системи від вірусних, хакерських та інших атак при збільшенні кількості каналів доступу та чисельності персоналу.

Мета

Основною метою даної статті є удосконалення підходу до визначення рівня безпеки системи, що розвивається.

Розв'язання задач

Розрахунок ризику повинен базуватися на прогнозі розвитку структури ІС, її функцій, кількості інформації, що зберігається, кількості запитів, їх складності з одного боку та інтенсивності і небезпеки атак (хакерських, вірусних та ін.) з іншого.

Зростання кількості доменів в зоні COM відображає історію розвитку Інтернету за останні роки [13]. Як видно з рис.1, бурхливе зростання кількості доменних імен в зоні COM в період з 1998-го по 2001 рік (період росту доткомів – період «зростання»), 2001-2002 більш стабільні значення (фаза «насичення»), сповільнився у 2002-му, в 2003 році відбувся спад (фаза «згасання»), а на початку 2004-го відновилося зростання реєстрацій (фаза «зародження» нового життєвого циклу).

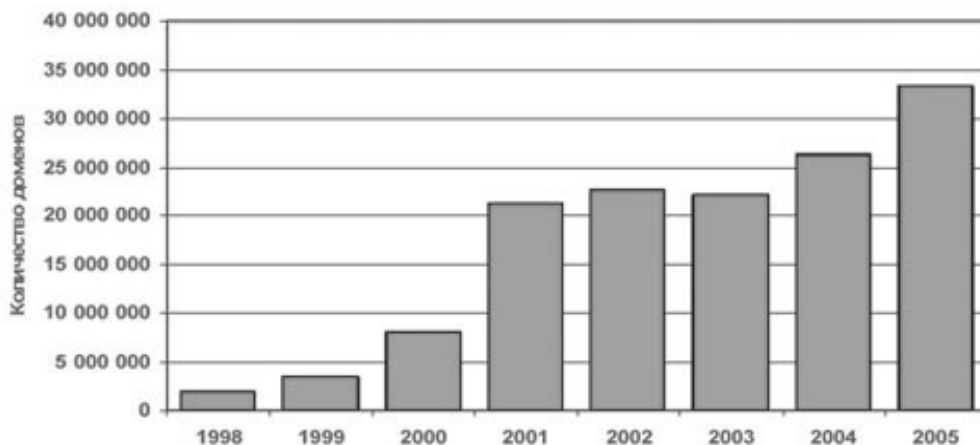


Рисунок 1 – Кількість зареєстрованих доменних імен другого рівня в домені COM в період з 1.01.1998 р. по 1.01.2005 р. (джерело: RU-CENTER). K –максимальна кількість доменних імен другого рівня, що може бути створена, r – параметр, що характеризує швидкість росту (розвитку), S_0 – початкова кількість доменних імен, S – кількість доменних імен на момент часу t .

Розвиток Інтернету у світі можна прослідкувати з графіка, представленого на рис.2.

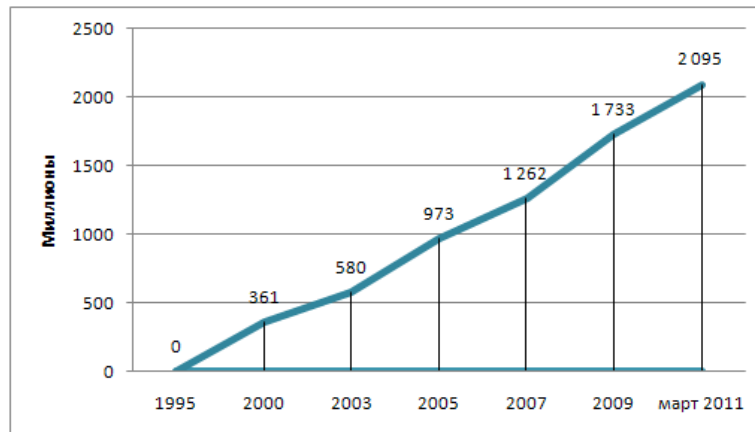


Рисунок 2 - Розвиток Інтернету у світі

З рис.1 і 2 видно, що процес розвитку може бути поданий послідовними фрагментами логістичної кривої. Таким чином, процес розвитку ІС можна подати системою рівнянь

$$\begin{cases} n(t) = \text{int} \left[\frac{t}{T} \right]; \\ \tau(t) = t - n(t) \cdot T; \\ S'(\tau) = \frac{K \cdot S_0 e^{r\tau}}{K + S_0 (e^{r\tau} - 1)}; \\ S(t) = S'(T) \cdot n(t) + S'(\tau), \end{cases} \quad (3)$$

де T - інтервал дії одної ділянки типу логістичної кривої; $\text{int}[\bullet]$ - функція виділення цілої частини аргументу; $n(t)$ - номер ділянки логістичного типу; $\tau(t)$ - інтервал часу від початку чергової ділянки логістичного типу; $S'(\tau)$ - окрема ділянка логістичного типу; $S(t)$ - крива розвитку.

Враховуючи особливість розвитку ІС, яка полягає у постійному пришвидшенні процесів розвитку, інтервал ділянки логістичної кривої будемо розглядати як функцію вигляду $T = \frac{T_0}{[n(t)]^\nu}$, де ν - показник темпу прискорення. Параметри K , S_0 , T , r підлягають ідентифікації для кожного процесу розвитку окремого параметру ІС.

Небезпеку несуть не тільки самі атаки, але й неузгодженість окремих аспектів розвитку, що призводить до порушення здатності ІС протистояти атакам і виконувати свої функції.

Для оцінювання ризику порушення безпеки за формулою (1) необхідно отримати оцінки ймовірностей P варіантів порушення безпеки в результаті неузгодженого розвитку окремих параметрів ІС. Очевидно, залежність ймовірності $P(\bar{x})$ від параметрів ІС повинна задовольняти умови:

- область визначення $0 \leq \bar{x} < \infty$;
- область значень $0 \leq P(\bar{x}) \leq 1$;

- вектор параметрів \bar{x} складається з двох підмножин: підмножини \bar{x}_1 параметрів, які менші за оптимальне значення \bar{x}_{10} , і підмножини \bar{x}_2 тих, які більші за оптимальне значення \bar{x}_{20} . Відповідно, збільшення значень \bar{x}_1 буде зменшувати ймовірність небезпеки, а збільшення \bar{x}_2 буде її збільшувати.

Апроксимуємо цю залежність функцією $P = e^{-k \frac{\bar{x}_1}{\bar{x}_2}}$, $\bar{x}_1 < \bar{x}_{10}$, $\bar{x}_2 > \bar{x}_{20}$, яка задовольняє усі умови і має додаткові переваги диференційованості на усій області визначення. Коефіцієнт k підлягає ідентифікації для кожної ІС і пари підмножин її параметрів $\{\bar{x}_1, \bar{x}_2\}$.

Вигляд апроксимуючої поверхні показаний на рис.3.

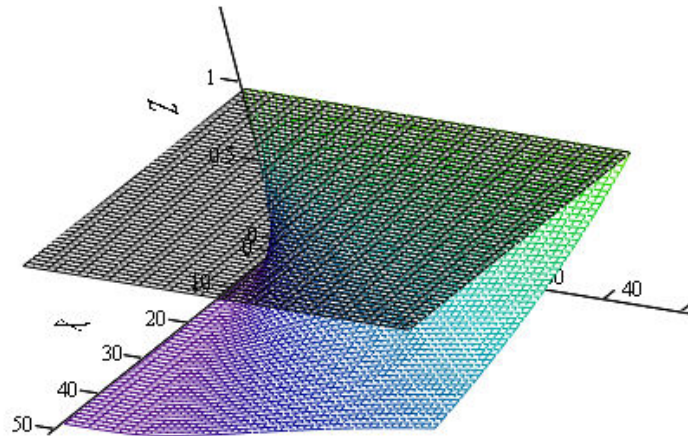


Рисунок 3 – Ймовірність втрати працездатності ІС при неузгодженому розвитку параметрів

Загальна ймовірність безпеки

$$P_{dang} = 1 - \prod_i (1 - P_i) \tag{4}$$

де P_i – ймовірність безпеки від кожного фактора.

Деякі фактори і характеристики безпеки неузгодженого розвитку наведені у табл. 1 (наведений лише фрагмент таблиці).

Таблиця 1 – Фактори безпеки неузгодженого розвитку

Параметри ІС, які розвиваються	Параметри ІС, які розвиваються неузгоджено (уповільнено)	Характер безпеки	Оцінка ймовірності безпеки (апроксимована модель)
1	2	3	4
Збільшення функцій	Розвиток системи захисту від атак	Ускладнення програмного забезпечення приводить до збільшення кількості точок вірусних і хакерських атак	$P = e^{-k_1 \frac{S_d}{n_f n_a}}$, де P - ймовірність втрати працездатності; n_f - кількість функцій; n_a - кількість атак; S_d - складність системи захисту; k_1 - коефіцієнт, який враховує розмірність величин та особливості конкретної ІС.
	швидкість обробки запитів	Затримка виконання процесів, перевантаження системи, відхилення запитів	$P = e^{-k_1 \frac{n_a}{n_f}}$, де P - ймовірність збоїв роботи; n_f - кількість функцій; n_a - швидкість обробки; k_1 - коефіцієнт, який враховує розмірність величин та особливості конкретної ІС.
	Обсяг пам'яті	Перевантаження системи, збої роботи	$P = e^{-k_1 \frac{S_d}{n_f n_a}}$, де P - ймовірність перевантаження системи; n_f - об'єм інформації, що вже зберігається; n_a - середній об'єм інформації, що зберігається при виконанні одної функції; S_d - об'єм пам'яті; k_1 - коефіцієнт, який враховує розмірність величин та особливості конкретної ІС.

продовження табл. 1

1	2	3	4
Збільшення кількості користувачів	Розвиток системи захисту від атак	Ускладнення програмного забезпечення приводить до збільшення кількості точок вірусних і хакерських атак	$P = e^{-k_1 \frac{n_a}{n_f}}$, де P - ймовірність втрати робочого часу; n_f - кількість користувачів; n_a - кількість атак; k_1 - коефіцієнт, який враховує особливості конкретної ІС.
	Необхідна швидкість обробки запитів	Затримка виконання процесів, перевантаження системи, відхилення запитів	$P = e^{-k_1 \frac{n_a}{n_f}}$, де P - ймовірність перевантаження системи; n_f - кількість користувачів; n_a - швидкість обробки запитів; k_1 - коефіцієнт, який враховує особливості конкретної ІС.
	Кількість каналів зв'язку	Перевантаження системи, збої роботи, Черги виконання операцій, втрата робочого часу	$P = e^{-k_1 \frac{n_a}{n_f}}$, де P - ймовірність перевантаження системи; n_f - кількість користувачів; n_a - швидкість обробки запитів; k_1 - коефіцієнт, який враховує особливості конкретної ІС.
	Необхідна швидкість каналів	Черги виконання операцій, втрата робочого часу	$P = e^{-k_1 \frac{n_a}{n_f}}$, де P - ймовірність втрати робочого часу; n_f - кількість користувачів; n_a - швидкість каналів; k_1 - коефіцієнт, який враховує особливості конкретної ІС.
	Необхідна швидкість обробки запитів	Затримка виконання процесів, перевантаження системи, відхилення запитів	$P = e^{-k_1 \frac{n_a}{n_f}}$, де P - ймовірність перевантаження системи; n_f - швидкість каналів; n_a - швидкість обробки запитів; k_1 - коефіцієнт, який враховує особливості конкретної ІС.
	Розвиток системи захисту від атак	Ускладнення програмного забезпечення приводить до збільшення кількості точок вірусних і хакерських атак	$P = e^{-k_1 \frac{n_a}{n_f}}$, де P - ймовірність втрати працездатності; n_f - швидкість каналів; n_a - кількість атак; k_1 - коефіцієнт, який враховує особливості конкретної ІС.
...

Оскільки, різноманітні фактори безпеки/небезпеки (див. вище) не являються незалежними, то проста адитивна формула (1) не підходить. Можна скористатися нечітким підходом на основі бази правил, оскільки метою роботи є отримання моделі для прогнозування ризиків розвитку ІС на основі комплексу факторів, кожен з яких змінюється у часі за законом (3), проте з різними параметрами.

Висновки

Розрахунок ризику базується на прогнозі розвитку структури ІС. Логістична крива використана при характеристиці розвитку різних сторін потенціалу ІС і її положення у зовнішньому середовищі. Будь-які зміни в ІС можуть приводити до погіршення рівня її безпеки. Небезпеку несуть не тільки атаки, але й неузгодженість окремих аспектів розвитку, що призводить до порушення здатності ІС виконувати свої функції. Наведені деякі фактори і характеристики безпеки неузгодженого розвитку. Для розрахунку ймовірності безпеки використана експоненційно-мультиплікативна апроксимація. Це дозволяє вра-

хувати як загрози безпеці, так і розвиток системи захисту, яка повинна включати як захист від атак, так і від структурно-параметрично-функціональних невідповідностей.

Список літератури

1. Information technology - Security techniques - Evaluation criteria for IT security -Part 1: Introduction and general model. -- ISO/IEC 15408-1.1999.
2. Information technology - Security techniques - Evaluation criteria for IT security -Part 2: Security functional requirements. -- ISO/IEC 15408-2.1999.
3. Information technology - Security techniques - Evaluation criteria for IT security -Part 3: Security assurance requirements. -- ISO/IEC 15408-3.1999.
4. ГОСТ Р ИСО/МЭК 15408-1-2002. Информационная технология. Методы и средства обеспечения безопасности. Критерии оценки безопасности информационных технологий. Часть 1. Введение и общая модель. -- М.: ИПК Издательство стандартов, 2002.
5. Грайворонський М. В., Новіков О. М. Безпека інформаційно-комунікаційних систем. – К.: Видавнича група BHV, 2009. – 608 с.
6. Домарев В.В. "Безопасность информационных технологий. Методология создания систем защиты" – К.: ООО "ТИД "ДС", 2002 – 688 с.
7. Гранатуров В.М. Экономический риск. – М.: Дело и Сервис, 1999. – 112 с.
8. Дуброва Т.А. Статистические методы прогнозирования в экономике. /М. Московский международный институт эконометрики, информатики, финансов и права, 2003. – 50 с.
9. Мартино, Дж. Технологическое прогнозирование / Дж. Мартино. М. : Прогресс, 1977. – 591 с.
10. Красильников, О. Ю. Структурные сдвиги в экономике / О. Ю. Красильников. Саратов : Изд-во Саратов. гос. ун-та, 2001. – 171 с.
11. Альтшуллер, Г. С. Творчество как точная наука. ТРИЗ / Г. С. Альтшуллер.М. : Сов. радио, 1979. – 184 с.
12. Передерій Л.В. Системний аналіз безпеки інформаційних систем/Науковий вісник Донбасу. – 2010. – №1 (9). [електронний ресурс] Режим доступу до журналу: http://almater.luguniv.edu.ua/magazines/elect_v/NN9/09plvbis.pdf
13. Internet World Stats. Usage and population statistic. [електронний ресурс] Режим доступу: www.internetworldstats.com

Відомості про авторів

Дубовой Володимир Михайлович – д.т.н., професор, завідувач кафедри КСУ, Вінницький національний технічний університет, Хмельницьке шосе, 95, м. Вінниця, 21021, тел. 59-81-57.

Дерман Галина Юрїївна – аспірантка кафедри КСУ, Вінницький національний технічний університет, Хмельницьке шосе, 95, м. Вінниця, 21021, e-mail: fortyna1000@mail.ru.

УДК 681.3:378.146

Н.О. РІЗУН

Дніпропетровський університет економіки та права, Дніпропетровськ

МЕТОДОЛОГІЯ СИСТЕМНОГО ЕКСПРЕС-АНАЛІЗУ ЯКОСТІ ТЕСТОВОГО МАТЕРІАЛУ

Анотація. Запропоновано методологію системного експрес-аналізу якості тестового матеріалу як комплексного поєднання технології організації навчального процесу із процедурою гарантованого отримання матриці результатів тестування за одне навчальне заняття із додатковою реалізацією системної експертної оцінки динамічних змін рівня знань та умінь осіб, які навчаються, у реальному масштабі часу.

Ключові слова: експрес-аналіз, якість тестового матеріалу, динамічні матриці результатів, експертна оцінка

Аннотация. Предложена методология системного экспресс-анализа качества тестового материала как комплексного объединения технологии организации учебного процесса с процедурой гарантированного получения матрицы результатов тестирования за одно учебное занятие с дополнительной реализацией системной экспертной оценки динамических изменений уровня знаний и умений обучающихся в реальном масштабе времени.

Ключевые слова: экспресс-анализ, качество тестового материала, динамические матрицы результатов, экспертная оценка

Abstract. The methodology of the system express-analysis of the testing material quality as the complex integration of technology of study process organization with the procedure of the guaranteed receipt of the testing results matrix during a single lesson (with the additional realization of system expert evaluation of dynamic changes of the level of students' knowledge and skills in the real time scale) if offered.

Key words: express-analysis, testing material quality, dynamic results matrix, expert evaluation

Вступ

Специфіка педагогічної технології полягає у тому, що якісний навчальний процес обов'язково повинен гарантувати досягнення мети, що поставлена. Однією з головних гарантій якісного результату навчання та, як наслідок, досягнення поставленої педагогічної мети, є оперативний зворотній зв'язок, який охоплює весь навчальний процес. Тобто, на усьому протязі навчання необхідна ефективна технологія оцінки поточних результатів та гнучка й оперативна корекція структури, состава та методології організації навчального процесу. Мета навчання формалізується у його результатах, що відбиті у діяльності осіб, що навчаються, причому таких, які викладач може надійно впізнати. Саме тому необхідним елементом ефективної технології навчання є аналіз якості інформації, що запропонована для вимірювання рівня знань та умінь осіб, що навчаються.

Тестові форми контролю знань студентів на сучасному етапі є одними з найбільш перспективних технологічних засобів організації зворотного зв'язку та підвищення ефективності процесу навчання у цілому. До безперечних переваг тестових технологій контролю знань відносяться стандартизація умов та результатів, оперативність та економічність, якісний диференційований характер оцінки, оптимальна складність, надійність, справедливість. Проте, серед недоліків цього інструменту відзначають небезпеку автоматичних помилок, втрату індивідуального підходу, низьку репродуктивність знань, неадекватну складність та інше [1, 2]. Однією з первинних причин виникнення проблем, що значно знижують ефективність ідентифікації результатів тестового контролю та ефективність управління навчальним процесом, є саме низька якість тестового матеріалу, що використовується, та відсутність сучасних практичних методологій та технологій її аналізу та контролю.

Аналіз останніх публікацій

Сучасні наукові дослідження по вдосконаленню класичної системи оцінки якості тестового матеріалу можливо розділити на наступні категорії: використання у матриці результатів тестування безперервної шкали [3, 4] або адаптивного шкалювання рівнів складності завдань [5, 6], що дозволяє більш тонко виміряти рівень знань студентів; створення комплексної експертизи якості тестових завдань та тестів при формуванні банку даних, що включає етапи попередньої, внутрішньої та технологічної експертизи по показникам надійності та валідності [7]; експертна оцінка якості тестових завдань [8], які, не зважаючи на переваги кожного з них, дозволяють сформулювати матрицю результатів тільки за умов наявності статистичної вибірки відповідей на усі питання тестового матеріалу не менш 100-150 особами, що навчаються [9, 10, 11, 12].

Невирішені частини загальної проблеми

Недоліками більшості розроблених на сучасний момент методів та моделей вдосконалення є теорії тестування є приділення більшої уваги розробці інтерфейсу та технології організаційної підтримки процесу тестування, та наявність ряду проблем із реалізацією технології оцінки та аналізу якості тестового матеріалу, а саме:

по-перше, через довгочасність та трудомісткість отримання повної матриці результатів тестування на початкових етапах тестового контролю системами використовується тестовий матеріал, якість якого не пройшла експертизи, та, як наслідок, ідентифікація результатів виміру знань не може визнаватися достатньо об'єктивною;

по-друге, формування матриці результатів згідно класичній теорії тестування припускає вірогідність повторного отримання одного й того ж тестового завдання особою, що тестується, із запам'ятовуванням тільки останньої відповіді, що значно знижує точність та адекватність результатів тестового контролю;

по-третє, очевидно, що результати обробки достатнього для формування матриці результатів статистичного масиву носять усереднений характер та не відображують індивідуальних характеристик та особливостей середовища учнів, у якому виконуються виміри;

у-четвертих, класична теорія тестування передбачає отримання статичного "знімку" результатів оцінки знань, але не передбачають аналізу динамічних змін рівня знань та умінь осіб, що навчаються, у реальному масштабі часу.

Тобто, згідно з класичною теорією тестування вірогідність отримати повну матрицю результатів відповідей на M із N питань за одно навчальне заняття дуже мала та складається з наступних складових:

вірогідності P_A випадіння кожному з K осіб, що тестуються, одного й того ж набору із M питань

$$\text{при вибірці із } N \text{ питань: } P_A = \left(\frac{1}{C_n^m} \right)^k.$$

Тоді, вірогідність отримання повної матриці результатів відповідей при вибірці випадковим образом, наприклад, 10 питань із 30 за один тестовий сеанс ($M=10$, $N=30$) для 10 осіб, що навчаються ($K=10$) дорівнює

$$P_A = \left(\frac{1}{C_{30}^{10}} \right)^{10} = \left(\frac{1}{30045015} \right)^{10} = 1,67 * E^{-75}, \quad (1)$$

тобто дуже мала та потребує проведення великої кількості тестових сеансів.

крім того, якщо, із урахуванням отриманої вірогідності P_A , із метою прискорення процесу отримання повної матриці результатів виконувати послідовно декілька сеансів тестування тих самих осіб, що навчаються (наприклад, на початку та наприкінці навчального заняття), на тому ж самому тестовому матеріалі, вірогідності P_B випадіння кожній з K осіб, що тестуються, різного (відмінного від отриманого під час попереднього сеансу) набору із M питань при вибірці із N питань при R сеансах тестування (наприклад $R=2$):

$$P_B = \left(\frac{1}{m!} \right)^{R-1} = \left(\frac{1}{10!} \right)^{2-1} = \left(\frac{1}{3628800} \right) = 2,76 * E^{-07}, \quad (2)$$

тобто теж дуже мала та, як вже відзначалось, свідчить про спотворення та надмірне усереднення отриманих результатів виміру знань та зниження точності та адекватності процесу тестового контролю.

Мета статті

Метою статті є вдосконалення класичної теорії тестового контролю шляхом розробки методології системного експрес-аналізу якості тестового матеріалу, що забезпечує оптимізацію процедури експертизи якості тестового матеріалу та зростання об'єктивності процесу оцінки рівня засвоєння навчального матеріалу через скорочення тривалості та зниження трудомісткості процесу формування повної кореляційної матриці результатів тестування та розширення спектру критеріїв експертизи якості тестового матеріалу завдяки уведенню у розгляд показників динамічних змін рівня знань та умінь осіб, що навчаються, у реальному масштабі часу.

Результати досліджень

Методологія системного експрес-аналізу якості тестового матеріалу, що пропонується автором, базується на наступних евристичних:

Проведення безпосередньо навчального заняття передбачає самостійне опанування навчального матеріалу за визначеною викладачем темою.

Тестовий матеріал повинен бути ієрархічно структурований та зкорельований за темами та окремими питаннями навчального матеріалу.

Попереднє тестування, що виконується шляхом подання на початку навчального заняття визначеної викладачем частки загальної кількості тестових завдань, дозволяє отримати інформацію про початковий рівень знань осіб, що навчаються.

Контрольне тестування, що виконується шляхом подання наприкінці навчального заняття частки тестових завдань, що залишилася, дозволяє отримати інформацію про динамічне кількісне та якісне змінювання знань осіб, що навчаються.

Гарантоване отримання матриці результатів тестування із двох неповних матриць – результатів попереднього та контрольного тестування у межах одного навчального заняття – доцільно використовувати

для виконання динамічного експрес-аналізу якості визначеного обсягу тестового матеріалу, що аналізується, із урахуванням індивідуальних особливостей середовища, у якому виконуються виміри.

Функціональна схема методології системного експрес-аналізу якості тестового матеріалу представлена на рис. 1. Алгоритм реалізації представленої функціональної схеми наступний:

У блоці 1 задається сигнал сукупності навчального матеріалу, що відзначений для розглядання на наступному навчальному занятті $X=\{X_1, X_2, \dots, X_n\}$, де n – кількість відібраних навчальних структурних елементів, що відібрані для розглядання на навчальному занятті, що планується.

У блоці 2 за сигналом $X=\{X_1, X_2, \dots, X_n\}$ із зазначеним викладачем випередженням задається сигнал структурованого навчального матеріалу $NM=\{NM_1, NM_2, \dots, NM_n\}$, із яким особа, що навчається, до визначеного часу (частіше усього – до наступного навчального заняття) повинна ознайомитися та самостійно засвоїти.

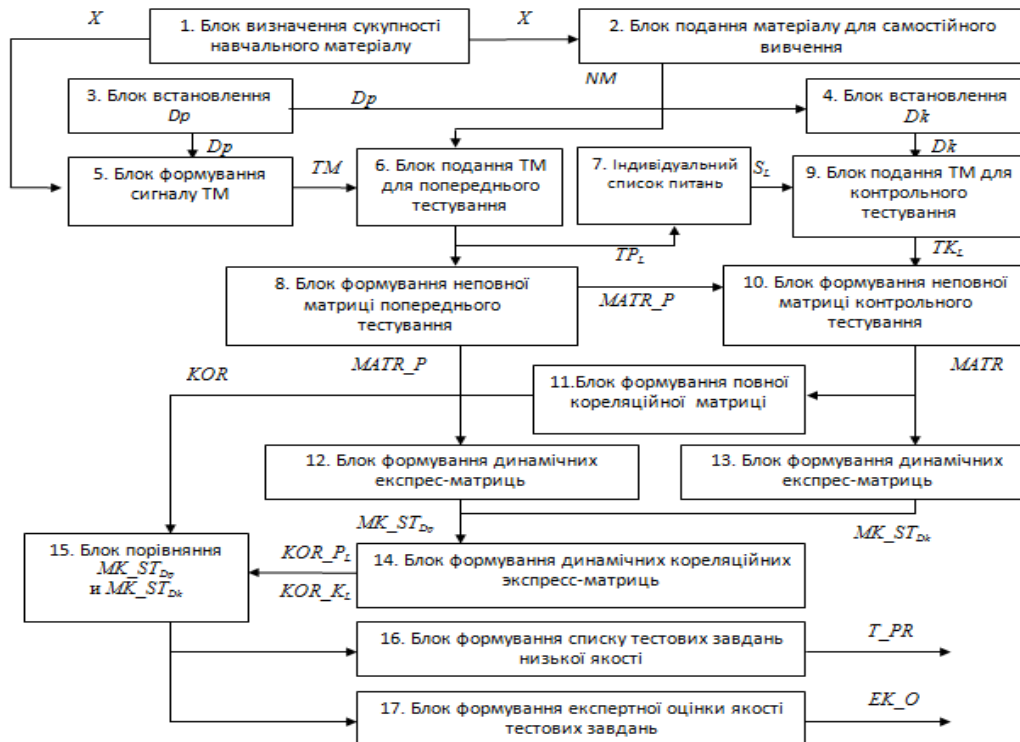


Рисунок 1 – Функціональна схема методології системного експрес-аналізу якості тестового матеріалу

У блоці 3 задається значення сигналу D_p відносної кількості (долі) тестових завдань із повної сукупності тестового матеріалу, що призначені для проведення попереднього тестування.

У блоці 4 формується значення сигналу D_k відносної кількості (долі) тестових завдань із повної сукупності тестового матеріалу, що призначені для проведення попереднього тестування:

$$D_k = 1 - D_p \quad (3)$$

У блоці 5 за сигналом $X=\{X_1, X_2, \dots, X_n\}$ формується сигнал сукупності тестового матеріалу $TM=\{TM_{ij}, \dots, TM_{mn}\}$, що відібраний для розглядання на запланованому навчальному занятті ($i=1, M_j$), де i – індекс тестового завдання за j -м структурним елементом навчального матеріалу, M_j – кількість тестових завдань у j -му структурному елементі, $K=\sum M_j$ – загальна кількість тестових завдань.

На початку навчального заняття у блоці 6 виконується подання сигналу тестових завдань $TP_L=\{TP_{ii}D_p\}$ для попереднього тестування у кількості KOL_p . Порядок подання – випадковий пропорційно з кожного j -го структурного елементу навчального матеріалу:

$$KOL_p = \sum D_p * M_j \quad (4)$$

За результатами попереднього тестування у блоці 7 формується сигнал проміжної неповної експрес-матриці результатів попереднього тестування $MATR_P$.

У таблиці 1 наведено приклад результату гіпотетичного сеансу навчання та виміру якості тестового матеріалу.

Вихідні дані, що використовувались:

кількість відібраних для розгляду на навчальному занятті, що планується, структурних елементів $n=3$;

кількість тестових завдань структурних елементах: $M_1=4$; $M_2=6$; $M_3=4$;

загальна кількість тестових завдань $K=14$;

відносна кількість (доля) тестових завдань із повної сукупності тестового матеріалу, що призначені для проведення попереднього тестування $D_p=0,5$, та, як наслідок, відносна кількість (доля) тестових завдань із повної сукупності тестового матеріалу, що призначені для проведення контрольного тестування $D_k=0,5$;

кількість осіб, що навчаються $L=12$;

A_{ks} – результат відповіді k -го користувача на s -е тестове завдання (може приймати значення $A_{ks}=\{0,1\}$).

Оскільки співвідношення кількості тестових завдань, що призначені для проведення попереднього та контрольного тестування $D_p=D_k$, отримана проміжна неповна експрес-матриця результатів попереднього тестування $MATR_P$, що заповнена рівно на половину.

Таблиця 1 – Проміжна неповна експрес-матриця результатів попереднього тестування $MATR_P$

		Тестові завдання													$\sum_{s=1}^{D_p * K} A_{ks}$	
		Питання 1				Питання 2				Питання 3						
		1	2	3	4	5	6	7	8	9	10	11	12	13		14
Студенти	1	1		0	0	1			0	0			1			3
	2		1			0	0		0		1		0		0	2
	3	1		1			1	0		1		0		1		5
	4		0			1	1		1		1		1		1	6
	5	0			1	1		0	1	0		1				4
	6			0			1	1		1		1		1	1	6
	7		0		1		0		1		0	1	1			4
	8	1		0		1		1	1	0			1			5
	9		1			0	0	0			1		0		0	2
	10	0		0						1	1	0		0	0	2
	11		0				1	0	1			1	1		1	5
	12	0			1	1		0	0		1			1		4

1. Додатково за результатами попереднього тестування у блоці 8 формується сигнал індивідуального списку тестових завдань $S_L=\{S_s, \dots, S_{D_p}\}$, одержаних для попереднього тестування, де s – індекс номеру тестового завдання, одержаного L -м користувачем у результаті генерації KOL_p тестових завдань із повної сукупності тестового матеріалу TP (у таблиці 2 наведено приклад результатів формування індивідуального списку тестових завдань у відповідності із даними таблиці 1).

Таблиця 2 – Індивідуальні списки тестових завдань S_L , одержаних для попереднього тестування

	$S_L 1$	$S_L 2$	$S_L 3$	$S_L 4$	$S_L 5$	$S_L 6$	$S_L 7$	$S_L 8$	$S_L 9$	$S_L 10$	$S_L 11$	$S_L 12$
Питання	1	2	1	2	1	3	2	1	2	1	2	1
	3	5	3	5	4	6	4	3	5	3	6	4
	4	6	6	6	5	7	6	5	6	8	7	5
	5	8	7	8	7	9	8	7	7	10	8	7
	8	10	9	10	8	11	10	8	10	11	11	8
	9	12	11	12	9	13	11	9	12	13	12	10
	12	14	13	14	11	14	12	12	14	14	14	13

2. Наприкінці заняття (після додаткового розглядання та обговорення на поточному занятті визначеної сукупності навчального матеріалу $X=\{X_1, X_2, \dots, X_n\}$) у блоці 9 виконується подання у випадковому порядку сигналу тестових завдань $TK_L=\{TP_{idk}\}$ для попереднього тестування у кількості:

$$KOL_k = D_k * K \tag{5}$$

При чому, обов’язковими умовами, що виконуються, є:

а) загальна сумарна кількість тестових завдань, отриманих кожною особою, що проходить етапи попереднього та контрольного тестування, дорівнює загальній кількості елементів тестового матеріалу $TM=\{TM_{ij}, \dots, TM_{mn}\}$, що відібраний для розглядання на запланованому навчальному занятті:

$$KOL_k + KOL_p = K \tag{6}$$

б) кожна особа, що проходить етапи попереднього та контрольного тестування, у процесі проходження цих етапів отримує у випадковому порядку усі питання тестового матеріалу $TM = \{TM_{ij}, \dots, TM_{mn}\}$ без повторів:

$$TK_L \cup TP_L = TM \quad (7)$$

$$TK_L \cap TP_L = \emptyset \quad (8)$$

3. За результатами контрольного тестування у блоці 10 формуються сигнали проміжної неповної експрес-матриці результатів контрольного тестування $MATR_K$ та, як результат виконання операції об'єднання множин $MATR_P \cup MATR_K$, повної експрес-матриці результатів тестування $MATR$ (приклад наведено у таблицях 3 та 4), отриманої за одно навчальне заняття.

Таблиця 3 – Проміжна неповна експрес-матриця результатів контрольного тестування $MATR_P$

		Тестові завдання												$\sum_{s=1}^{D_k * K} A_{ks}$			
		Питання 1				Питання 2				Питання 3							
		1	2	3	4	5	6	7	8	9	10	11	12		13	14	
Студенти	1		0					1	1			1	1		1	0	5
	2	0		1	1			1		1		0		1			5
	3		1		0	1			1		0		0		0		3
	4	1		1	1			0		1		1		0			5
	5		1	0				1			1		1	1	1		6
	6	0	1		1	1			1		0		1				5
	7	1		1		1			1		0				1	1	6
	8		0		1		0				1	1		1	1		5
	9	0		0	0				0	0		0		1			1
	10		0		0	0	1	0	0				1				2
	11	1		1	1	1					1	1			1		7
	12		1	0				1			0		0	0		0	2

Таблиця 4 – Повна експрес-матриця результатів тестування $MATR$

		Тестові завдання												$\sum_{s=1}^{14} A_{ks}$		
		Питання 1				Питання 2				Питання 3						
		1	2	3	4	5	6	7	8	9	10	11	12		13	14
Студенти	1	1	0	0	0	1	1	1	0	0	1	1	1	1	0	8
	2	0	1	1	1	0	0	1	0	1	1	0	0	0	0	6
	3	1	1	1	0	1	1	0	1	1	0	0	0	1	0	8
	4	1	0	1	1	1	1	0	1	1	1	1	1	0	1	11
	5	0	1	0	1	1	1	0	1	0	1	1	1	1	1	10
	6	0	1	0	1	1	1	1	1	1	0	1	1	1	1	11
	7	1	0	1	1	1	0	1	1	0	0	1	1	1	1	10
	8	1	0	0	1	1	0	1	1	0	1	1	1	0	1	9
	9	0	1	0	0	0	0	0	0	0	1	0	0	1	0	3
	10	0	0	0	0	0	1	0	0	1	1	0	1	1	0	5
	11	1	0	1	1	1	1	0	1	1	1	1	1	1	1	12
	12	0	1	0	1	1	1	0	0	0	1	0	0	1	0	6

2. На основі сигналу матриці повної експрес-матриці результатів тестування $MATR$ у блоці 11 формується сигнал повної кореляційної експрес-матриці результатів тестування KOR (приклад наведено у таблиці 5).

У наведеній таблиці стовбець REZ_1 є інформаційним елементом повної кореляційної експрес-матриці результатів тестування KOR та утримує данні про рівень якості тестового завдання:

$$\checkmark - \text{достатній } (R_{xy} \geq 0,5) \quad (9)$$

$$! - \text{середній } (0,3 < R_{xy} < 0,5) \quad (10)$$

$$\times - \text{недостатній } (R_{xy} \leq 0,29) \quad (11)$$

Таким чином, завдяки використанню тестового контролю знань у якості допоміжного технологічного інструменту заохочення студентів щодо навчання, а також контролю та моніторингу зростання рівня знань та оцінки зрозумілості навчального матеріалу, одночасно, без додаткових витрат часу на проведення експертизи якості тестового матеріалу, надається можливість отримати повну

кореляційну експрес-матрицю результатів тестування, аналіз якої дозволяє оперативно виявити проблемні з точки зору якості тестові завдання вже на перших етапах використання нового тестового матеріалу.

Таблиця 5 – Повна кореляційна експрес-матриця результатів тестування *KOR*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	R_{xy}	<i>REZ_1</i>
1	1,00	-0,67	0,51	0,00	0,58	0,00	0,17	0,51	0,00	-0,19	0,51	0,35	-0,19	0,33	0,53	✓
2	-0,67	1,00	-0,17	0,00	-0,19	0,00	-0,17	-0,17	0,00	-0,19	-0,51	-0,71	0,19	-0,33	-0,35	×
3	0,51	-0,17	1,00	0,24	0,10	-0,12	-0,03	0,37	0,51	-0,29	0,03	-0,12	-0,29	0,17	0,37	!
4	0,00	0,00	0,24	1,00	0,41	-0,13	0,24	0,48	0,00	0,00	0,48	0,25	-0,41	0,71	0,60	✓
5	0,58	-0,19	0,10	0,41	1,00	0,41	0,10	0,68	-0,19	-0,33	0,68	0,41	0,11	0,58	0,78	✓
6	0,00	0,00	-0,12	-0,13	0,41	1,00	-0,48	0,12	0,35	0,00	0,12	0,25	0,41	0,00	0,33	!
7	0,17	-0,17	-0,03	0,24	0,10	-0,48	1,00	0,03	-0,17	-0,29	0,37	0,24	-0,29	0,17	0,18	×
8	0,51	-0,17	0,37	0,48	0,68	0,12	0,03	1,00	0,17	-0,49	0,66	0,48	-0,10	0,85	0,85	✓
9	0,00	0,00	0,51	0,00	-0,19	0,35	-0,17	0,17	1,00	-0,19	-0,17	0,00	-0,19	0,00	0,22	×
10	-0,19	-0,19	-0,29	0,00	-0,33	0,00	-0,29	-0,49	-0,19	1,00	-0,10	0,00	-0,33	-0,19	-0,31	×
11	0,51	-0,51	0,03	0,48	0,68	0,12	0,37	0,66	-0,17	-0,10	1,00	0,84	-0,10	0,85	0,85	✓
12	0,35	-0,71	-0,12	0,25	0,41	0,25	0,24	0,48	0,00	0,00	0,84	1,00	0,00	0,71	0,67	✓
13	-0,19	0,19	-0,29	-0,41	0,11	0,41	-0,29	-0,10	-0,19	-0,33	-0,10	0,00	1,00	-0,19	-0,09	×
14	0,33	-0,33	0,17	0,71	0,58	0,00	0,17	0,85	0,00	-0,19	0,85	0,71	-0,19	1,00	0,85	✓
															R_{xy}	0,39

Однак, оскільки повна кореляційна експрес-матриця отримана у результаті злиття двох неповних матриць результатів тестування тих самих осіб, що навчаються, але у різні моменти часу та з урахуванням різного рівня знань (до та після навчального заняття), з метою урахування особливостей отримання інформаційного масиву, що використовується, а також підвищення об'єктивності та адекватності проведення експрес-експертизи якості тестового матеріалу автором запропоновано додаткові експертні технології:

На базі неповних експрес-матриць результатів попереднього MATR_P та контрольного тестування MATR_K формується сигнал динамічних експрес-матриць-стовпців результатів попереднього МК_STDp (блок 12) та контрольного тестування МК_STDk (блок 13) за кожним тестовим завданням.

На базі сигналів динамічних експрес-матриць-стовпців результатів попереднього МК_STDp та контрольного тестування МК_STDk за кожним тестовим завданням у блоці 14 формується сигнал динамічних кореляційних експрес-матриць-стовпців результатів попереднього KOR_PL та контрольного тестування KOR_KL (приклад наведено у таблиці 6):

Таблиця 6 – Динамічні кореляційні експрес-матриці-стовпці результатів попереднього *KOR_PL* та контрольного *KOR_KL* тестування

№ тестового завдання	<i>KOR PL</i>	<i>REZ 2</i>	<i>KOR KL</i>	<i>REZ 3</i>
1	0,47	!	0,63	✓
2	-0,92	×	0,04	×
3	0,27	×	0,55	✓
4	1,00	✓	0,83	✓
5	0,78	✓	0,81	✓
6	0,89	✓	-0,30	×
7	0,58	✓	0,66	✓
8	0,75	✓	0,95	✓
9	0,12	×	0,62	✓
10	-0,20	×	0,63	✓
11	0,47	!	0,70	✓
12	0,81	✓	0,41	!
13	0,88	✓	0,00	×
14	0,98	✓	0,73	✓
R_{xy}	0,49		0,53	

Стовпці REZ_2 та REZ_2 також є інформаційними елементами динамічних кореляційних експрес-матриць-стовпців результатів попереднього KOR_PL та контрольного KOR_KL тестування та утримують динамічну інформацію про рівень якості тестового завдання (згідно 9, 10, 11), визначену у різні моменти часу та із урахуванням різних рівнів знань (до та після навчального заняття)

У блоці 15 виконується порівняння сигналів повної кореляційної експрес-матриці KOR та динамічних кореляційних експрес-матриць-стовпців результатів попереднього KOR_PL та контрольного тестування KOR_KL.

За результатами порівняння формується сигнали:

блок 16 – сигнал списку із Z тестових завдань $T_PR=\{T_PRz\}$, що викликають сумнів із точки зору їх якості (на основі сигналу стовпця REZ_1 експрес-матриця результатів тестування KOR із значеннями – середній $0,3 < R_{xy} < 0,5$ (!) та недостатній $R_{xy} \leq 0,29$ (*));

блок 17 – сигнал експертної оцінки тестових завдань, що викликають сумнів із точки зору їх якості $EK_O=\{EK_Oz\}$.

Прикладом формування списку експертних оцінок тестових завдань, що викликають сумнів із точки зору їх якості $EK_O=\{EK_Oz\}$, де EK_Oz складається з кортежу $EK_Oz = \langle EK_ONz; EK_OVz \rangle$ (EK_ONz – група, до якої відноситься тестове завдання; EK_OVz – експертний висновок), може бути наступний алгоритм (таблиця 7):

Таблиця 7 – Алгоритм експертної оцінки тестових завдань, що викликають сумнів із точки зору їх якості

EK_ON_z	Умови визначення	EK_OV_z	
		Тестове завдання	Структурний елемент навчального матеріалу
<i>Немає росту знань</i>			
1	$REZ_1_z < 0,29$ $REZ_2_z < 0,29$ $REZ_3_z < 0,29$	– занадто складне питання; – некоректно сформульоване;	недостатньо якісно роз'яснено на навчальному занятті
<i>Спостерігається зниження знань</i>			
2	$REZ_1_z > 0,5$ $REZ_2_z < 0,29$ $REZ_2_z < 0,29$ $REZ_2_z / REZ_1_z > 0,5$	– некоректно (не вірогідно) сформульовані дистрактори; – можливий факт вгадування;	недостатньо якісно роз'яснено на навчальному занятті
<i>Спостерігається різке підвищення знань</i>			
3	$REZ_1_z < 0,29$ $REZ_2_z > 0,5$ $REZ_2_z < 0,5$ $REZ_1_z / REZ_2_z > 0,5$	у питанні або дистракторах проглядається відповідь;	не рекомендується видавати матеріал за цією темою на самостійне розглядання; досить легке питання (рекомендується використовувати тільки для проміжного контролю)

Приклад результатів експертної оцінки якості тестових завдань наведено у таблиці 8.

Таблиця 8 – Приклад результатів експертної оцінки якості тестових завдань

№ тестового завдання	$EK_ON_z=1$	$EK_ON_z=2$	$EK_ON_z=3$
2	$REZ_1_z = -0,92$ $REZ_2_z = 0,00$ $REZ_3_z = -0,35$	-	-
3	-	-	$REZ_1_z = 0,27$ $REZ_2_z = 0,66$ $REZ_3_z = 0,37$
6	-	$REZ_1_z = 0,89$ $REZ_2_z = -0,30$ $REZ_3_z = 0,33$	-
9	-	-	$REZ_1_z = 0,12$ $REZ_2_z = 0,62$ $REZ_3_z = 0,22$
10	-	-	$REZ_1_z = -0,20$ $REZ_2_z = 0,68$ $REZ_3_z = -0,31$
13	-	$REZ_1_z = 0,88$ $REZ_2_z = 0,00$ $REZ_3_z = -0,09$	-

Запропонована методологія реалізована у програмній оболонці "Мережне тестування", яка застосовується на кафедрі економічної кібернетики і математичних методів в економіці Дніпропетровського університету економіки та права ім. А. Нобеля.

Висновки

Таким чином, запропонована методологія системного експрес-аналізу якості тестового матеріалу поетапно та комплексно дозволяє:

1. Оптимізувати процедуру проведення експертизи якості тестового матеріалу завдяки розробці комплексу навчально-методологічних засобів, що поєднують безпосередньо початковий процес та процес експрес-аналізу та виміру якості тестового матеріалу.

2. Знизити трудомісткість процесу формування повної кореляційної матриці результатів тестування за означеним навчальним матеріалом шляхом формалізації методології її поетапного гарантованого (із вірогідністю 1) отримання із двох неповних матриць – результатів попереднього та контрольного тестування у межах одного навчального заняття.

3. Підвищити ефективність інструментів оцінки якості тестового матеріалу та, як наслідок, об'єктивності процесу виміру рівня засвоєння навчального матеріалу через обґрунтування доцільності застосування методики експертної оцінки якості тестового матеріалу за додатковими динамічними показниками контролю змінювання знань осіб, що навчаються, у реальному масштабі часу.

4. Гарантувати зростання якості тестового матеріалу, що використовується для контролю знань осіб, що навчаються, з метою адекватної та об'єктивної оцінки рівня їх навчальних досягнень та прийняття ефективних управлінських рішень щодо організації та управління навчальним процесом у цілому.

Література

1. Лучинин А.С. Психодіагностика. Конспект лекцій. □ Феникс, 2004 □ 256 с.
2. Аванесов В.С. Основы научной организации педагогического контроля в высшей школе / В.С. Аванесов // Педагогическая диагностика. – 2002. – №1 С.41- 43.
3. Белоус Н., Куцевич И., Белоус И. Методика определения качества тестовых заданий, оцениваемых по непрерывной шкале / Н. Белоус, И. Куцевич, И. Белоус // International Book Series "Information Science and Computing". The paper is selected from XVth International Conference "Knowledge-Dialogue-Solution". – Kyiv, 2009. – С. 127-133.
4. Мамонтова М.Ю. Квалиметрический подход к моделированию оценки качества академической подготовки студентов / М.Ю. Мамонтова // Успехи современного естествознания – 2007. – № 12 – С. 145.
5. Федорук П.І. Використання адаптивних тестів в інтелектуальних системах контролю знань / П.І. Федорук // Штучний інтелект. – 2008. – № 3. – С. 380–387.
6. Спосіб проведення комп'ютерного тестування знань студентів. [Текст]: патент на корисну модель 58657 Україна: МПК G06F 7/00; Замовник та патентовласник: Тараненко Ю.К., Ризун Н.О. □ № u 2010 09376, заявл. 26.07.2010, опубл. 26.04.2011, Бюл. № 8, 2011 р. – 14 с.
7. Спосіб виміру рівня знань учнів при комп'ютерному тестуванні [Текст]: патент на корисну модель № 51559 Україна: МПК G06F 7/00; Замовник та патентовласник: Тараненко Ю.К., Ризун Н.О. № u 200913726, заявл. 28.12.2009, опубл. 26.07.2010, Бюл. № 14, 2010 р.
8. Максимова О.А. Технология комплексной экспертизы качества тестовых материалов для контроля учебных достижений обучающихся / О.А. Максимова // Журнал научных публикаций аспирантов и докторантов. – 2008. – № 10 – С. 140-146.
9. Автоматизована система тестування, навчання та моніторингу [Текст]: патент на корисну модель № 43616 (13)U. Україна: МПК G09B 7/00 . Ціделко В. Д., Яремчук Н. А., Шведова В. В.; Заявник та патентовласник: Ціделко В. Д., Яремчук Н. А., Шведова В. В. – № u200902620, заявл. 23.03.2009, опубл. 25.08.2009. Бюл. 16, 2009 р.
10. Напрасник С.В., Таранов В.Б., Шкиль А.С. Технологии подготовки тестовых заданий и проведения тестирования в системе OpenTEST. <http://opentest.com.ua/tehnologii-podgotovki-testovykh-zadaniy-i-provedeniya-testirovaniya-v-sisteme-opentest>].
11. Ким В.С. Тестирование учебных достижений. Монография. □ Уссурийск: Издательство УГПИ, 2007. – 214 с.

Відомості про авторів

Ризун Н.О. – доцент кафедри економічної кібернетики та математичних методів в економіці Дніпропетровського університету економіки та права. Україна, м. Дніпропетровськ, вул. Наб. Леніна, 18. Контактний тел.: (056) 713-43-99, моб. 0505-363-16-61, e-mail: n_fedo@mail.ru

УДК 004.89+336.713

Т. О. САВЧУК, С. І. ПЕТРИШИН

Вінницький національний технічний університет, Вінниця

ОЦІНЮВАННЯ РЕЗУЛЬТАТІВ МОДЕЛЮВАННЯ ПРОЦЕСУ КЛАСТЕРНОГО АНАЛІЗУ НАДЗВИЧАЙНИХ СИТУАЦІЙ НА ЗАЛІЗНИЧНОМУ ТРАНСПОРТІ

Анотація. У статті проведено оцінювання результатів моделювання процесу кластерного аналізу надзвичайних ситуацій на залізничному транспорті за допомогою програмного продукту, що базується на модифікованому алгоритмі k-means та програмного засобу DEDUCTOR.

Ключові слова: моделювання, Data Mining, надзвичайна ситуація, залізничний транспорт, кластеризація, кластер.

Аннотация. В статье проведено оценивание результатов моделирования процесса кластерного анализа чрезвычайных ситуаций на железнодорожном транспорте с помощью программного продукта, основанном на модифицированном алгоритме k-means и программного средства DEDUCTOR.

Ключевые слова: моделирование, Data Mining, чрезвычайная ситуация, железнодорожный транспорт, кластеризация, кластер.

Abstract. The paper carried the results of evaluation of process modeling of cluster analysis of emergencies on the railways with the help of a software product based on a modified k-means algorithm and software tools DEDUCTOR.

Key words: Modeling, Data Mining, emergencies, rail transport, clustering, cluster.

Вступ

Характеристики оцінювання результатів моделювання кластерного аналізу надзвичайних ситуацій як складного процесу [1, 2] визначаються як властивостями його складових, так і характером взаємодії між ними. Слід врахувати такі особливості процесу оцінювання:

- його стан описується потужним вектором динамічних змінних;
- виявляє якісні зміни поведінки його складових;
- включає нелінійні взаємодії основних його складових і обернені зв'язки між ними, які також містять нелінійності.

Серед основних задач, що є актуальними в процесі оцінювання результатів кластеризації надзвичайних ситуацій на залізничному транспорті, є розробка його моделі шляхом відтворення зв'язків і відношень між основними його складовими [2, 3].

Для дослідження процесу оцінювання результатів кластерного аналізу надзвичайних ситуацій на залізничному транспорті з використанням фізичного моделювання, як одного з найбільш поширених на практиці підходів [2], в якості фізичної моделі може виступати:

- процес фізичної природи, що описується аналогічним математичним апаратом;
- процес аналогічної фізичної природи, але в другій області параметрів (масштабна модель).

Оскільки підбір процесу фізичної природи аналогічного процесу оцінювання результатів кластерного аналізу і відповідного математичного апарату, є складним, можна зробити висновок про недоцільність застосування даного виду моделювання для розв'язання поставленої задачі [2].

Іншим підходом, також достатньо популярним в практиці дослідження, є математичне моделювання, яке дозволяє оцінювати результати кластерного аналізу надзвичайних ситуацій на залізничному транспорті, враховуючи його якісні та кількісні характеристики, що є доцільним для проведення оцінювання результатів функціонування програмних засобів, що базуються на відповідних алгоритмах.

Отже, оцінювання результатів кластеризації за допомогою засобів, що базуються на класичному та модифікованому алгоритмах k-means для аналізу надзвичайних ситуацій на залізничному транспорті, доцільно проводити за допомогою математичного моделювання.

Постановка задачі

Провести оцінювання результатів моделювання процесу кластерного аналізу надзвичайних ситуацій за допомогою програмного засобу DEDUCTOR для аналізу таких ситуацій, який базується на класичному алгоритмі k-means, та розробленого програмного засобу, що базується на модифікованому алгоритмі k-means [4] (в подальшому – Program analysis of emergency situations (PAES)) з метою визначення якості кластеризації при аналізі надзвичайних ситуацій на залізничному транспорті.

Формалізація процесу оцінювання результатів кластеризації надзвичайних ситуацій на залізничному транспорті

Оцінювання результатів кластеризації надзвичайних ситуацій на залізничному транспорті можна здійснити шляхом використання критеріальних величин.

Існує потужна множина критеріїв, що можуть бути використані для аналізу надзвичайних ситуацій на залізничному транспорті, серед них було обрано частку загального розкиду, точково-бісеріальний метод кореляції та узагальнену дисперсію в класах, яких достатньо для оцінки якості результатів розбиття [5].

Нехай множина надзвичайних ситуацій на залізничному транспорті розбита на k кластерів G_i ($i = \overline{1, k}$).

Для визначення частки загального розкиду надзвичайних ситуацій між кластерами T необхідно ввести такі три характеристики ступеню розсіювання надзвичайних ситуацій [1, 5, 6]:

- загальне розсіювання S , що визначається як

$$S = \sum_{i=1}^n a^2(Y_i, \bar{Y}), \quad (1)$$

де Y_i - вектор даних про i -ту надзвичайну ситуацію;

$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$ - загальний центр ваги;

n - кількість надзвичайних ситуацій, що аналізуються;

$a^2(Y_i, \bar{Y})$ - квадрат відстані між i -ю надзвичайною ситуацією та загальним центром ваги;

- міжкластерний розкид B

$$B = \sum_{z=1}^k n_z a^2(\bar{Y}_z, \bar{Y}), \quad (2)$$

де $\bar{Y}_z = \frac{1}{n_z} \sum_{Y_i \in G_z} Y_i$ - центр ваги z -го кластера надзвичайних ситуацій на залізничному транспорті;

n_z - кількість надзвичайних ситуацій на залізничному транспорті в кластері G_z ;

$a^2(\bar{Y}_z, \bar{Y})$ - квадрат відстані між центром ваги z -го кластера та загальним центром ваги;

- розкид всередині кластерів надзвичайних ситуацій на залізничному транспорті

$$W = \sum_{z=1}^k W_z, \quad (3)$$

де $W_z = \sum_{Y_i \in G_z} a^2(Y_i, \bar{Y})$.

Оскільки кластерний аналіз надзвичайних ситуацій базується на модифікованому алгоритмі k -means [4], то

$$S = W + B. \quad (4)$$

Тоді, частка загального розкиду надзвичайних ситуацій на залізничному транспорті T може бути визначена як [4, 5]:

$$T = 1 - \frac{W}{S} \quad (5)$$

Частка загального розкиду надзвичайних ситуацій T є нормованою величиною ($0 \leq T \leq 1$). При цьому, наближення значення до 0 свідчить про нижчу якість розбиття надзвичайних ситуацій на залізничному транспорті на таксони.

Точково-бісеріальний коефіцієнт кореляції R_b між надзвичайними ситуаціями, що аналізуються, визначається таким чином. Кожній парі надзвичайних ситуацій на залізничному транспорті Y_i та Y_j ставиться у відповідність дві величини – відстань між ними та індекс еквівалентності δ_{ij} [4, 5]

$$\delta_{ij} = \begin{cases} 1, \text{ якщо } Y_i \text{ та } Y_j \text{ належать одному кластеру;} \\ 0, \text{ в протилежному випадку.} \end{cases} \quad (6)$$

Коефіцієнт R_b підраховується як коефіцієнт кореляції між a_{ij} та бінарною величиною δ_{ij} по всіх парах надзвичайних ситуацій, які аналізуються, що дає [5, 6]

$$R_b = \frac{(\bar{a}_b - \bar{a}_w) \sqrt{\frac{f_w f_b}{n_a^2}}}{s_a}, \quad (7)$$

де \bar{a}_b – середня відстань між надзвичайними ситуаціями із різних кластерів;

\bar{a}_w – середня відстань між надзвичайними ситуаціями із одного кластера;

f_w – кількість відстаней між надзвичайними ситуаціями, що потрапили в один кластер;

f_b – кількість відстаней між надзвичайними ситуаціями із різних кластерів;

n_a – загальна кількість відстаней;

s_a – стандартне відхилення відстаней.

Узагальнена дисперсія в класах надзвичайних ситуацій на залізничному транспорті H є однією з характеристик ступеню розсіювання надзвичайних ситуацій на залізничному транспорті, що належать одному класу, і обраховується за формулою [5, 6]:

$$H = \det\left(\sum_{l=1}^k n_l W_l\right), \quad (8)$$

де $\det\left(\sum_{l=1}^k n_l W_l\right)$ – визначник матриці.

Елементи $w_{qm}(l)$ вибіркової коваріаційної матриці W_l можуть бути визначені як

$$w_{qp}(l) = \frac{1}{n_l} \sum_{Y_i \in G_l} (y_i^{(q)} - \bar{y}^{(q)}(l))(y_i^{(p)} - \bar{y}^{(p)}(l)), \quad q, p = 1, 2, \dots, m, \quad (9)$$

де $y_i^{(p)}$ – p -та характеристика надзвичайної ситуації на залізничному транспорті Y_i ;

$\bar{y}^{(p)}(l)$ – середнє значення p -ї компоненти, підраховане за надзвичайними ситуаціями l -го класу.

Відносний показник якості розбиття множини надзвичайних ситуацій на таксономії може бути визначений як

$$K = \frac{\bar{T}' + \bar{R}'_b + \bar{H}'}{3} \quad (10)$$

де T' , R'_b і H' – відносні значення показників T , R_b і H .

Таким чином, процес оцінювання результатів кластеризації надзвичайних ситуацій на залізничному транспорті буде включати такі етапи (рисунок 1):

- 1) моделювання процесу кластерного аналізу за допомогою за допомогою програмного засобу DEDUCTOR та за допомогою PAES;
- 2) обрахування частки загального розкиду T надзвичайних ситуацій між кластерами;
- 3) обрахування точково-бісеріального коефіцієнта кореляції R_b між надзвичайними ситуаціями, що аналізуються;
- 4) обрахування узагальненої дисперсії в класах надзвичайних ситуацій на залізничному транспорті H ;
- 5) обрахування відносного показника якості розбиття множини надзвичайних ситуацій на таксономії.

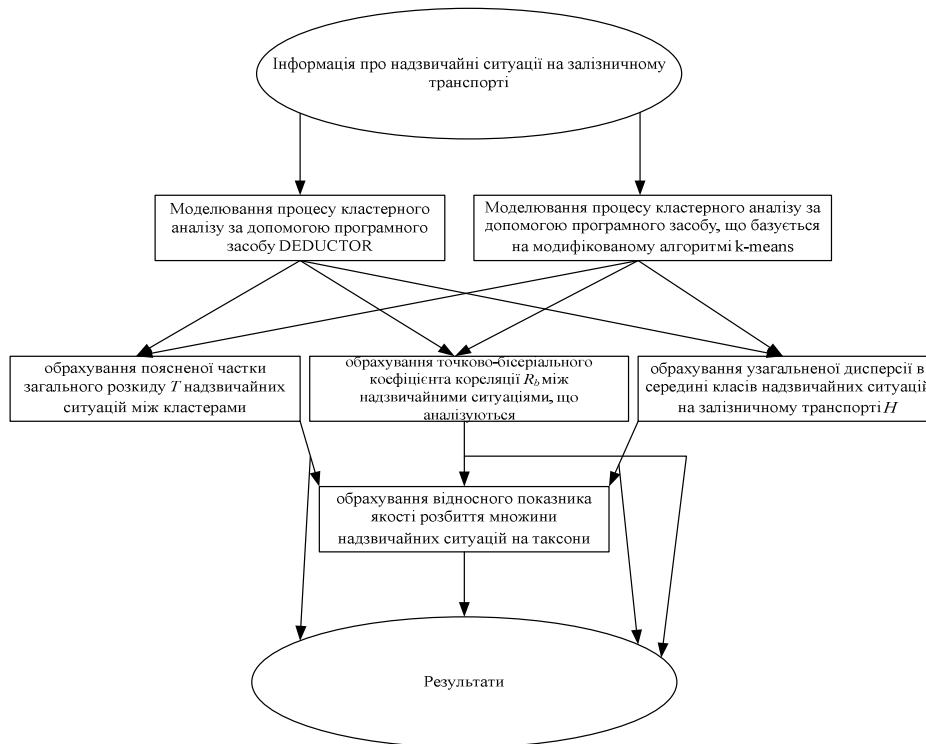


Рисунок 1 - Етапи процесу оцінювання результатів кластеризації надзвичайних ситуацій на залізничному транспорті

Оцінювання результатів процесу кластеризації надзвичайних ситуацій на залізничному транспорті

При використанні програмного засобу DEDUCTOR для аналізу надзвичайних ситуацій та PAES, було проведено 3 експерименти:

- в першому – було обрано 5 значущих параметрів та характеристик надзвичайних ситуацій, що в повній мірі відображають стан таких ситуацій;
- в другому – окрім цих п’яти було додано ще одну – менш значущу;
- в третьому – окрім значущих було додано ще 5 не важливих характеристик та параметрів при ідентифікації надзвичайних ситуацій.

Частку загального розкиду T (таблиця 1) визначений за результатами моделювання процесів кластерного аналізу надзвичайних ситуацій показав, що якість розбиття, яке виконане PAES є вищою, оскільки $T_{PAES} > T_{Deductor}$ (рисунок 2).

Точково-бісеріальний коефіцієнт кореляції R_b (таблиця 1) визначений за результатами моделювання процесів кластерного аналізу надзвичайних ситуацій показав, що якість розбиття, яке виконане програмним засобом DEDUCTOR для аналізу надзвичайних ситуацій є вищою, оскільки $R_{bDeductor} > R_{bPAES}$, проте різниця між отриманими значеннями є незначною та за певних умов можна вважати, що, $R_{bDeductor} \geq R_{bPAES}$, або $R_{bDeductor} = R_{bPAES}$, а отже, можна зробити висновок про те, що якість розбиття в обох випадках (за даним критерієм) є однаковою (рисунок 2).

Таблиця 1 – Оцінювання результатів кластерного аналізу надзвичайних ситуацій на залізничному транспорті

Засіб	Алгоритм кластеризації	Кількість кластерів			Частка загального розкиду, T			Точково-бісеріальний коефіцієнт кореляції, R_b			Узагальнена дисперсія в класах, H			Відносний показник якості розбиття
		1	2	3	1	2	3	1	2	3	1	2	3	
Програмний засіб DEDUCTOR для аналізу надзвичайних ситуацій	k-means	3	3	4	0,8005	0,8018	0,8399	0,586	0,516	0,469	0,721	0,744	0,782	75,4%
PAES	Модифікований k-means	3	3	3	0,8199	0,8233	0,8475	0,563	0,512	0,482	0,672	0,689	0,715	78,31%

Узагальнена дисперсія в класах H (таблиця 1) визначена за результатами моделювання процесів кластерного аналізу надзвичайних ситуацій показала, що якість розбиття, яке виконане програмним засобом DEDUCTOR для аналізу надзвичайних ситуацій є нижчою, оскільки $H_{\text{Deductor}} > H_{\text{PAES}}$, що свідчить про більші відхилення від центра кластерів надзвичайних ситуацій при розбитті програмним засобом DEDUCTOR для аналізу надзвичайних ситуацій (рисунок 2).

Відносний показник якості розбиття множини надзвичайних ситуацій на таксони (таблиця 1) визначений за результатами моделювання процесів кластерного аналізу надзвичайних ситуацій показав, що якість розбиття, яке виконане програмним засобом DEDUCTOR для аналізу надзвичайних ситуацій є нижчою на 2,91%, що свідчить про меншу якість розбиття множини надзвичайних ситуацій на кластери програмним засобом DEDUCTOR для аналізу надзвичайних ситуацій (рисунок 2).

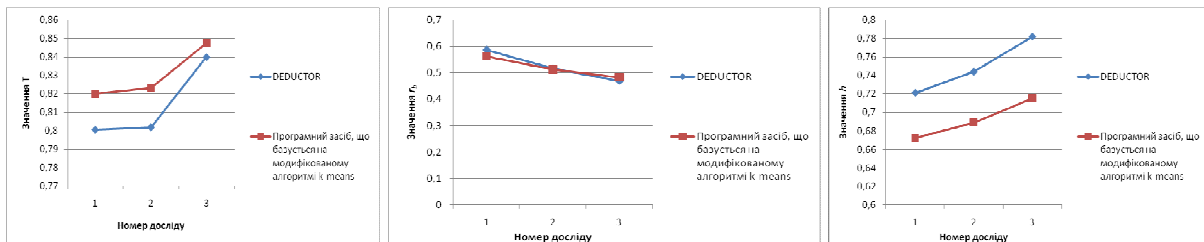


Рисунок 2 - Порівняння результатів моделювання процесу кластерного аналізу надзвичайних ситуацій на залізничному транспорті

Проведене дослідження показало, що удосконалення класичного алгоритму k-means шляхом введення нових параметрів до цільової функції привело до підвищення якості кластеризації надзвичайних ситуацій на залізничному транспорті на 2,9%, а отже, модифікований алгоритм доцільно застосовувати для аналізу надзвичайних ситуацій на залізничному транспорті.

Висновки

Таким чином, за результатами моделювання процесу оцінювання результатів кластерного аналізу надзвичайних ситуацій на залізничному транспорті можна зробити висновок, що модифікований алгоритм k-means має такі переваги по відношенню до класичного алгоритму k-means, що покладений в основу програмного засобу DEDUCTOR, як

- 1) якість розбиття за часткою загального розкиду T є вищою;
- 2) якість розбиття за узагальненою дисперсією в класах H є вищою;
- 3) у випадку точково-бісеріального коефіцієнта кореляції R_b – якість розбиття практично однакова;
- 4) за відносним показником якості розбиття – модифікований алгоритм k-means має перевагу 2,9% при кластеризації надзвичайних ситуацій на залізничному транспорті в порівнянні із звичайним алгоритмом k-means.

Список літератури

1. Савчук Т.О., Петришин С.І. Визначення евклідової відстані між надзвичайними ситуаціями на залізничному транспорті під час кластерного аналізу//Наукові праці Вінницького національного технічного університету. – Серія «Інформаційні технології та комп'ютерна техніка». – 2010. – Випуск №3(2010). - http://www.nbu.gov.ua/e-journals/vntu/2010_3/2010

2. В.М. Томашевский Моделирование систем – К. Видавнича група ВНУ, 2005 – 352 с.
3. Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. Методы и модели анализа данных: OLAP и Data Mining. – СПб.: БХВ-Петербург, 2004 – 336с.
4. Савчук Т.О., Петришин С.І. Розробка модифікованого алгоритму K-MEANS для аналізу надзвичайних ситуацій на залізничному транспорті// Матеріали конференції, Обчислювальний інтелект (результати, проблеми, перспективи): Матеріали 1-ї Міжнародної науково-технічної конференції (10-13 травня 2011 р.), - Черкаси, 2011, - С. 236-237.
5. Айвазян С.А., Бухштабер В.М., Енюков И.С. Прикладная статистика: Классификация и снижение размерности. – М.:Финансы и статистика, 1989. – 607 с.
6. Мандель И.Д. Кластерный анализ. – М.:Финансы и статистика, 1988. – 176с.

Відомості про авторів

Савчук Тамара Олександрівна – кандидат технічних наук, доцент, професор кафедри Комп’ютерних наук, Вінницький національний технічний університет, Хмельницьке шосе, 95, м. Вінниця, 21021, тел.0664124037, savchtam@rambler.ru

Петришин Сергій Іванович – аспірант кафедри Комп’ютерних наук Вінницького національного технічного університету, Хмельницьке шосе, 95, м. Вінниця, 21021.

КОМП'ЮТЕРНІ СИСТЕМИ ТА КОМПОНЕНТИ

УДК 621.375.024

О. Д. АЗАРОВ, С. В. БОГОМОЛОВ

Вінницький національний технічний університет, Вінниця

ПОХИБКИ ЛІНІЙНОСТІ ПРЕЦИЗІЙНИХ ПЕРЕТВОРЮВАЧІВ СТРУМ-НАПРУГА ТА НАПРУГА-НАПРУГА НА БАЗІ ДВОТАКТНИХ ПІДСИЛЮВАЧІВ СТРУМУ

Анотація. Розглянуто похибки лінійності прецизійних перетворювачів струм-напруга та напруга-напруга на базі двотактних підсилювачів струму. Виведено аналітичні співвідношення, що описують похибки статичної передатної характеристик із урахуванням параметрів схеми, зокрема, вхідного і вихідного опорів, опору навантаження та коефіцієнтів передачі.

Ключові слова: прецизійність, перетворювач струм-напруга, перетворювач напруга-напруга, похибка лінійності, структурно-функціональна організація, двотактний підсилювач струму, автокоригування.

Аннотация. Рассмотрены погрешности линейности прецизионных преобразователей ток-напряжение и напряжение-напряжение на базе двухтактных усилителей тока. Выведены аналитические соотношения, описывающие погрешности статической передаточной характеристики с учетом параметров схемы, в частности, входного и выходного сопротивлений, сопротивления нагрузки и коэффициентов передачи.

Ключевые слова: прецизионность, преобразователь ток-напряжение, преобразователь напряжение-напряжение, погрешность линейности, структурно-функциональная организация, двухтактный усилитель тока, автокорректировка.

Abstract. The error of linear precision converters current-voltage and voltage-voltage push-pull amplifiers based on the current. Analytical value describing the error of static transfer characteristics of the subject circuit parameters, including input and output resistance, load resistance and transfer factors.

Key words: precision, current-voltage converter, voltage-voltage converter, linearity error, structural and functional organization, push-pull current amplifier, automatic adjustment.

Вступ

Перетворювачі струм-напруга (ПСН) та напруга-напруга (ПНН) є аналоговими вузлами, що використовуються у багатьох електронних пристроях, зокрема, багаторозрядних АЦП і ЦАП, які у свою чергу входять до складу високоточних систем вимірювання, опрацювання та реєстрування сигналів [1-3]. Деякі метрологічні характеристики цих перетворювачів, такі, як лінійність передатної характеристики, слабо піддаються коригуванню, тому прецизійні ПСН та ПНН, здебільшого, будують на базі операційних підсилювачів з великим коефіцієнтом передачі (10^6 - 10^8) [1,4]. Проте, такий підхід, як правило, обмежує їх швидкодію.

Кроком вперед у цьому напрямку є використання, двотактних підсилювачів струму (ДПС), які мають високу лінійність і в яких не вимагається досягнення надвисоких значень коефіцієнтів передачі.

Актуальність

Традиційно побудову ПСН та ПНН здійснюють шляхом застосування операційного підсилювача, який увімкнено по схемі із глибоким зворотнім зв'язком [3, 5-7]. Проте застосування традиційних однотактних підсилювачів постійного струму має свої недоліки, а саме: вузьку смугу повної неспотвореної потужності та обмежену швидкість наростання вихідного сигналу. Це все призводить до різкого збільшення нелінійності із зростанням частоти вхідного сигналу [8]. Кращі показники у цьому плані мають ДПС, які випускаються серійно провідними фірмами, зокрема Analog Device, Maxim, ON Semiconductor, Intersil [1,4]. Проте і вони мають свої недоліки: низьке підсилення на каскад і, як результат, для досягнення потрібних коефіцієнтів підсилення, потрібне збільшення кількості підсилювальних каскадів [6], що у свою чергу, погіршує динамічні характеристики.

Авторами запропоновано ряд методів структурно-функціональної організації прецизійних ПСН та ПНН на базі ДПС [9]. При цьому доцільно оцінити переваги та недоліки застосування ДПС із низьким, середнім та високим вхідними опорами. Відомі аналітичні співвідношення, що описують передатні характеристики ПСН та ПНН із урахуванням більшої кількості параметрів підсилювача, таких, як: параметри зовнішніх і внутрішніх кіл та опору навантаження, дозволяють більш точно оцінити передатні характеристики. Проте здійснено оцінку впливу лише вхідного опору ДПС на похибку лінійності ПСН та ПНН [9].

Водночас матеріал, присвячений аналізу похибок лінійності прецизійних ПСН та ПНН на базі ДПС в науково-технічній літературі подається епізодично і є неструктурованим. Тому тема статті, присвячена похибкам лінійності прецизійних перетворювачів стум-напруга та напруга-напруга на базі двотактних підсилювачів струму є актуальною.

Мета

Аналіз похибок лінійності прецизійних перетворювачів стум-напруга та напруга-напруга на базі двотактних підсилювачів струму із урахуванням параметрів елементів зовнішніх і внутрішніх кіл підсилювача та опору навантаження.

Задачі

1. Отримати аналітичні співвідношення похибок лінійності статичних передатних характеристик перетворювачів струм-напруга та напруга-напруга.
2. Дослідити вплив параметрів елементів зовнішніх і внутрішніх кіл підсилювача та опору навантаження на рівень похибки лінійності.
3. Проаналізувати запропонований метод структурно-функціональної організації прецизійного двотактного підсилювача струму із коригуванням зсуву нуля.

Розв’язання задач

Для дослідження впливу параметрів зовнішніх і внутрішніх кіл підсилювача та опору навантаження на похибку лінійності прецизійних ПСН та ПНН на базі ДПС, доцільно розглянути їх схеми заміщення, які зображено на рис. 1. Операційний режим роботи підсилювача здійснено шляхом уведення кола глибокого зворотного зв’язку за допомогою резистора R_M .

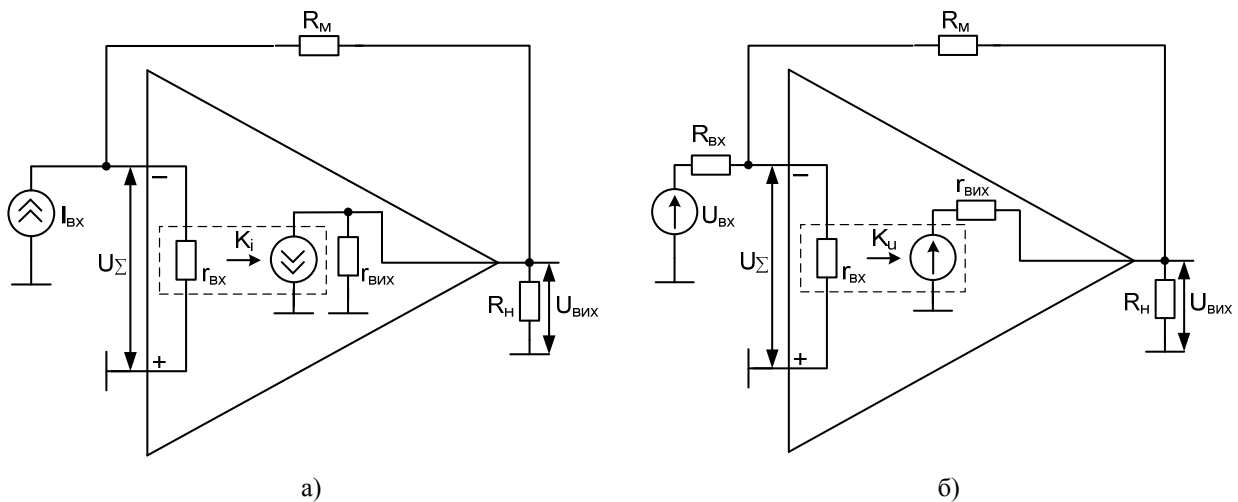


Рисунок 1 – Схеми заміщення перетворювачів а) струм-напруга; б) напруга-напруга

Залежно від типу перетворювача вхідний сигнал може бути заданим, як у вигляді струму I_{vx} , так і у вигляді напруги U_{vx} , а вихідним сигналом є напруга U_{vix} . У літературних джерелах [1-3, 7, 10], описуються ПСН та ПНН, передані характеристики яких мають вигляд:

$$\text{ПСН: } U_{vix} \approx I_{vx} \cdot R_M; \quad \text{ПНН: } U_{vix} \approx -U_{vx} \cdot \frac{R_M}{R_{vx}}. \quad (1)$$

Проте у цих формулах не враховуються характеристики самого підсилювача, а саме: вхідний r_{vx} та вихідний r_{vix} опори, опір навантаження R_n , які впливають на рівень похибок лінійності ПСН та ПНН.

Залежно від чинників, що впливають на передатну характеристику, можна виділити такі рівні деталізації, а саме врахування:

$$R_M, r_{vx}, K_i \text{ (або } K_u \text{)};$$

$$R_M, R_n, r_{vx}, K_i \text{ (або } K_u \text{)};$$

$$R_M, R_n, r_{vx}, r_{vix}, K_i \text{ (або } K_u \text{)},$$

де K_i, K_u – коефіцієнти передачі по струму і по напрузі – K_u , що визначаються при розірваній петлі зворотного зв’язку [9].

При першому рівні деталізації передатна характеристика ПСН має вигляд [9]:

$$U_{вих} = I_{вх} \cdot \frac{r_{вх} + R_M \cdot K_i}{1 + K_i}.$$

Абсолютна похибка [11] визначається, як:

$$\Delta U_{вих} = U_{вих} - U'_{вих}, \quad (2)$$

де $U'_{вих} = I_{вх} \cdot \frac{R_M \cdot K_i}{1 + K_i}$ – напруга на виході ПСН, при $r_{вх} = 0$.

Враховуючи вищевказані вирази та вираз (2), отримуємо:

$$\Delta U_{вих} = I_{вх} \cdot \frac{r_{вх}}{1 + K_i} \quad (3)$$

Відносна похибка [11] визначається, як:

$$\delta = \frac{\Delta U_{вих}}{U_{вих}} \cdot 100\%. \quad (4)$$

Враховуючи вищевказані вирази, отримуємо вираз для опису відносної похибки ПСН при ненульовому значенні $r_{вх}$.

$$\delta = \frac{r_{вх}}{R_M \cdot (1 + K_i)} \cdot 100\%. \quad (5)$$

При другому рівні деталізації передатна характеристика ПСН має вигляд [9]:

$$U_{вих} = I_{вх} \cdot \frac{r_{вх} + R_M \cdot K_i}{1 + K_i + \frac{R_M - r_{вх}}{R_H}}.$$

Із урахуванням рівностей (2) і (4), отримуємо співвідношення для опису абсолютної та відносної похибок ПСН:

$$\Delta U_{вих} = I_{вх} \cdot \frac{r_{вх} \cdot R_H \cdot (1 + K_i) - R_M \cdot K_i \cdot (R_M - r_{вх})}{(1 + K_i) \cdot [R_M - r_{вх} + R_H \cdot (1 + K_i)]}, \quad (6)$$

$$\delta = \frac{r_{вх} \cdot R_H \cdot (1 + K_i) - R_M \cdot K_i \cdot (R_M - r_{вх})}{R_M \cdot (1 + K_i) \cdot [R_M - r_{вх} + R_H \cdot (1 + K_i)]} \cdot 100\%. \quad (7)$$

При третьому рівні деталізації передатна характеристика ПСН має вигляд [9]:

$$U_{вих} = I_{вх} \cdot \frac{r_{вх} + R_M \cdot K_i}{\left(1 + \frac{R_M}{R_{сер}}\right) \cdot (1 + K_i) - \frac{r_{вх} + R_M \cdot K_i}{R_{сер}}},$$

причому $R_{сер} = \frac{r_{вих} \cdot R_H}{r_{вих} + R_H}$.

Із урахуванням рівностей (2) і (4), отримуємо співвідношення для опису абсолютної та відносної похибок ПСН:

$$\Delta U_{вих} = I_{вх} \cdot \frac{r_{вх} \cdot R_{сер} \cdot (1 + K_i) - R_M \cdot K_i \cdot (R_M \cdot (1 + 2 \cdot K_i) - r_{вх})}{(1 + K_i) \cdot [R_M \cdot (1 + 2 \cdot K_i) - r_{вх} + R_{сер} \cdot (1 + K_i)]}; \quad (8)$$

$$\delta = \frac{r_{вх} \cdot R_{сер} \cdot (1 + K_i) - R_M \cdot K_i \cdot (R_M \cdot (1 + 2 \cdot K_i) - r_{вх})}{R_M \cdot (1 + K_i) \cdot [R_M \cdot (1 + 2 \cdot K_i) - r_{вх} + R_{сер} \cdot (1 + K_i)]} \cdot 100\%. \quad (9)$$

Беручи до уваги аналітичні вирази (6-9), які описують абсолютну та відносну похибки ПСН із вищими рівнями деталізації та враховуючи параметри підсилювача ($r_{вх} \ll R_M$, $r_{вих} \ll R_H$, $K_i \gg 1$), шляхом математичних спрощень, отримуємо вирази (3) і (5) із нижчим рівнем деталізації для ПСН. Вирази (3) і (5) враховують критичні параметри ПСН. Найкритичнішим чинником спотворення похибок передатної характеристики є наявність U_Σ , що викликано впливом $r_{вх}$ [5].

Передатна характеристика ПНН при першому рівні деталізації має вигляд [9]:

$$U_{вих} = U_{вх} \cdot \frac{r_{вх} \cdot R_M \cdot K_U}{R_M \cdot (r_{вх} + R_{вх}) - (K_U - 1) \cdot r_{вх} \cdot R_{вх}}. \quad (10)$$

Із урахуванням рівностей (2) і (4), отримуємо співвідношення для опису абсолютної та відносної похибок ПСН:

$$\Delta U_{вих} = U_{вх} \cdot \frac{R_M \cdot [r_{вх} \cdot R_{вх} + R_M \cdot (r_{вх} + R_{вх})]}{R_{вх} \cdot [R_M \cdot (r_{вх} + R_{вх}) - (K_U - 1) \cdot r_{вх} \cdot R_{вх}]}; \quad (11)$$

$$\delta = - \frac{r_{вх} \cdot R_{вх} + R_M \cdot (r_{вх} + R_{вх})}{R_M \cdot (r_{вх} + R_{вх}) - (K_U - 1) \cdot r_{вх} \cdot R_{вх}} \cdot 100\%. \quad (12)$$

При другому рівні деталізації отримаємо збіжність передатної характеристики ПНН із виразом (10), адже $U_{вих} = U_{вих.вн}$ і сигнал повністю передається в навантаження. Проте на практиці підсилювач має $r_{вих} \neq 0$, враховуючи який, отримаємо передатну характеристику ПНН [9] для третього рівня деталізації:

$$U_{вих} = U_{вх} \cdot \frac{r_{вх} \cdot R_M \cdot R_H \cdot (K_U \cdot R_M + r_{вих})}{\left[r_{вих} \cdot R_H + R_M \cdot (r_{вих} + R_H) \right] \cdot \left[r_{вх} \cdot R_{вх} + R_M \cdot (r_{вх} + R_{вх}) \right] - \frac{-r_{вх} \cdot R_{вх} \cdot R_H \cdot (K_U \cdot R_M + r_{вих})}{}}$$

Із урахуванням рівностей (2) і (4), отримуємо співвідношення для опису абсолютної та відносної похибок ПНН:

$$\Delta U_{вих} = U_{вх} \cdot \frac{R_M \cdot \left[r_{вх} \cdot R_{вх} \cdot R_H \cdot (K_U \cdot R_M + r_{вих}) + \left[r_{вих} \cdot R_H + R_M \cdot (r_{вих} + R_H) \right] \cdot \left[r_{вх} \cdot R_{вх} + R_M \cdot (r_{вх} + R_{вх}) \right] - \frac{-r_{вх} \cdot R_{вх} \cdot R_H \cdot (K_U \cdot R_M + r_{вих})}{}}{R_{вх} \cdot \left[r_{вих} \cdot R_H + R_M \cdot (r_{вих} + R_H) \right] \cdot \left[r_{вх} \cdot R_{вх} + R_M \cdot (r_{вх} + R_{вх}) \right] - \frac{-r_{вх} \cdot R_{вх} \cdot R_H \cdot (K_U \cdot R_M + r_{вих})}{}}; \quad (13)$$

$$\delta = - \frac{r_{вх} \cdot R_{вх} \cdot R_H \cdot (K_U \cdot R_M + r_{вих}) + \left[r_{вих} \cdot R_H + R_M \cdot (r_{вих} + R_H) \right] \cdot \left[r_{вх} \cdot R_{вх} + R_M \cdot (r_{вх} + R_{вх}) \right] - \frac{-r_{вх} \cdot R_{вх} \cdot R_H \cdot (K_U \cdot R_M + r_{вих})}{}}{\left[r_{вих} \cdot R_H + R_M \cdot (r_{вих} + R_H) \right] \cdot \left[r_{вх} \cdot R_{вх} + R_M \cdot (r_{вх} + R_{вх}) \right] - \frac{-r_{вх} \cdot R_{вх} \cdot R_H \cdot (K_U \cdot R_M + r_{вих})}{}} \cdot 100\% \quad (14)$$

Враховуючи, що на практиці параметри підсилювача співвідносяться, як: $r_{вх} \gg R_M$, $r_{вх} \gg R_{вх}$, $R_M \approx R_{вх}$, $r_{вих} \ll R_H$, $K_U \gg 1$ та беручи до уваги аналітичні вирази (11-14), шляхом математичних спрощень, отримуємо вирази (8) і (9) із нижчим рівнем деталізації для ПНН. Вирази (8) і (9) враховують найкритичніші параметри ПНН на основі ДПС. Істотний внесок на появу похибок передатної характеристики є наявність U_{Σ} [5]. Тому для побудови прецизійних ПСН доцільно використовувати ДПС із низьким вхідним опором $r_{вх}$. Традиційно для зменшення U_{Σ} використовують підсилювачі із високим коефіцієнтом підсилення по струму K_i . Сучасні підсилювачі, які випускаються провідними фірмами світу, забезпечують коефіцієнти підсилення на рівні 120...140 дБ [12]. Підвищення коефіцієнту підсилення по струму K_i ДПС можливе, шляхом введення одно-, дво-, трохступневих проміжних каскадів (ПК) та двотактних вихідних каскадів (ДВК) [13]. При цьому похибка лінійності оберненопропорційна коефіцієнту підсилення по струму K_i .

Зменшити похибку лінійності можна додатково збільшивши глибину зворотного зв'язку R_M . Проте такий підхід призводить водночас до зменшення швидкодії [10].

Для побудови ПНН доцільно використовувати ДПС на основі комплементарних польових транзисторів [9]. Такі ДПС хоч і мають високий вхідний опір, проте можуть бути використані при роботі із вхідними сигналами у вигляді напруги та джерелами сигналів ємнісного характеру, а також забезпечують підсилення.

Специфікою ДПС на біполярних транзисторах є наявність ненульового вхідного струму, що призводить до появи похибки зсуву нуля $\Delta U_{зс0}$, причому $\Delta U_{зс0} = \Delta I_{зс0} \cdot R_M$, при $I_{вх} = 0$.

Для зменшення похибки зсуву нуля авторами запропоновано метод структурно-функціональної організації ДПС із автокоригуванням зсуву нуля, схему якого наведено на рис. 2.

Вона складається з блока напруг зміщення (БНЗ), джерел струму I1 та I2, які з транзисторами VT1 і VT2 задають режим роботи каскадів схеми.

Блок автокоригування нуля (БАН), який складається з імітатора вхідного каскаду (ІВК), перетворювачів струмів (ПС1, ПС2), відбивачів струму (ВС1, ВС2) і транзисторів VT2-VT12, забезпечує автокори-

гування струму $I_{зс0}$ параметричним методом таким чином, що $\Delta I_{зс0} \rightarrow 0$. Це, у свою чергу, дозволяє зменшити похибку зсуву нуля на 1-2 порядки.

Двонаправлений відбивач струму (ДВС) разом із ВС3-ВС6, забезпечують автоматичне симетрування підсилення по верхньому і нижньому каналах підсилення, які побудовано на вхідному (ВК), ПК1, ПК2 і ДВК.

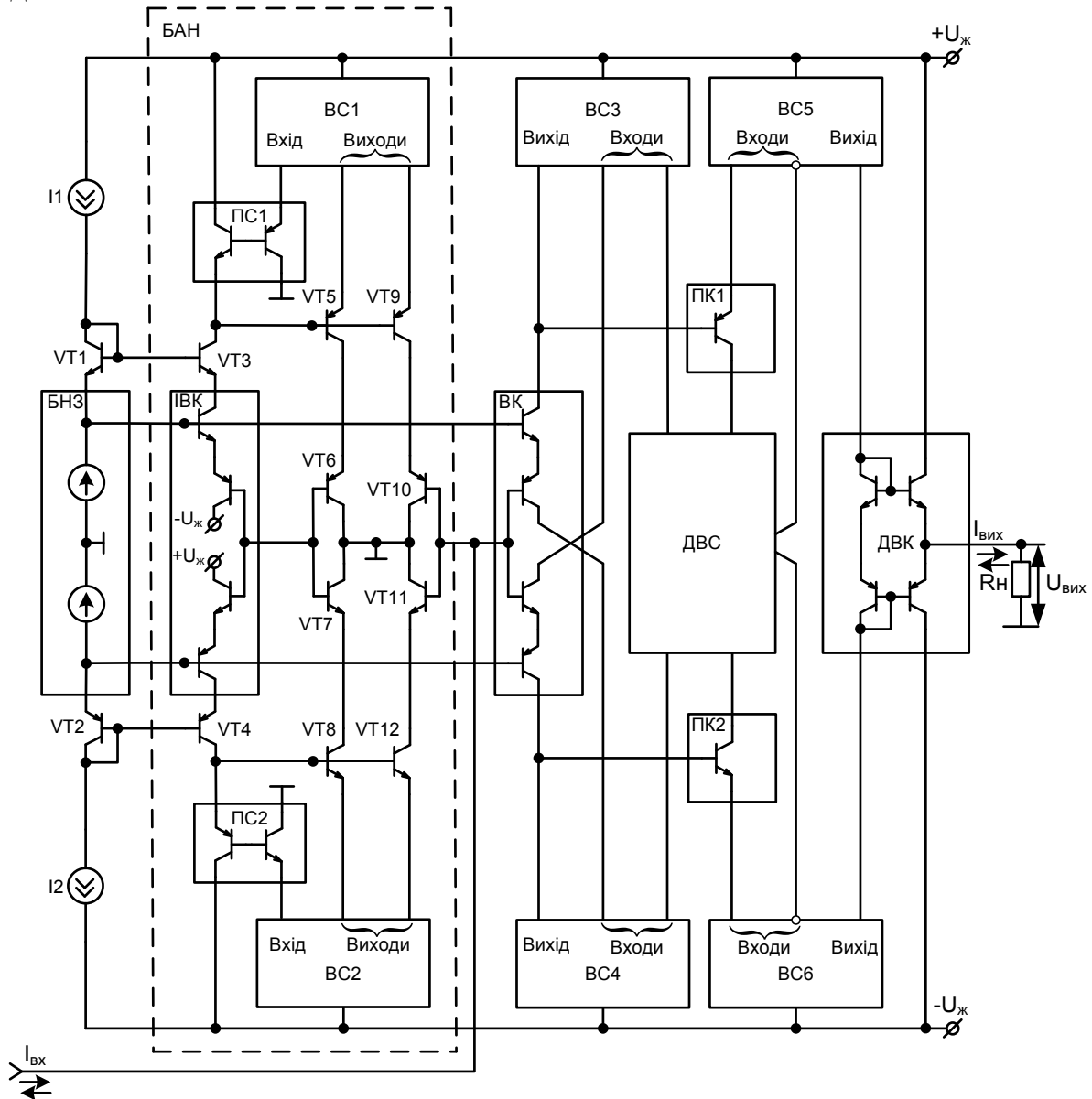


Рисунок 2 – Структурно-функціональна організація ДПС із автокоригуванням зсуву нуля

ДПС, який побудовано за розглянутою структурно-функціональною організацією, забезпечує такі характеристики:

- коефіцієнт підсилення K_i : 125,87 дБ;
- вхідний опір: 3380 Ом;
- частота одиничного підсилення: 796,2 МГц;
- струм зсуву нуля: 89,9 нА;
- діапазон вихідного струму: ± 1 мА;
- діапазон вихідної напруги: ± 10 В;
- похибка лінійності $\delta_L = 0.000016\%$.

Висновки

1. Виведено аналітичні співвідношення, що дозволяють описати похибки лінійності передатних характеристик ПСН та ПНН із урахуванням параметрів внутрішніх елементів підсилювача, зовнішніх кіл та опору навантаження. Це дозволяє точніше оцінити значення похибок передатних характеристик.
2. Досліджено вплив параметрів зовнішніх і внутрішніх кіл підсилювача та опору навантаження на появу похибок лінійності прецизійних ПСН та ПНН. Доведено, що лінійність ДПС можна покращити, використовуючи вхідні каскади із низьким вхідним опором.
3. Проаналізовано запропонований метод структурно-функціональної організації прецизійного ДПС із автокоригуванням зсуву нуля. Показано, що це дозволяє зменшити похибку зсуву нуля на 1-2 порядки.

Список літератури

1. Walt Kesler. ANALOG-DIGITAL CONVERSION / Walt Kesler – ADI Central Application Department, March 2004. – 1127 p.
2. Alan B. Grebene. Bipolar and MOS analog integrated circuit design / Alan B. Grebene – New Jersey: Wiley Classic Library, 2002. – 915 p.
3. Волович Г.И. Схемотехника аналоговых и аналого-цифровых электрон-ных устройств / Волович Г.И. – М.: Издательский дом «Додэка-XXI», 2005. – 528 с.
4. Walter G. Jong, Op Amp applications handbook / Walter G. Jong – Analog Devices series, 2004, – 895 p.
5. Титце У. Полупроводниковая схемотехника: [справочное руководство] / У. Титце, К. Шенк; [пер. с нем.] –М.: 1982. – 512 с., ил.
6. Фолкенберри Л. Применение операционных усилителей и лнейных ИС / Фолкенберри Л. [пер. с англ.] –М.: Мир, 1985. – 572 с., ил.
7. Пейтон А. Дж. Аналоговая электроника на операционных усилителях / А. Дж.Пейтон, В. Волш – М.: БИНОМ, 1994 – 352 с.: ил. – ISBN 5-7503-0013-7
8. Азаров О.Д. Основи теорії аналого-цифрового перетворення на основі надлишкових позиційних систем числення: [монографія] / Азаров О.Д. – Вінниця: УНІВЕРСУМ – Вінниця, 2004. – 257 с.
9. Азаров О.Д. Перетворювачі струм-напруга та напруга-напруга на базі двотактних підсилювачів струму / О.Д. Азаров, С.В. Богомолів // Інформаційні технології та комп'ютерна інженерія. Вінницький національний технічний університет – 2011. – №2(21). – С. 4-11.
10. Уин Палмер. Быстродействующий прецизионный усилитель-преобразователь сопротевлений / Уин Палмер // Электроника. Серия: методы, схемы, аппаратура – 1988. – №1. – С. 77-82.
11. Касаткин А.С. Электротехника: [учеб. пособие для вузов] /А.С. Касаткин, М.В. Немцов– изд. 4-е, перераб. – М.:Энергоатомиздат, 1983. – 440 с., ил.
12. Макаренко В. Операционные усилители со сверхмалыми искажениями для высококачественных аудиоприложений / В. Макаренко // Электронные компоненты и системы – 2011. – №12(172). – С.34-38.
13. Азаров О.Д. Двотактні підсилювачі постійного струму із симетричною структурою / О.Д. Азаров, В.А. Гарнага, С.В. Богомолів // Вісник Хмельницького національного університету – 2008. – №4(113). – С. 20-24.

Відомості про авторів

Азаров Олексій Дмитрович – д. т. н., професор, завідувач кафедри обчислювальної техніки, Вінницький національний технічний університет, Хмельницьке шосе, 95, м. Вінниця, 21021, тел. 58-02-25.

Богомолів Сергій Віталійович – аспірант кафедри обчислювальної техніки, Вінницький національний технічний університет, Хмельницьке шосе, 95, м. Вінниця, 21021, тел. +38-097-131-83-78, e-mail: bogomolovsergiy@rambler.ru.

УДК 621.382.3

О. Д. АЗАРОВ, В. Є. ЯЦИК

Вінницький національний технічний університет, Вінниця

МЕТОДИ ПОКРАЩЕННЯ СТАТИЧНИХ ХАРАКТЕРИСТИК ВІДБИВАЧІВ СТРУМУ

Анотація. У статті розглянуто методи покращення статичних характеристик відбивачів струму за рахунок зменшення їх вхідного опору й істотного збільшення вихідного. Аналізуються запропоновані підходи щодо схемно-функціональної організації вказаних пристроїв на базі параметричного і компенсаційного відбивача Уілсона. Виведено аналітичні співвідношення для оцінювання вхідного і вихідного опорів запропонованих структур. Здійснено порівняльний аналіз статичних характеристик запропонованих і відомих схем. Показано, що запропоновані пристрої істотно перемагають базові за вищезгаданими параметрами.

Ключові слова: параметричний відбивач струму, відбивач струму Уілсона, вхідний опір, вихідний опір.

Аннотация. В статье рассмотрены методы улучшения статических характеристик отражателей тока за счет уменьшения их входного сопротивления и существенного увеличения выходного. Анализируются предложенные подходы к схемно-функциональной организации указанных устройств на базе параметрического и компенсационного отражателя Уилсона. Выведены аналитические соотношения для оценки входного и выходного сопротивлений предложенных структур. Осуществлен сравнительный анализ статических характеристик предложенных и известных схем. Показано, что предложенные устройства существенно превосходят базовые по вышеуказанным параметрам.

Ключевые слова: параметрический отражатель тока, отражатель тока Уилсона, входное сопротивление, выходное сопротивление.

Abstract. The article reviews methods of improving the static characteristics of current mirrors by reducing their input resistance and a substantial increase in output. Analyzed the proposed approaches to functional organization of these devices based on parametric and compensation Wilson current mirror. Been derived the analytic relations for the evaluation of input and output resistances of the proposed structures. Realizable comparative analysis of static characteristics of the proposed and known schemes. It is shown that the proposed device essentially win basic on the above parameters.

Key words: current mirror, Wilson current mirror, input resistance, output resistance.

Вступ

Відбивачі струму (ВС) або струмові дзеркала (current mirrors) [1] є складовою частиною багатьох аналогових і аналогово-цифрових схем [1, 2]. Параметри цих елементів, зокрема, вхідний і вихідний опори можуть істотно впливати як на статичні, так і на динамічні характеристики електронних пристроїв [3], які є складовими частинами систем вимірювання, перетворення й опрацювання сигналів.

Актуальність

Існує певне розмаїття схем ВС, побудованих на біполярних транзисторах (БТ), описаних, зокрема, в [1, 2, 4, 5, 6]. При цьому треба відзначити, що основою більшості схем є або найпростіший параметричний ВС у вигляді композиції: давач струму (діод) - відбивач, або компенсаційний ВС із зворотнім зв'язком типу схеми Уілсона [1, 2]. Перевагою останнього є збільшений вихідний опір і відповідно менша похибка завдання коефіцієнта передачі. Треба також відзначити, що схеми ВС на БТ є економічнішими порівняно зі схемами на польових транзисторах, оскільки вимагають меншої напруги живлення.

Значну кількість існуючих публікацій, присвячених ВС [1, 5, 7, 8, 9], спрямовано на розв'язання задачі підвищення вихідного опору. Водночас, можливості зменшення вхідного опору практично не розглядаються. При цьому слід відзначити, що для ряду пристроїв, наприклад, таких як високолінійні широкосмугові підсилювачі струму [10] вхідний опір ВС є критичним і може істотно впливати на їх як статичні, так і динамічні характеристики. Крім того питання, присвячені комплексному покращенню статичних характеристик ВС, а саме зменшенню вхідного та збільшенню вихідного опорів у рамках єдиної схеми у науково-технічній літературі практично не розглянуто.

У зв'язку з вищевказаним тема статті “Методи покращення статичних характеристик відбивачів струму” є актуальною.

Мета

Покращення статичних характеристик ВС за рахунок зменшення їх вхідного та збільшення вихідного опорів шляхом застосування запропонованих нових схемно-структурних методів.

Задачі

Аналіз запропонованих методів схемно-структурної організації ВС зі зменшеними вхідними і збільшеними вихідними опорами;

1. Виведення аналітичних співвідношень для оцінювання вхідних і вихідних опорів ВС, побудованих на базі запропонованих методів покращення їх статичних характеристик.

Розв'язання задач

Найпростіший параметричний ВС, побудований як композиція діод (транзистор в діодному вмиканні) – відбивач (каскад із загальним емітером (ЗЕ)) має вхідний опір, значення якого визначається на основі співвідношення, наведеного в [11]:

$$\Gamma_{\text{вх}} \approx \frac{\Phi_t}{I_d},$$

де Φ_t – термопотенціал, значення якого дорівнює 25,86 мВ при температурі $t=27\text{ }^\circ\text{C}$;

I_d – струм, що протікає через діод.

Питання зменшення $r_{вх}$ такого ВС практично не розглянуто в науково-технічній літературі, тому проаналізуємо його детальніше. Кардинальним шляхом розв'язання цієї задачі є застосування схем із введенням активного від'ємного зворотного зв'язку, що стабілізує напругу на вході ВС під час змінення вхідного струму.

Можна вказати принаймні два варіанти схемно-функціональної організації ВС параметричного типу з низьким вхідним опором, що ілюструються на рисунку 1. Застосування на входах схем окремо генератора робочого струму I_p і вхідного струму $I_{вх}$ здійснено для зручності виконання комп'ютерного схемотехнічного аналізу.

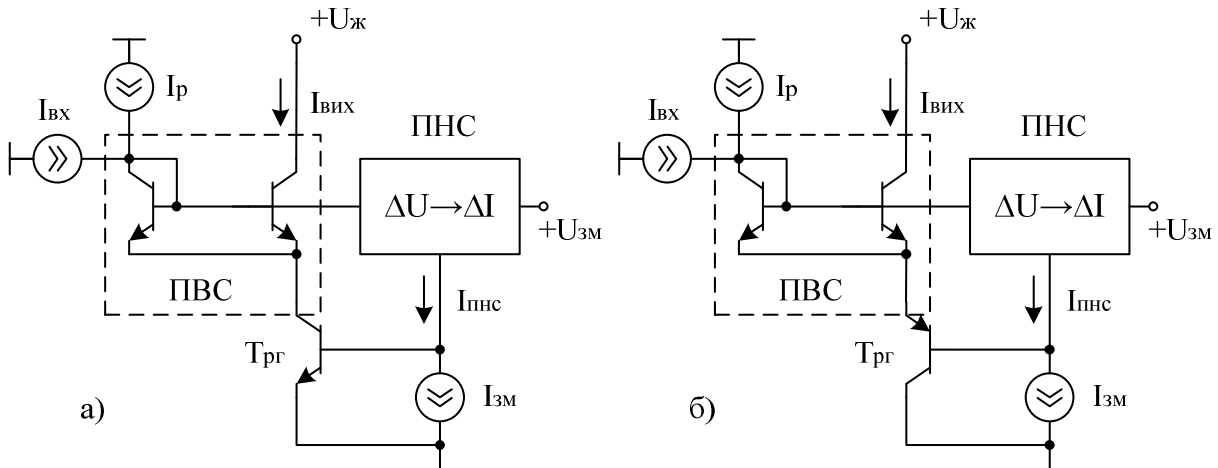


Рисунок 1 – Ілюстрація методів схемно-функціональної організації параметричних ВС із низьким вхідним опором із реалізацією регуляторів на основі каскадів: а) ЗЕ; б) ЗК

Перший метод (рис. 1 а)) передбачає застосування регулятора $T_{рг}$ у вигляді транзисторного каскаду загальний емітер (ЗЕ), а другий (рис. 1 б)) каскаду загальний колектор (ЗК). До контуру зворотного зв'язку уведено перетворювач напруга-струм (ПНС) ($\Delta U \rightarrow \Delta I$). Тут $\Delta U = U_{вх} - U_{зм}$ – різниця напруга, що знімається з діода параметричного відбивача струму (ПВС) і з виходу джерела напруги змінення. При цьому слід зазначити, що для першого випадку треба мати ПНС в якому збільшення ΔU призведе до зростання вихідного струму $I_{пнс}$, а в другому випадку - до зменшення. Обидві схеми працюють таким чином. Якщо вхідний струм $I_{вх}$ збільшується, то збільшується $U_{вх}$. При цьому ПНС формує приріст струму $\Delta I_{пнс}$, що привідкриває транзистор $T_{рг}$, тобто збільшує струм, який протікає через ПВС і стабілізує $U_{вх}$. Робочі точки ПНС і $T_{рг}$ задаються відповідно робочим струмом I_p та струмом змінення $I_{зм}$. На рисунку 2 наведено варіанти схемних реалізацій ПНС.

Оскільки наведені схеми ВС містять контури зворотного зв'язку (ЗЗ) [11] у вигляді композиції ПНС і регуляторів $T_{рг}$, то у загальному випадку малосигнальний вхідний опір пристрою визначається не стільки вхідним опором ПВС, скільки вхідним опором ПНС, а саму у вигляді:

$$r_{вх} = \frac{r_{вх}}{1 + \chi \cdot K_i} \quad (1)$$

де $r_{вх}$ – вхідний опір ПНС при розірваній петлі ЗЗ;

K_i – коефіцієнт підсилення струму в контурі ЗЗ;

χ – коефіцієнт розгалуження струму на виході регулятора. При цьому слід відзначити, що оскільки ПВС має два кола, а саме: діод – давач і каскад ЗЕ, то розгалуження I_p у ці кола має бути приблизно однаковими. Тому можна вважати, що $\chi \approx 0,5$.

Треба відзначити, що коефіцієнти передачі (відбиття) для всіх наведених схем дорівнюють:

$$K_{вб} = \frac{I_{вих}}{I_{вх}} \approx 1.$$

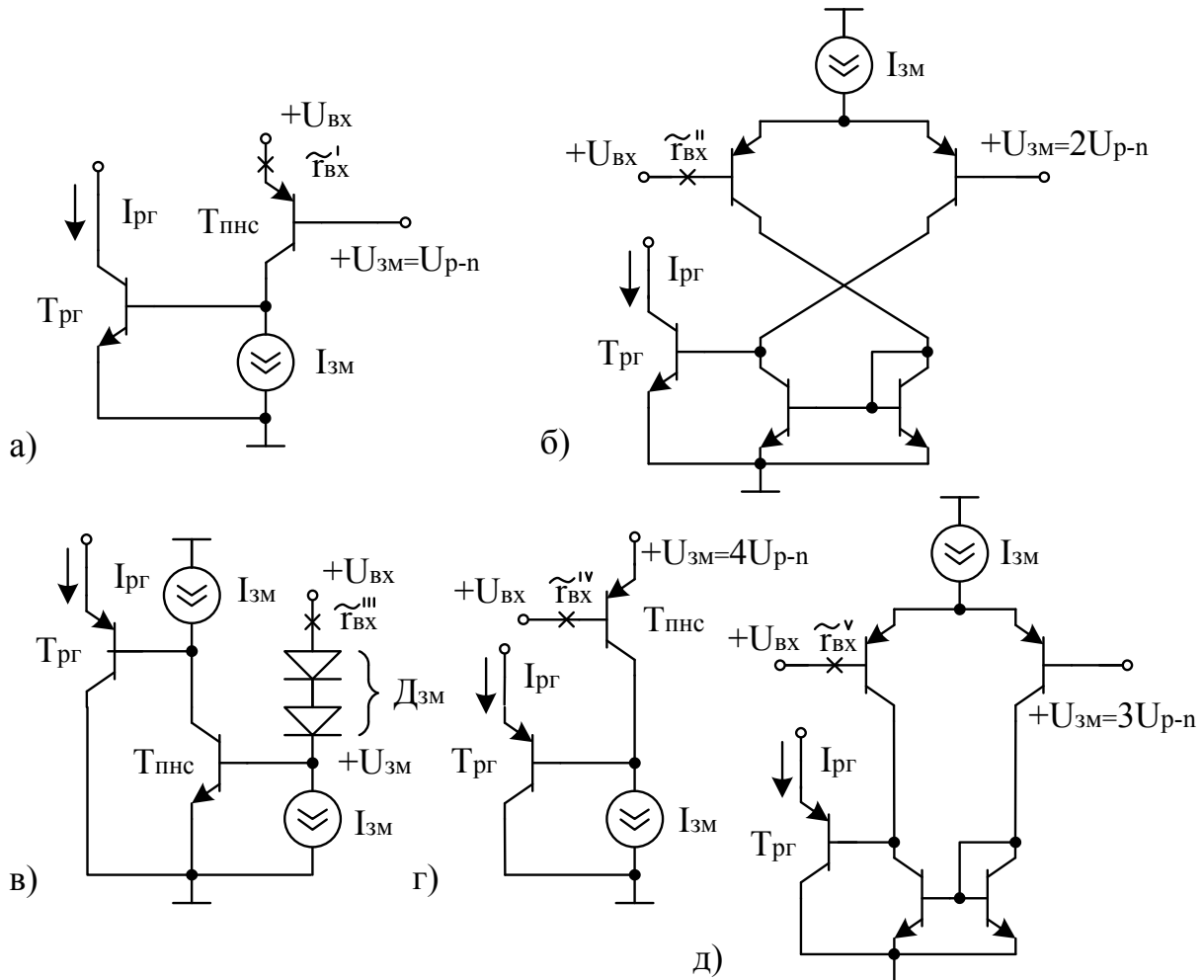


Рисунок 2 – варіанти реалізації ПНС для регуляторів на основі каскаду: а), б) ЗЕ; в), г), д) ЗК.

У випадку використання ПНС, зображеного на рисунку 2 а), маємо:

$$K_i' = \alpha \cdot \beta_{\text{pr}},$$

де α – коефіцієнт передачі струму каскаду ЗБ (ПНС);

β_{pr} – коефіцієнт передачі струму каскаду Т_{pr} (ЗЕ).

Застосовуючи аналітичні вирази, наведені в [6], можна показати що:

$$\overline{r_{\text{BX}}'} = r_{\text{BXЗБ}} = r_e + \frac{r_{\delta}}{1 + \beta},$$

тому:

$$r_{\text{BX}}' = \frac{r_e + \frac{r_{\delta}}{1 + \beta}}{1 + 0,5 \cdot \alpha \cdot \beta_{\text{pr}}} \approx \frac{2 \cdot r_e}{\beta_{\text{pr}}}.$$

Для схеми, зображеної на рисунку 2 б), ПНС являє собою диференційний каскад з активним навантаженням у вигляді ВС, тому:

$$\overline{r_{\text{BX}}''} = 2 \cdot r_e \cdot (1 + \beta_{\text{p-n-p}}).$$

Коефіцієнт передачі струму по контуру ЗЗ дорівнює:

$$K_i'' = 2 \cdot K_{i\text{пкс}} \cdot K_{i\text{pr}} = 2 \cdot \beta_{\text{p-n-p}} \cdot \beta_{\text{n-p-n}}.$$

Підставляючи значення $\overline{r_{\text{BX}}''}$ і K_i'' в (1) маємо:

$$r_{\text{ВХ}}'' = \frac{2 \cdot r_e \cdot (1 + \beta_{\text{p-n-p}})}{\beta_{\text{p-n-p}} \cdot \beta_{\text{n-p-n}}} \approx \frac{2 \cdot r_e}{\beta_{\text{n-p-n}}}$$

У випадку схеми (рис. 2 в)) без урахування опорів діодів зміщення маємо:

$$r_{\text{ВХ}}''' = 2 \cdot r_e \cdot (1 + \beta_{\text{n-p-n}})$$

тому по аналогії з попереднім випадком отримаємо:

$$r_{\text{ВХ}}'''' = \frac{r_e \cdot (1 + \beta_{\text{n-p-n}})}{0,5 \cdot \beta_{\text{p-n-p}} \cdot \beta_{\text{n-p-n}}} \approx \frac{2 \cdot r_e}{\beta_{\text{p-n-p}}}$$

Для схеми (рис. 2 г)) вхідний опір дорівнює:

$$r_{\text{ВХ}}^{\text{IV}} = r_e \cdot (1 + \beta_{\text{p-n-p}}),$$

при цьому:

$$r_{\text{ВХ}}^{\text{IV}} = \frac{r_e \cdot (1 + \beta_{\text{p-n-p}})}{0,5 \cdot \beta_{\text{p-n-p}}^2} \approx \frac{2 \cdot r_e}{\beta_{\text{p-n-p}}}$$

У випадку схеми (рис. 2 д)):

$$r_{\text{ВХ}}^{\text{V}} = 2 \cdot r_e \cdot (1 + \beta_{\text{p-n-p}}),$$

тому:

$$r_{\text{ВХ}}^{\text{V}} = \frac{2 \cdot r_e \cdot (1 + \beta_{\text{p-n-p}})}{\beta_{\text{p-n-p}}^2} \approx \frac{2 \cdot r_e}{\beta_{\text{p-n-p}}}$$

Під час розрахунку робочих точок ВС слід мати на увазі, що рівень I_p для схем, наведених на рисунках 2 а) і 2 в) повинен бути приблизно удвічі більше ніж для схем 2 б), г) і д). Крім того слід брати до уваги, що оскільки транзистори ПНС і T_p працюють при низьких рівнях напруги живлення $1 \div 3 U_{p-n}$, то значення їх малосигнальних коефіцієнтів передачі β будуть дещо зниженими ніж для вихідного транзистора ПВС, який працює в номінальному режимі.

Значення малосигнальних вхідних опорів ПВС для випадків застосування ПНС, зображених на рисунках 2 а) – д), при робочих струмах $\sim 1 \text{mA}$ (для інтегральних транзисторів фірми Intersil) зведено в таблиці 1.

Таблиця 1 – Порівняльна таблиця вхідного опору для різних схем ВС

$r_{\text{ВХ}}$ (Ом)	$r_{\text{ВХ}}'$	$r_{\text{ВХ}}''$	$r_{\text{ВХ}}'''$	$r_{\text{ВХ}}^{\text{IV}}$	$r_{\text{ВХ}}^{\text{V}}$	$r_{\text{ВХ}}$ ПВС	$r_{\text{ВХ}}$ ВСУ
	0,85	0,92	1,56	1,72	1,72	28,9	57,8

Ці результати добре узгоджуються з даними, отриманими шляхом схемотехнічного аналізу, здійсненого за допомогою інтегрованого пакету MicroCap – 9 [12]. При цьому слід зауважити, що значення вхідних опорів для усіх розглянутих запропонованих схем значно менше ніж для ПВС і ВСУ.

Аналогічним чином можна будувати компенсаційні ВС з низьким $r_{\text{ВХ}}$ на базі схем Уїлсона, структури яких зображено на рисунку 3 а). Причому, якщо застосовувати ПНС і T_g , наведені на рисунку 2 а) – д), то значення вхідних опорів будуть близькими до наведених у таблиці 1.

Принципово іншим шляхом зменшення $r_{\text{ВХ}}$ компенсаційних ВС є додаткове уведення вхідного каскаду (Вх К) ЗБ, що має значно менший вхідний опір ніж каскад ЗЕ. Приклад побудови модернізованої схеми Уїлсона із застосуванням на вході каскаду ЗБ наведено на рисунку 3 б). Для задання робочої точки тут використовуються джерела напруги зміщення $U_{\text{зм}}$ і струму зміщення $I_{\text{зм}}$. Малосигнальний вхідний опір цієї схеми дорівнює:

$$r_{\text{ВХ}} = \frac{r_{\text{ВХ ЗБ}}}{1 + \chi \cdot K_{i33}},$$

$$\text{де } r_{\text{ВХ ЗБ}} = r_e + \frac{r_6}{1 + \beta_{\text{p-n-p}}};$$

$K_{i33} \approx \beta$ – коефіцієнт передачі струму зі входу схеми на емітер T_g ;

$\chi \approx 0,5$ – коефіцієнт передачі струму емітера $T_{рг}$ на діод внутрішнього ВС.

Таким чином малосигнальний опір модернізованої схеми Уїлсона:

$$r_{вх\text{ МУ}} = \frac{r_e + \frac{r_{\delta}}{1 + \beta_{p-n-p}}}{1 + 0,5 \cdot \beta_{n-p-n}}$$

Альтернативою схеми Уїлсона може слугувати компенсаційний ВС, що містить Вх К ЗБ і вихідний каскад ЗЕ. Для завдання робочої точки Вх К і вихідного каскадів $T_{вих}$ застосовуються джерела напруги $U_{зм}$ і струму зміщення $I_{зм}$. Схема працює таким чином: вхідний струм $I_{вх}$ із виходу Вх К поступає одночасно на бази транзисторів компенсатора $T_{кмс}$ і вихідного транзистора $T_{вих}$. Причому розгалуження колекторного струму каскаду ЗБ у бази вказаних транзисторів здійснюється приблизно в рівних частках, пропорційних $\beta_{кмс}$ і $\beta_{вих}$, тому $I_{вх} \approx I_{кмс}$. Водночас це співвідношення залежить від ідентичності параметрів $T_{кмс}$ і $T_{вих}$, а також (хоча і не значно) від напруг колектор-емітер.

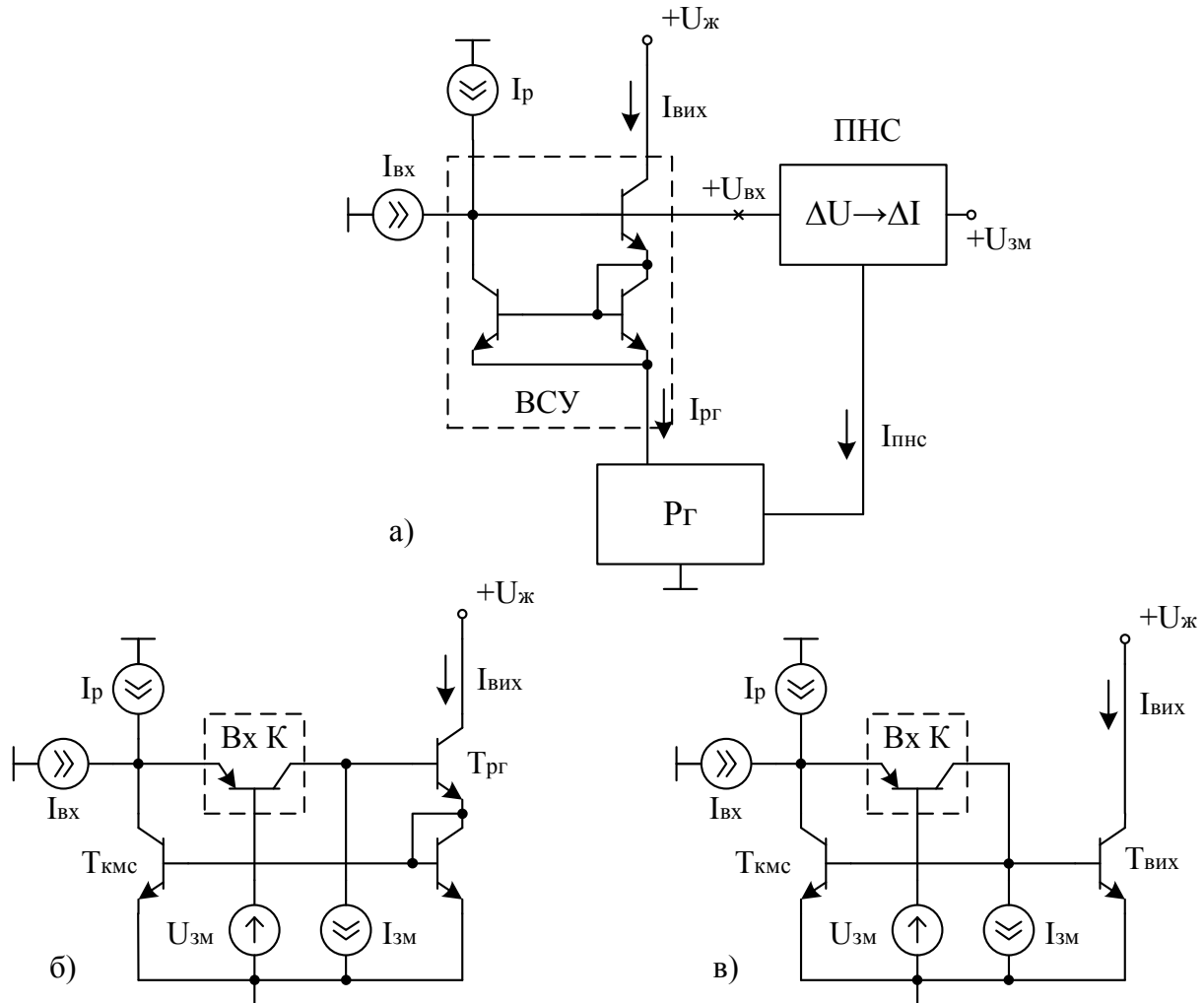


Рисунок 3 – Схемно-структурна організація компенсаційних ВС із низьким : модернізованого ВСУ із зовнішнім активним ЗЗ а); із застосуванням вхідного каскаду ЗБ б); в) із вхідним каскадом ЗБ і вихідним ЗЕ

Малосигнальний $r_{вх}$ розглянутої схеми компенсаційного ВС:

$$r_{вх\text{ КМС}} = \frac{r_{вх\text{ ЗБ}}}{1 + \chi \cdot \beta_{n-p-n}},$$

де $\chi \approx 0,5$ - коефіцієнт розгалуження колекторного струму Вх К. Остаточно маємо:

$$r_{\text{ВХ КМС}} = \frac{r_e}{1 + 0,5 \cdot \beta_{\text{п-р-п}}} \approx \frac{2 \cdot r_e}{\beta_{\text{п-р-п}}}$$

Результати комп'ютерного моделювання показують, що значення $r_{\text{ВХ}}$ для структури, зображеної на рис. 3 а) збігається з наведеними у табл. 1. Вхідні опори схем (рис. 3 б), в)) є приблизно однаковими і при $I_{\text{зм}} \approx 1$ мА дорівнюють $r_{\text{ВХ}} \approx 0,7$ Ом.

Важливою статичною характеристикою ВС є також вихідний опір $r_{\text{ВІХ}}$. Власне саме він характеризує степінь "неідеальності" генератора струму. При цьому слід зазначити, що найнижчий $r_{\text{ВІХ}}$ має параметричний ВС, для якого:

$$r_{\text{ВІХ}} \approx r_{\text{К}}^* = \frac{U_{\text{Ерлі}}(I_{\text{Г}})}{I_{\text{К}}}$$

де $U_{\text{Ерлі}}(I_{\text{Г}})$ – напруга Ерлі, яка для інтегральних малопотужних транзисторів п-р-п дорівнює 40÷70 В, а для р-п-р транзисторів 15÷20 В.

Таким чином для ПВС при $I_{\text{ВІХ}} = 1$ мА вихідний опір має значення на рівні десятків кОм. ВС Уілсона має [1]:

$$r_{\text{ВІХ}} = 0,5 \cdot r_{\text{К}} = 0,5 \cdot r_{\text{К}}^* \cdot (1 + \beta)$$

Безумовно це істотно більше, ніж у ПВС. Водночас, у багатьох випадках: прецизійні схеми, підсилювачі постійного струму та інших вимагається значно вищий рівень $r_{\text{ВІХ}}$.

У науково-технічній літературі [1,5] розглядаються шляхи збільшення вихідного опорів генераторів і ВС. Водночас, вони є недостатньо ефективними або вимагають складної реалізації (відслідкований вихід). Так застосування на виході ПВС або ВС Уілсона (ВСУ) каскадних схем дозволяють досягти максимального вихідного опору на рівні $r_{\text{ВІХ}} \text{ ВС} = r_{\text{К}}$, що усього вдвічі більше ніж у поширеного ВСУ.

Автори пропонують методи збільшення $r_{\text{ВІХ}} \text{ ВС}$, що базуються на уведенні зовнішніх і внутрішніх контурів ЗЗ. Це дозволяє істотно нейтралізувати вплив ефекту Ерлі. Причому тут за основу береться постулат про те, що максимальний вихідний опір каскадів ЗЕ (з емітерним резистором) і ЗБ не може перевищувати опору колекторного переходу $r_{\text{К}} = r_{\text{К}}^* \cdot (1 + \beta)$ [11]. Причиною цього є залежність (хоча і незначна) струму бази від напруги колектор-емітер. Тому суть запропонованих підходів щодо компенсації $I_{\text{б}}$ вихідного транзистора полягає у введенні контурів від'ємного зворотного зв'язку між базою $T_{\text{ВІХ}}$ і внутрішнім генератором струму.

Суть методів схемно-функціональної організації ВС із високим $r_{\text{ВІХ}}$ ілюструється на рисунку 4. Схеми, зображені на рис. 4 а) і б) являють собою ПВС, до виходу який підключено вихідні буферні транзистори на основі каскадів ЗБ.

Тут бази вихідних транзисторів $T_{\text{ВІХ}}$ підключено до контурів зовнішнього або внутрішнього від'ємного ЗЗ. Причому у схемі зображений на рис. 4 а) контур ЗЗ являє собою генератор струму зміщення $I_{\text{зм}}$, діод зміщення $D_{\text{зм}}$ і внутрішній додатковий ВС, вихід якого підключено до входу пристрою. Дія ЗЗ відбувається таким чином, що збільшення $I_{\text{б}}$ внаслідок змінення $U_{\text{к}}$ призводить до зменшення струму відбиття $I_{\text{вб}}$ внутрішнього ВС і відновлення $I_{\text{ВІХ}}$. Якщо $I_{\text{б}}$ зменшується, то $I_{\text{вб}}$ збільшується і $I_{\text{ВІХ}}$ знову поновлюється. У схемі на рис. 4 б) організовано контур внутрішнього ЗЗ і $D_{\text{зм}}$ замінено транзистором зміщення $T_{\text{зм}}$, котрий разом із $I_{\text{зм}}$ утворює контур зовнішнього ЗЗ. Причому дія ЗЗ тут є аналогічно, як і для першого випадку. На рисунку 4 в) і г) зображено схеми компенсаційних ВС, виходи яких організовано аналогічно тому як і для розглянутих перших двох. Діод $D_{\text{зм}}$ тут слугує для вирівнювання β транзистора $T_{\text{Г}}$ і компенсатора. Останній варіант (рис. 4 д)) являє собою схему ВСУ, на виході якої ЗЗ організовано за допомогою транзистора $T_{\text{зм}}$ і генератора струму зміщення $I_{\text{зм}}$.

Для оцінювання вихідних опорів розглянутих пристроїв доцільно проаналізувати узагальнену мало-сигнальну схему заміщення вихідного каскаду (рис. 4 е)), в якій вилучено другорядні елементи. Тут $I_{\text{Г}}$ – струм генератора, що формується транзистором $T_{\text{Г}}$, $r_{\text{Г}}$ – внутрішній опір, а $r_{\text{ВХ}}^*$ – вхідний опір вихідного каскаду на транзисторі $T_{\text{ВІХ}}$. Слід відзначити одну принципову деталь – вихідний опір вихідного каскаду ВС обмежується двома складовими. Перша – це опір колекторного переходу $r_{\text{К}}$, значення якого з урахуванням дії ЗЗ дорівнює:

$$\bar{r}_k = \frac{r_k}{1 - K_{i33}} = \frac{r_k^* \cdot (1 + \beta)}{1 - K_{i33}}$$

де r_k, r_k^* – опори колекторних переходів відповідно каскаду ЗБ і ЗК;

β – коефіцієнт передачі по струму каскаду ЗЕ;

K_{i33} – коефіцієнт передачі струму по петлі ЗЗ.

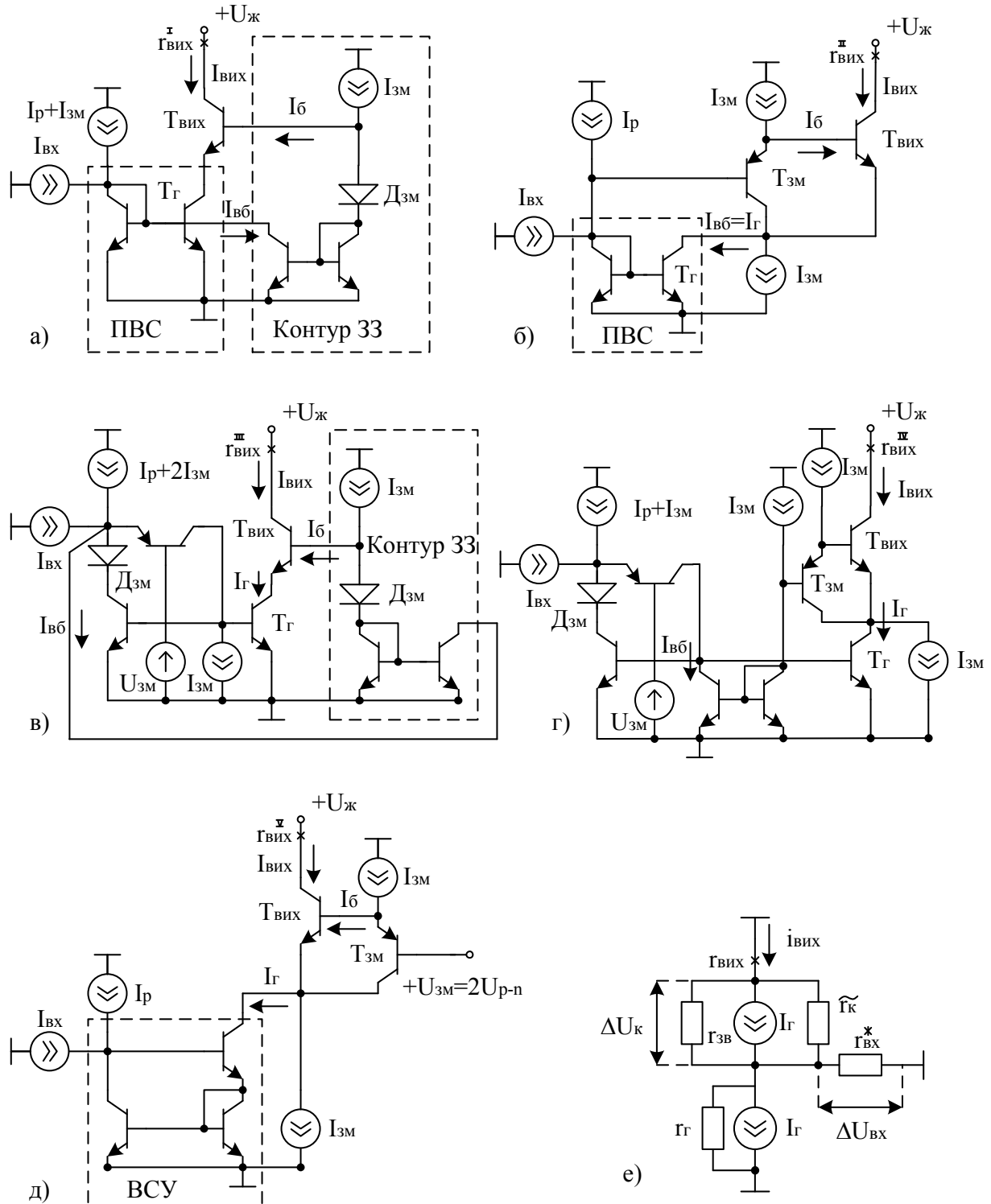


Рисунок 4 – схеми ВС із високим $\Gamma_{\text{вих}}$ на основі: а), б) ПВС; в), г) компенсаційних ВС; д) ВСУ; е) мало-сигнальна схема заміщення вихідного каскаду.

Друга складова – це опір $r_{зв}$, що залежить від рівня зворотної передачі приросту колекторної напруги ΔU_k на емітер $T_{вих}$. При цьому приріст напруги $\Delta U_{вх}$ впливає на рівень I_r у зв'язку з кінцевим значенням r_k . Звідси маємо:

$$r_{зв} = r_r \cdot K_{зв},$$

де $K_{зв} = \frac{\Delta U_{ж}}{\Delta U_{вх}}$ – коефіцієнт зворотної передачі приросту колекторної напруги ΔU_k на емітер $T_{вих}$.

Скориставшись малосигнальною схемою заміщення БТ [11], неважко показати, що:

$$K_{зв} \approx \frac{U_{Ерлі}}{\Phi_t},$$

тому:

$$r_{зв} = r_r \cdot \frac{U_{Ерлі}}{\Phi_t} \approx r_r \cdot \frac{r_k^*}{r_e}.$$

Враховуючи вищевідзначене, вихідний малосигнальний опір схеми дорівнює:

$$r_{вих} = \frac{\Delta U_k}{i_{вих}} = r_{зв} \parallel \overline{r_k} = \frac{\frac{r_k^* \cdot (1 + \beta) \cdot r_r}{(1 - K_{i33}) \cdot r_e}}{\frac{1 + \beta}{1 - K_{i33}} + \frac{r_r}{r_e}}.$$

Під час аналізу $r_{вих}$ треба відзначити той чинник, що складова $\overline{r_k}$ істотно залежить від співвідношення складових різниці $(1 - K_{i33})$. При цьому, якщо K_{i33} наближається до 1, то $\overline{r_k}$ різко зростає, і навіть може виникнути повна компенсація або перекомпенсація. За таких умов можливо, що $\overline{r_k} \gg r_{зв}$. Тоді:

$$r_{вих} = r_{зв} = r_r \cdot K_{зв}.$$

За такої ситуації для підвищення $r_{вих}$ треба використовувати у схемі внутрішній генератор струму з якомога більшим r_r , наприклад ВСУ. Підсумки схемотехнічного аналізу малосигнальних $r_{вих}$ при $I_{зм} = 1$ мА схем ВС, наведених на рисунку 4, зведено у таблиці 2.

Таблиця 2 – Порівняльна таблиця вихідного опору схем ВС

$r_{вих}$ (МОм)	$r_{вих}^I$	$r_{вих}^{II}$	$r_{вих}^{III}$	$r_{вих}^{IV}$	$r_{вих}^V$	$r_{вих}$ ПВС	$r_{вих}$ ВСУ
	80,3	72,3	97,1	176,6	262,1	0,067	3,7

Отримані результати свідчать про те, що запропоновані методи дозволяють істотно (майже на два порядки) збільшити вихідний опір ВС порівняно з ВСУ і на 3–4 порядки порівняно з ПВС.

Висновки

1. Запропоновано методи схемно-структурної організації відбивачів струму з покращеними статичними характеристиками, а саме зі зменшеним вхідним і підвищеним вихідним опором. Це досягається введенням контурів внутрішнього і зовнішнього зворотного зв'язків по струму.

2. Виведено аналітичні співвідношення для оцінювання вхідних і вихідних опорів параметричного і компенсаційних відбивачів струму. Доведено, що значення вхідних опорів запропонованими методами можуть бути зменшені на один – два порядки, а значення вихідних опорів підвищені на один – два порядки при заданому коефіцієнті передачі.

Список літератури

1. Титце У. Полупроводниковая схемотехника. / У. Титце, К. Шенк [пер. с нем. ДМК Пресс]. – М. : – ДМК Пресс, 2008 – Т. 1. – 832 с.
2. Grebene A. Bipolar and MOS Analog Integrated Circuit Design / Alan B. Grebene – , 2003. – 879 p.
3. Кестер У. Аналогово-цифровое преобразование. / У. Кестер. – М.: Техносфера, 2007. – 1016 с.
4. Гребен А. Б. Проектирование аналоговых интегральных схем. / А. Б. Гребен [Пер. с англ. Энергия] – М.: Энергия, 1976. – 256 с. с ил.
5. Коннели Дж. Аналоговые интегральные схемы. Элементы, схемы, системы и применения. / Дж. Коннели. – М.: Мир, 1977. – 439 с. с ил.
6. Степаненко И. П. Основы микроэлектроники. Учебное пособие для вузов. / И. П. Степаненко – 2-е изд. перераб. и доп. – М.: Лаборатория Базовых Знаний. 2003. – 488 с.: ил.
7. United States Patent № 3936725 Herbert Anton Schneider Current mirrors Feb. 3, 1976
8. United States Patent № 0181257 A1 Hugo Veenstra Current mirror Aug. 17, 2006
9. United States Patent № 5461343 Ryan P. Foran Current mirror circuit Oct. 24, 1995
10. Азаров О.Д. Двотактні підсилювачі постійного струму для багато розрядних перетворювачів форми інформації, що самокалібруються : монографія / О.Д. Азаров, В.А. Гарнага. – Вінниця : ВНТУ, 2011. – 156 с.
11. Степаненко И. П. Основы теории транзисторов и транзисторных схем. / И. П. Степаненко – 2-е из. перераб. и доп. – М.: «Энергия», 1977. – 616 с.: ил.
12. Амелина М. А. Программа схемотехнического моделирования Micro-Cap 8. / М. А. Амелина, С. А. Амелин – М. : Горячая линия-Телеком, 2007. – 464 с.

Відомості про авторів

Азаров Олексій Дмитрійович – д.т.н., проф., завідувач кафедри обчислювальної техніки ВНТУ, azarov2@vntu.edu.ua.

Яцик Володимир Євгенійович – студент ВНТУ, yatsyk.ve@gmail.com.

УДК 681.3

В. П. КОЖЕМ'ЯКО, Р. М. НОВИЦЬКИЙ, К. В. КОЖЕМ'ЯКО

Вінницький національний технічний університет, Вінниця

АНАЛІЗ ПАРАЛЕЛЬНИХ МЕТОДІВ ТА ЗАСОБІВ ОПТИКО-ЕЛЕКТРОННОЇ ОБРОБКИ ТА ПОРІВНЯННЯ ЗОБРАЖЕНЬ

Анотація. В статті проведено аналітичний огляд сучасних паралельних методів та засобів порівняння зображень з використанням оптико-електронних систем око-процесорного типу. Розглянуто класифікаційний аналіз ознак зображень в процесах паралельного порівняння образної інформації та переваги застосування методу KVP-перетворення для задач паралельної око-процесорної обробки та порівняння зображень.

Ключові слова: методи порівняння зображень, ознаки зображень, KVP-перетворення, око-процесор.

Аннотация. В статье проведен аналитический обзор современных параллельных методов и средств сравнения изображений с использованием оптико-электронных систем глаз-процессорного типа. Рассмотрены классификационный анализ признаков изображений в процессах параллельного сравнения образной информации и преимущества применения метода KVP-преобразования для задач параллельной глаз-процессорной обработки и сравнения изображений.

Ключевые слова: методы сравнения изображений, признаки изображений, KVP-преобразования, глаз-процессор.

Abstract. In the article the analytical review of modern methods and means of parallel comparison of images using optical-electronic eye-type processor. We consider the classification analysis of images features in a parallel comparison of imagery data and the advantages of the method KVP-transformations for eye problems in parallel-processor data processing and image comparison.

Key words: methods of image comparison, image attributes, KVP-transformations, eye-processor.

Вступ

Порівняння як основна операція обробки образної інформації входить до складних процедур розпізнавання і обробки майже на всіх етапах оброблення візуальної інформації у сучасних комп'ютерних засобах, системах технічного зору, а також у спеціалізованих око-процесорних структурах.

Задачі порівняння зображень носять неявно виражений комплексний характер та включають ряд основних етапів [1]: сприйняття візуального поля; попереднє оброблення вхідної інформації; сегментацію вхідної інформації; нормалізацію виділених об'єктів; центрування вхідного зображення по відношенню до еталону; безпосередній опис та класифікацію зображення.

Такий обов'язковий етап, як інтерпретація зображення включається частково в етап сегментації і остаточно зважається на етапі порівняння. Важливим етапом при порівнянні є попереднє оброблення [1-3] та зокрема центрування [4], що значно впливає на загальну ефективність порівняння: на час оброблення та на достовірність кінцевого результату.

Основною проблемою завдання порівняння зображень є прийняття рішення про те, чи відносяться вхідні елементи вхідного зображення до класу, що належить еталонному зображенню і являє собою певний еталон або патерн. При порівнянні зображення з еталонами [1] виникає ряд специфічних проблем при створенні систем порівняння: зображення порівнюються на складному фоні; зображення еталона і вхідних зображень відрізняються положенням у полі зору; вхідні зображення не збігаються з еталонами через випадкові завади; відмінності вхідних й еталонних зображень виникають за рахунок зміни освітленості, підсвічування, локальних завод; еталони й зображення можуть відрізнити геометричні перетворення, включаючи афінні та проєктивні [1,3,4].

Для розв'язання задачі порівняння застосовуються різні методи сегментації, нормалізації та розпізнавання зображень. У цій сфері розвиваються все більше нових ідей та методів.

Структурний метод [1] виконує обробку та пов'язане з нею порівняння виходячи з припущення про аналогію між структурою зображення і синтаксисом описової мови, яка дозволяє перенести методи теорії формальних структурних граматики для опису зображень при їх порівнянні.

Недоліки структурного методу проявляються в тому, що деякі засоби з мовної лінгвістики не пристосовані належним чином до вирішення задач розпізнавання та порівняння образної інформації.

Алгебраїчний метод виконує детермінований підхід до вирішення задач порівняння зображень. В основі алгебраїчного методу [5] лежить ідея індуктивного прояву математичних об'єктів шляхом узагальненого індуктивного визначення. В цьому методі реалізуються основні моделі і виконуються операції над ними, які дозволяють послідовно логічним шляхом та математичними визначеннями утворювати нові моделі.

Однак алгебраїчний метод має недоліки ймовірносних моделей. Проте алгебраїчні моделі мають значно менші обмеження в підвищенні точності розділення класів об'єктів і фонів та адаптивності до часткових спотворень опису об'єктів за умови збереження основних їх частин.

Кореляційний метод [1,6] ґрунтується на визначенні коефіцієнтів кореляції між попередньо сформованими еталонами (патернами) і зображеннями об'єктів, що аналізуються; на основі попередньої статистичної обробки зображень з обчисленням координат автокореляційних функцій [1]. В процесі

попередньої обробки кореляційного методу здійснюється власне порівняння зображень (кореляція) з еталоном. Кореляційна функція при цьому визначається [1]:

$$R(\xi, \eta) = \int_{\Omega'} B(x, y) B_e(x - \xi, y - \eta) dx dy \quad (1)$$

де $B_e(x, y)$ – двовірна функція розподілу яскравості еталона; Ω' – область визначення функцій $B(x, y)$ і $B_e(x, y)$.

Даний метод порівняння зображень, незважаючи на оригінальність, має свої недоліки. В першу чергу, для поділу ознак на класи необхідно проводити сканування зображення, при цьому виникає помилка процесу оброблення. Крім того потрібно здійснювати поворот вхідного зображення по відношенню до еталона. Для визначення світлового центру зображення і переносу початку координат потрібна велика кількість обчислень, що неминуче призводить до збільшення обчислювальних операцій обробки, значного ускладнення технічної реалізації та в кінцевому варіанті до збільшення часу процесу порівняння.

Геометричний метод [1] розглядає концепцію образу та відповідних йому ознак. Об'єктивними характеристиками образів є їх ознаки. Принцип порівняння полягає в тому, що системі висувається реалізація окремих класів (навчаюча послідовність). Після цього система сама генерує правило кінцевого вирішення, обираючи його з деякого класу правил, закладених в неї при створенні та навчанні.

Основним недоліком такого підходу є значні похибки при визначенні мір (шаблонів) близькості компактних множин в периферійному створюємому просторі. При повороті на 15 градусів, або аксіальному зміщенні ймовірність точного порівняння значно зменшується (5-10%). Це потребує попередньої операції центрування.

Спектральний метод [7] базується на тому, що у якості ознак об'єктів при порівнянні або розпізнаванні зображень використовують просторово-часовий спектр $S(\omega_x, \omega_y)$, або перетворення Фур'є F функції яскравості $B(x, y)$ (2) виду [7]:

$$S(\omega_x, \omega_y) = F\{B(x, y)\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} B(x, y) e^{-j(\omega_x x + \omega_y y)} dx dy \quad (2)$$

Хоча порівняння і обробка зображень спектральними методами здійснюється в оптичному діапазоні, вони становлять інтерес при побудові цифрових фільтруючих, кореляційних та розпізнавальних пристроїв при виборі інформативних ознак.

До недоліків, які притаманні цьому методу можна віднести: чутливість до зсуву вхідного зображення відносно узгодженого фільтра, що призводить до зміщення функції кореляції на виході операції обробки зображення; чутливість до повороту зображення відносно еталону, що потребує необхідності проведення операції обертання зображення, а це значно збільшує кількість операцій та відповідно машинний час; неінваріантність до масштабування вхідного зображення [1]. Це також потребує попереднього оброблення, а саме операції точного центрування зображення, що аналізується по відношенню до еталонного.

У методі на базі теорії нечітких множин [8] ключовим визначенням є нечіткі ознаки. Кінцеве рішення про відповідність зображення до визначеного класу приймається за максимальним значенням коефіцієнта відповідності, що припадає на цей клас. Достовірність обробки та порівняння забезпечується коректним застосуванням операцій максимізації – max та мінімізації – min за допомогою нечіткої логіки, що працює з експертною матрицею знань [1,8]. Як правило в якості нечіткої логіки використовується нейронні мережі або нейронні елементи. До недоліків слід віднести великий коефіцієнт похибки та необхідність до тривалого навчання нейронної експертної системи.

Метод суміщення з еталонами [1] – історично перший метод розпізнавання, основний принцип якого полягає в порівнянні зображення з наперед сформованим набором ідеальних еталонів. Системи порівняння, побудовані на цьому методі, можна розділити на: системи із зберіганням еталонних наборів; системи на базі фотографічних масок; системи, що використовують електричні параметри.

Метод просторово-зв'язаних процесів [1] застосовується при структурному аналізі зображень. Він представляє собою багатокрокову структуру, що досліджує часткові зв'язки між структурними компонентами зображення. З метою зниження чутливості до спотворень та шумів протягом формування зображення застосовується метод представлення зображення з узагальненим просторовим W-зв'язаним спектром.

Перевагою цього методу порівняння зображень [1] є незначне збільшення відношення сигнал/шум, що забезпечується відповідним порівнянням W-зв'язаного спектра зображень з їх Q-спектрами.

В теперішній час існує величезна кількість принципів, методів та алгоритмів порівняння зображень. Серед основних можна виділити статистичні, лінгвістичні або нечіткі підходи [1,9]; детерміновані та неперервно-групові підходи [10]. Методи порівняння зображень класифікують в залежності від типу інформативних ознак або від типу операцій обробки інформації (алгебраїчні, структурні, геометричні, кореляційні, спектральні, обробка на базі нечітких множин) [10]. В залежності від подання вхідного зображення виділяють також кореляційні, ознакові та синтаксичні методи [1]. Є також спеціалізовані методи порівняння зображень, які не є настільки гнучкими та універсальними, однак вже тривалий час використовуються для вирішення конкретних практичних задач.

Метод зондів використовується для порівняння літер і цифр. Він базується на аналізі деяких ознак, істотних для даного зображення, і дає можливість класифікувати рукописні знаки, написані з деякими розбіжностями в розмірі та стилі написання [10]. На зображення, яке порівнюється, накладається спеціально підібрана система електронзондів, які розглядаються як координати в просторі ознак.

Метод маркування зображень [1] полягає в тому, що зображенню штучно надаються ознаки, які визначають його в процесі порівняння. Суть методу полягає в тому, що кожне зображення до самого порівняння має пройти спеціальну обробку в пристрої маркування.

Квазітопологічний метод. Зображення типу букв алфавіту, можна трактувати як плоскі графи, а різне написання однієї і тої ж букви розглядається як гомеоморфні перетворення деякого ідеального знаку. Цей метод порівняння може бути реалізований з використанням графів за допомогою слідкуючої розгортки, що здійснює обхід зображення по зовнішньому контуру. Даний метод не може забезпечити якісного порівняння у випадку поганої якості надрукованих знаків і має не високу швидкість процесу порівняння.

Метод потенціальних функцій [1] використовує подання зображень у вигляді векторів в просторі вхідних сигналів. Алгоритм цього методу базується на основній гіпотезі про характер функцій, що розділяють відповідні образам множини.

Метод стохастичної апроксимації [1] використовує поняття функцій регресії та їх коренів. Ідея застосування цього методу для вирішення задач порівняння пов'язана з вибором функції втрат такою, щоб вона дозволяла організувати рекурентну процедуру.

Метод мінімізації емпіричного ризику, який прямо залежить від існування рівномірної збіжності частот до ймовірностей за класами подій. Алгоритми даного методу можуть бути використані, якщо сміність класу правил алгоритму невелика.

Метод допустимих перетворень дозволяє одночасно з оцінкою шуканого параметра знайти і оцінку параметра перетворення, тим самим для кожного зображення можна вказати еталон максимальної схожості, а також показати, чим зображення відрізняється від початкового еталону. Крім того метод є інваріантним до припустимих перетворень.

Методи порівняння зображень для чисельних задач використовуються в основному при програмній реалізації систем пошуку і порівняння схожих зображень (image retrieval systems). До них відносяться: метод збільшення яскравості; метод текстурної фільтрації; метод детектора границь на базі оператора Собеля; метод порівняння за допомогою гістограм. Розглянемо коротко дані методи.

Метод збільшення яскравості полягає в тому, що при порівнянні зображень, значення яскравостей їх елементів (пікселів) в сумі може бути збільшена або зменшена, що дозволяє фіксувати різницю. Чим ця різниця менша, тим яскравіший елемент зображення. По сумарній різниці яскравостей, або різниці кольорів (при порівнянні яскравостей кольорових зображень) приймається рішення про відповідність, або не відповідність вхідного зображення еталону. До недоліків цього методу можна віднести велике розходження при порівнянні зображень різної роздільної здатності, а також за умов різної фонові освітленості, що вносить значну похибку порівняння.

Метод текстурної фільтрації; передбачає фільтрацію на базі пошуку характерних ознак певної області (сегмента) в зображенні, що порівнюється. Такими областями можуть бути структури, що повторюються або структури з певними геометричними формами. Недоліки методу текстурної фільтрації такі ж як і в методі сегментації, оскільки в цих методах передбачена деяка ступінь подібності між собою.

Метод детектора границь на базі оператора Собеля передбачає використання детектора границь на основі оператора Собеля, який виконує вимірювання двовимірного просторового градієнта на зображенні і виявляє області з великим значенням цього параметра. Ці області і відповідають краям зображення. Це метод використовується для оцінки модуля градієнта в кожній точці чорно-білого зображення. Порівняння зображень в такий спосіб є достатньо точним, але даний метод має вагомий недолік – високу чутливість до зміни роздільної здатності та чіткого розміщення центрів зображень.

Метод порівняння за допомогою гістограм проводить порівняння частоти розташування ліній градацій відтінків кольору або градацій відтінків сірого в напівтонових зображеннях. Розходження і

оцінка порівняння базується при використанні різних характеристик роздільної здатності і градацій, наприклад ущільненні ліній в областях від 0 до 255. Порівняння гістограм дозволяє виявляти розходження не помітні людському оку, що робить цей метод досить високоточним, але разом з тим і повільним при реалізації на послідовних алгоритмах і апаратних архітектурах.

Основні труднощі практичної реалізації цих методів полягають у складності опису конкретного зображення і великому об’ємі обчислювальних операцій над їх параметрами. Для цих спеціалізованих методів необхідно наперед визначити лімітовані критерії порівняння, які мають дуже вузьку галузь застосування.

Узагальнюючі результати аналізу відомих методів порівняння представлені у вигляді удосконаленої класифікації на рис. 1, в якій враховано обов’язкові операції попередньої обробки зображення перед порівнянням: фільтрацію та центрування.

На основі аналізу існуючих методів порівняння встановлено, що ефективність та достовірність кінцевого результату більшості з них значно залежить від розташування вхідного зображення по відношенню до еталона. Це потребує попереднього оброблення у вигляді центрування зображень.

Узагальнена класифікація методів порівняння зображень (рис.1) об’єднує методи за 4-ма підходами до розв’язання задачі порівняння: кореляційні, ознакові, синтаксичні та на основі нейромереж.

Одними з новітніх методів оброблення є методи з використанням трьох основних паралельних перетворень зображень, що розроблені учнями наукової школи проф. Кожем’яко В.П. До них належать: KVP [11-14], Q[2,16-19] та П [16-19] – перетворення. Ці три методи обробки зображень дозволяють усунути більшість недоліків, що притаманні вище приведеним методам оброблення зображень. Але ефективність та достовірність їх результатів також залежить від точності розташування вхідного та еталонного зображень, що потребує попереднього центрування цих зображень. Слід розглянути їх більш детально.

Подальший розвиток структур обробки зображень око-процесорного типу привів до створення Q-перетворення [2,16,17,19] та способу око-процесорної обробки зображень на основі нього[2].

Метод KVP-перетворення [11-14] базується на представленні і перетворенні вхідної інформації зображення у вигляді логіко-часових функцій (ЛЧФ), які несуть інформацію у вигляді тривалостей груп часових інтервалів [11-14]. Апарат ЛЧФ є подальшим розвитком апарату векторно-перемікаючих функції проф. З. Л. Рабиновича.

Q-перетворення образної інформації [2,18] використовується в процесі око-процесорного оброблення (рис. 1) та базується на методі узагальненого контурного препарування зображень [13].

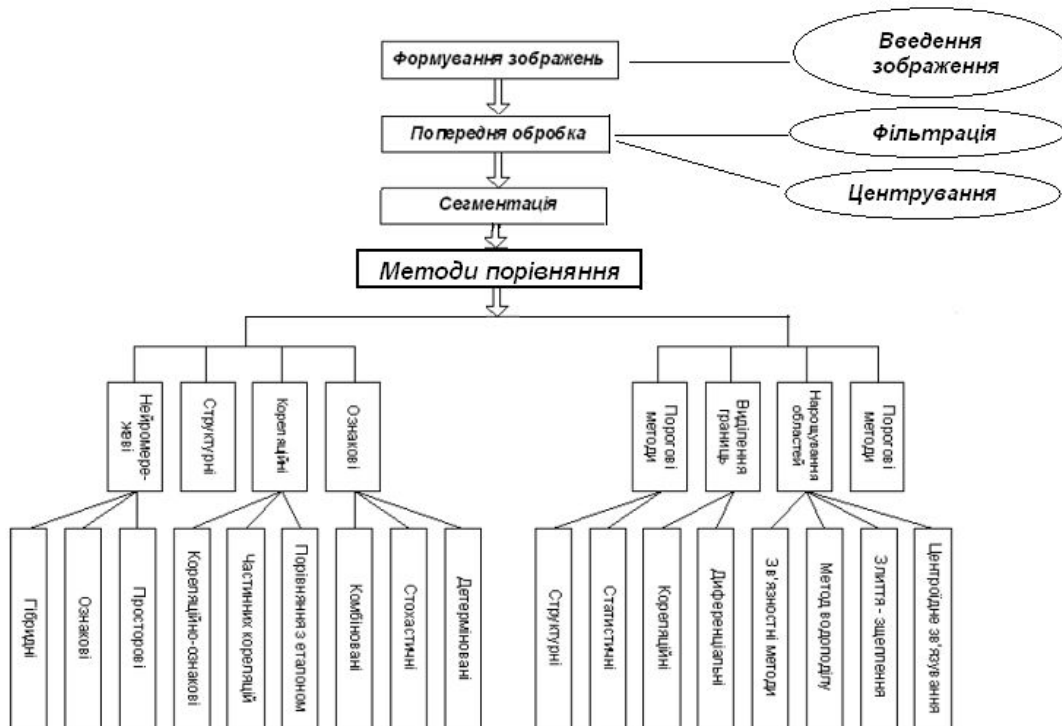


Рисунок 1 – Узагальнена класифікація відомих методів порівняння зображень

Суть Q-перетворення зображень полягає у формуванні сукупності узагальнених контурних препаратів [2] зображення та знаходженні часткових сум qr, k – коефіцієнтів одиничних відліків зображення у вікнах розміром r, k при $k=1 \dots (rt-1)$ [2,18,19]. Причому r та k обирають кратними степеню 2 [2,13,19].

Якщо значення просторових сум $q(t,k)$ розмістити у центрах відповідних їм вікон, то в результаті формується нове зображення, яке можна розглядати як Q-спектр зображення [2,13]. Таким чином, основна відмінність застосування Q-перетворення для порівняння зображень від кореляційних методів полягає в тому, що немає необхідності суміщати зображення, що порівнюються, і підраховувати кількість одиничних відліків, які збіглися, та потім порівнювати отримані суми. При використанні Q-перетворення суми одиничних відліків (q-коефіцієнти) формують для кожного із зображень [2].

Метод ПП-перетворення (паралельно-ієрархічне перетворення) при око-процесорному обробленні (рис. 1) передбачає узагальнене просторово-зв'язане препарування зображень [2,16,18,19]. Його суть полягає в одночасному використанні послідовності масивів інформації, які утворюють множини інформаційних полів на різноманітних рівнях ієрархії, рекурсивному формуванні нових послідовностей інформаційних потоків на різноманітних рівнях ієрархії, що дозволяє реалізувати стратегію багаторівневої взаємодії: «від загального до часткового» [2,18]. В результаті прямого мережного ПП-перетворення утворюється одновимірна матриця $\varphi(t, p)$ [2,18]:

$$\varphi(t, p) = \Phi(j, i, t, p) \sum_{j=1}^S \sum_{i=1}^{n_s} \mu(i, j), \quad (3)$$

де $\Phi(j, i, t, p)$ – ядро нелінійного прямого ПП-перетворення над елементами $\mu(i, j)$ вхідного зображення з розмірністю $S \times n_s$.

Базовим для ПП-перетворення є метод узагальненого просторово-зв'язаного препарування (УПЗП) [1], суть якого полягає у перетворенні багатоградацийних за інтенсивністю зображення у двоградацийні за допомогою формування просторово-зв'язаних узагальнених контурів [2].

На відміну від узагальненого Q-перетворення, що включає виділення образів за допомогою функції маски послідовні УПЗП образи формують сукупності функцій додатних, від'ємних та нульових УПЗП препаратів. ПП перетворення є мережним перетворенням з формуванням узагальнених мережних контурних препаратів і їх попарним ранговим порівнянням [2].

Метод ПП-перетворення зображень при порівнянні не тільки зручний для оброблення методами «картинної» логіки [2], але й дозволяє обчислювати кореляційні функції з підвищеною точністю, швидкістю, завадостійкістю, зокрема, до повороту зображень. При порівнянні з іншими методами перетворення зображень з розкладанням у визначеній системі базисних функцій, ПП-перетворення використовує розкладання за адаптивною системою базисних мережних функцій у вигляді рядів, формування яких залежить від структури самого зображення. Це і спрощує процедуру перетворення, знижує середньоквадратичну помилку, але збільшує кількість і ресурс машинного часу, що є основним недоліком ПП-перетворення. Тому область застосування ПП-перетворення обмежується паралельними оптико-електронними системами око-процесорного типу [2].

У табл. 1 наведено основні показники ефективності відомих методів порівняння зображень.

Таблиця 1 – Основні характеристики відомих методів порівняння зображень

Назва методу порівняння	Середні характеристики ефективності методу				
	Достовірність результату порівняння, %	Сер. кількість машинних тактів (зобр. з розмір. $N \times N$ *)	Можливість паралелізму операцій	Середня вартість опер., грн.	Допуст. зміщення зобр., град.
1	2	3	4	5	6
Структурний	80	$N \times N \log 2N$	Ні	$1.8 \cdot 10^{-8}$	<10
Алгебраїчний	70	$N \times N \log 2N$	Ні	$0.7 \cdot 10^{-8}$	<8
Кореляційний	85	$N \log 2N$	Так	÷	<15
Геометричний	80-85	$N \log 2N$	Так	÷	<15
Спектральний	60	$N \log 2N$	Так	÷	<10
Метод на базі теорії нечітких множин	40-50	$2N \log 2N$	Так	$0.7 \cdot 10^{-4}$	<15
Суміщення з еталонами	75-80	$N \times N$	Ні	$0.6-1.3 \cdot 10^{-8}$	<5
Просторово-зв'язаних процесів	55	$2N \log 2N$	–	–	<7

продовження табл. 1

1	2	3	4	5	6
Метод зондів	–	$N \log_2 N \times N$	Ні	–	–
Маркування зображень	45	N^2	Ні	$1.7 \cdot 10^{-6}$	–
Потенціальних функцій	40	$\sim N^2$	Так	$0.7 \cdot 10^{-4}$	< 10
Стохастичної апроксимації	50-65	$N \log_2 N \times N$	Ні	$0.7 \cdot 1 \cdot 10^{-8}$	< 8
Мінімізації емпіричного ризику	–	$N \times N \log_2 N \times N$	Ні	÷	< 10
Допустимих перетворень	60-70	$N \times N \log_2 N$	Так	÷	–
Метод збільшення яскравості	65-70	$1/2 \log_2 N$	Так	÷	< 9
Текстурна фільтрація	60-70	$N \times N \log_2 N$	Ні	÷	< 10
Детектор границь на базі оператора Собеля	80-85	$1/2 N^2 \log_2 N$	Ні	÷	< 7
Метод гістограм	80-85	$N \times N$	Так	÷	< 12
KVP-перетвор.	80-85	$\log_2 N$	Так	$0.5 \cdot 1 \cdot 10^{-2}$	< 15
Q-перетвор.	70-80	–	Так	$0.5 \cdot 1 \cdot 10^{-3}$	< 15
III-перетвор.	70-85	–	Так	$1 \cdot 10^{-3}$	–

* Значення приведені за умов розмірності квадратних зображень $N \times N$ при реалізації моделей методів на послідовній апаратній платформі, окрім паралельних методів KVP-, Q- та III-перетворень із середньою продуктивністю обчислень $2 \cdot 10^{10} - 1.5 \cdot 10^{12}$ оп/с.

На практиці різноманітність розв’язуваних задач за допомогою розглянутих методів обробки і порівняння зображень завжди перевищує можливість цих методів і засобів їх апаратної обробки. Тому створення вузько спеціалізованих систем менш ефективне, ніж створення універсального методу і універсальних засобів апаратної обробки з підвищеною адаптацією до інформаційної задачі або до варіації завдосигнальних обставин.

Аналіз результатів досліджень відомих методів порівняння показує, що в усіх існуючих методах необхідним є забезпечення чіткого і точного позиціонування вхідного та еталонного зображень, що передбачає обов’язкове їх центрування [20], з метою виключень аксіальних та радіальних зміщень, які значно впливають на достовірність рішення.

Крім того дуже важливим є виконання операцій попереднього аналізу, з метою виключення за допомогою фільтрації негативних факторів. Забезпечення цих операцій попереднього оброблення дозволить значно скоротити обчислювальні ресурси на виконання складних обчислень в процесі порівняння зображень та підвищити достовірність сучасних методів порівняння зображень.

Типи оптоелектронних засобів для паралельного порівняння візуальної інформації

Існуючі оптоелектронні пристрої [5] для паралельного порівняння зображень, виконують функцію порівняння зображень без застосування процесу сканування й тому мають два паралельних оптичних входи, на які подаються порівнювані зображення А та В, і оптичний вихід, на якому формується вихідне зображення С. Завдяки відсутності процесу сканування, порівняння зображення виконується за один такт, тобто час порівняння не залежить від розмірності зображення, а залежить тільки від швидкодії обраної оптоелектронної елементної бази.

Залежно від виду вхідного зображення пристрої для паралельного порівняння зображень діляться на 3 види: компаратори зображень; пристрої для порівняння-віднімання зображень [5]; пристрої порівняння зображення з порогом (граничні пристрої).

Компаратори зображень є аналого-цифровими пристроями, що видають на виході цифровий сигнал – бінарне зображення, яке пов’язане із вхідними аналоговими сигналами (напівтоновими зображеннями) нелінійною залежністю типу [5]:

$$c_{ij} = \begin{cases} 1, \text{ якщо } a_{ij} \geq b_{ij}, \\ 0, \text{ якщо } a_{ij} < b_{ij}. \end{cases} \quad (4)$$

де $A = \{ a_{ij} \}$ – перше вхідне напівтонове (багатоградаційне) зображення; $B = \{ b_{ij} \}$ – друге вхідне напівтонове зображення; $C = \{ c_{ij} \}$ – вихідне бінарне зображення.

Пристрої для порівняння-віднімання зображень відрізняються від компараторів тим, що виконують аналогову нелінійну операцію «обмежена різниця», яка описується виразом [5]:

$$c_{ij} = \begin{cases} a_{ij} - b_{ij}, & \text{якщо } a_{ij} \geq b_{ij}, \\ 0, & \text{якщо } a_{ij} < b_{ij}. \end{cases} \quad (5)$$

Пристрої порівняння-віднімання зображень більш функціонально гнучкі, ніж компаратори зображень, оскільки дозволяють не тільки зрівняти два зображення, але й оцінити і визначити ступінь різниці між ними.

Пристрої для визначення ступеня збігу зображень призначені для використання в оптоелектронних кореляторах, системах розпізнавання зображень та системах технічного зору.

Основний недолік електронних кореляторів полягає в сильній залежності часу обчислення кореляційної функції від розмірів зображення, що порівнюється та еталонного зображень, що значно ускладнює реалізацію порівняння в режимі реального часу.

Дані пристрої порівняння, як компаратори й пристрої для порівняння-віднімання зображень, мають два оптичних картинних входи, але вихід у них один – електричний, на якому видається аналоговий сигнал, пропорційний ступеню збігу двох зображень, у якості якого може використовуватися одна з наступних величин [5]: скалярний добуток векторів (кожен вектор – набір пікселів зображення); евклідова відстань у багатовимірному просторі пікселів 2-х зображень; манхетенська відстань у багатовимірному просторі 2-х зображень.

Серед основних сфер застосування операції порівняння зображень можна назвати: оптоелектронні паралельні аналогово-цифрові процесори; оптоелектронні корелятори для систем технічного зору; оптоелектронні штучні нейронні мережі.

В оптичних голографічних кореляторах час обробки мало залежить від складності та розмірності зображень (за рахунок паралельності обробки) і визначається часом введення й виведення зображень у корелятор. Однак цей тип кореляторів має досить вагомий недолік – складність та дороговизна реалізації.

Прикладом пристрою для порівняння зображень може бути пристрій на основі паралельних багатофункціональних оптоелектронних схем (ПБОЕС) [5], який включає оптичний операційний підсилювач. Пристрої для порівняння зображень на ПБОЕС характеризуються високою роздільною здатністю (до 300 ліній/мм), але низькою швидкістю (одиниці-десятки мілісекунд) і поганою рівномірністю характеристик по апертурі (розкид 10-15%), що обмежує точність порівняння. На ПБОЕС можна будувати пристрої для порівняння зображень всіх трьох типів (рис. 2) [5].

Пристрої для порівняння зображень на основі ПБОЕС можуть профункціонувати як компаратори зображень. Основними з них є такі:

- 1) на основі оптоелектронних ІС, що представляють собою безшовні набори планарних лінійок інтегральних мікросхем, що сполучають фотоприймачі й інтегральні електронні компаратори;
- 2) на основі матриць світловипромінювальних фототиристорів;
- 3) на основі SEED-приладів (self-electrooptic effect device) [4].



Рисунок 2 – Паралельні оптоелектронні пристрої для порівняння зображень

Недоліком пристроїв порівняння зображень на основі фотодіодів й інтегральних схем є: мала кількість пікселів у матриці, великий розмір одного пікселя, технологічний розкид параметрів фотодетекторів, що знижує точність порівняння. Недоліком пристроїв порівняння зображень на основі світловопромінювальних фототиристорів є мала кількість пікселів у матриці, великий розмір одного пікселя, досить великий споживаний струм (5...10 мА на кожен в фототиристор), відносно велика енергія перемикавання, необхідність подачі негативної імпульсної напруги живлення для скидання інформації в матрицю, труднощі при технологічному виготовленні.

Найбільш поширеними є методи та засоби порівняння, які базуються на виділенні ознак зображень. Вони дозволяють з великою достовірністю виконувати порівняння об’єктів, причому достовірність порівняння прямо пропорційна кількості присутніх ознак [17].

Ознаки – найбільш стійкі характеристики образу, які зберігають свої значення при різних просторово-часових трансформаціях образу, таких як зсуви, повороти, зміни масштабу, а також при зашумленні образу. Стійкість ознак дозволяє людині впевнено орієнтуватися в навколишньому середовищі, а їх об’єктивний характер відкриває шлях до моделювання процесів порівняння зображень [1,2].

Серед задач, що розв’язуються за допомогою теорії розпізнавання образів, є задачі автоматичного порівняння і розпізнавання в реальному часі зображень у видимому, інфрачервоному та інших діапазонах, тобто двовимірних полів.

Виділяються дві основні проблеми [1], без вирішення яких немислимий процес порівняння: попередня обробка вхідної інформації; конструювання системи ознак.

Основою процесу порівняння є вибір оптимального набору ознак [1,4]. Без його задовільного розв’язання ніяке ускладнення вирішальних правил не може дати необхідного ефекту. Від вдалого вибору залежить алгоритм порівняння та складність апаратної реалізації.

Авторами вдосконалено відому класифікацію ознак [1], у якій визначено нові типи вторинних геометричних ознак (рис. 3.) Розглянемо детальніше кожен з типів ознак.



Рисунок 3 – Загальна класифікація ознак об’єктів

Первинні ознаки використовуються для виділення в загальних полях зображень тих областей, які відповідають об’єктам. При цьому виділення областей (сегментація зображень) може здійснюватися на основі деяких інтегральних, диференціальних, або локальних інформативних характеристик аналізованого зображення [7].

До числа найпоширеніших інтегральних ознак, що використовуються для сегментації, відносяться гістограми елементів зображень або елементів їх контурів, що реалізуються на основі таких параметрів, як яскравість, колір, контраст, або похідних від них параметрів.

До диференціальних ознак відносяться градієнти, вищі похідні яскравостей пікселів, що реалізуються через відповідні диференціальні оператори.

Вторинні ознаки відрізняються від первинних тим, що аналізована інформація для вторинних ознак повинна включати не лише відповідні параметри сегментованих областей об’єктів, але й правила перетворення первинних ознак у форми, що будуть використовуватися при подальшому порівнянні об’єктів [7]. При цьому ефективність реалізації вторинної ознаки залежить одночасно від повноти і надлишковості представленої інформації. Особливий інтерес з погляду підвищення ефективності порівняння представляють саме такі вторинні геометричні ознаки як: центр зв’язності, центр мас та вісь

орієнтації зображення, оскільки дозволяють найбільш точно забезпечити умови збіги вхідного та еталонного зображень при використанні структур компараторів.

Інваріанти моментів являють собою спеціальні функції від центральних нормованих моментів областей зображень об'єктів і забезпечують інваріантність значень даних ознак до ортогональних перетворень областей об'єктів в площині зображень: переносу, симетрії, повороту, зміни масштабу.

Морфологічні ознаки мають кращу, ніж інваріанти моментів, селективність – за рахунок великої орієнтації на детальність опису форм об'єктів, що представляються булевою алгеброю.

Структурні ознаки являють собою вторинні образи, які називаються непохідними елементами або примітивами, а також співвідношення між ними, що характеризують структуру образу.

Текстурні ознаки перевершують структурні морфологічні ознаки і рівносильні інваріантам моментів або стохастичним інваріантам. В основі цих ознак лежить поняття текстури – найважливішої характеристики зорових моделей зображень.

Логічні ознаки об'єктів, які потрібно порівняти, розглядаються як елементарні висловлювання, що приймають два значення істинності з повною точністю відповідно до наявності або відсутності деяких властивостей чи елементів у об'єктах порівняння.

Детерміновані ознаки приймають конкретні числові значення, які можуть розглядатися як координати точки в ознаковому просторі, що відповідає даному об'єкту. Найчастіше під час порівняння великомасштабних зображень використовують такі системи детермінованих ознак, як топологічний опис об'єкта, спектральний та інтегральний опис зображень.

Імовірнісні ознаки – ознаки, випадкові значення яких розподілені по всіх класах об'єктів.

Найпоширеніші методи виділення ознак зображень: метод центрування зображень [1,2,4]; методи виділення центральних моментів [4].

Переваги застосування методу KVP-перетворення для задач паралельного око-процесорного порівняння зображень

Одним із якісно нових підходів у системі методології сучасних інтелектуальних структур є наукові дослідження, які здійснюються в напрямі створення образного комп'ютера, що відповідають пріоритетним напрямкам розвитку науки в Україні.

Створення пристроїв око-процесорного типу [1-4,16] базується на оригінальному способі перетворення інформації, який максимально наближений до природного паралельного введення і оброблення, за рахунок відокремлення одночасно деякої кількості визначників різних типів [1].

Метод KVP-перетворення базується на використанні: принципу квантування часу будь-якого типу сигналу на часові інтервали, тобто перехід до ЛЧФ (4), в яких інформація представляється у вигляді тривалостей імпульсів; універсальної ЛЧФ – «ключ-функції», яка дозволяє повністю описувати вхідну інформацію; паралельному обробленні вхідних масивів інформації у вигляді тривалостей часових інтервалів; простих операцій (додавання, віднімання, множення) за допомогою тривалостей часових інтервалів (перехід до багатозначної логіки); забезпечення внутрішньої класифікації ЛЧФ способом формування та відокремлення окремих інтервалів ЛЧФ та формування первинних ознак (визначників) вхідних даних з подальшим визначенням їх якісних та кількісних характеристик.

KVP-перетворення складається з чотирьох етапів [2,15]: прийом інформації; перетворення сигналів на ЛЧФ; виділення визначників об'єкта аналізу; якісний розподіл визначників. За методом оброблення даних у процесі KVP-перетворення поділяють на: паралельне оброблення та оброблення зі зв'язністю [2].

Перевагою KVP-перетворення є застосування квантових перетворень сигналів на ЛЧФ. При цьому, усі сигнали вхідного образу є функціями єдиного типу, що обумовлює відсутність помилок при переході від обробки сигналів одного типу до іншого. Також, універсальний сигнал дозволяє підвищити точність результату та уникнути методичної помилки при обробці вхідних інформаційних паралельних сигналів від зображення.

ЛЧФ [1,2] – це послідовність часових виразів, які правильні тільки на визначених проміжках. ЛЧФ наближені до апарату векторно-часових перемикаючих функцій, а наявність їх часових тотожностей дозволяє виражати часові функції, які не можна описати за допомогою звичайних перемикаючих функцій. Він може бути представлений для опису та проектування систем обробки інформації на рівні операцій над безперервними процесами у вигляді

$$f(t, t_i, a_i) = \begin{cases} t - t_i, t_i \leq t \leq t_i + a_i, \\ 0, t_i > t > t_i + a_i. \end{cases} \quad (6)$$

де t , t_i – поточне та початкове значення часового аргументу; a_i – період існування, еквівалент вхідної інформації у вигляді суми значень елементів дискретизації відрізка існування часового аргументу

ЛЧФ мають мінімальний часовий інтервал з однаковою амплітудою [1,2]. Найпростіша ЛЧФ – це функція квантована за рівнем зі значимістю 2 – приймає значення "0" або "1". Виходячи з цього можна сказати, що за основу математичного опису ЛЧФ можна взяти математичний опис функцій Хаара або Уолша [1, 2]. У системах око-процесорного типу ЛЧФ отримується шляхом квантування світла часом. У вказаному способі око-процесорного розпізнавання зображень з виділенням ознак паралельно проектується зображення та перетворення.

Висновки

1. В статті проведено аналітичний огляд сучасних паралельних методів та засобів порівняння зображень з використанням оптико-електронних систем око-процесорного типу. Це дозволило визначити їх переваги та недоліки, а також оцінити базові характеристики.

2. Вдосконалено загальну класифікацію ознак об'єктів, що дозволило включити спеціалізовані ознаки: центр мас, вісь орієнтації та центр зв'язності, які роблять можливим суттєве підвищення достовірності порівняння зображень шляхом виконання якісної попередньої обробки.

3. Встановлено: орієнтовані на операцію порівняння зображень методи потребують попереднього оброблення у вигляді центрування вхідного та еталонного зображень; для реалізації швидкісного порівняння зображень в режимі реального часу на оптоелектронній елементній базі оптимальним є метод КВР-перетворення з орієнтацією його на операцію центрування зображень; існуючі моделі КВР-перетворення потребують вдосконалення, з орієнтацією на більший паралелізм і більшу кількість ознак, які впливають на точність кінцевого результату.

Список використаних джерел

1. Квантові перетворювачі на оптоелектронних логіко-часових середовищах для око-процесорної обробки зображень: [Монографія] / В.П. Кожем'яко, Т.Б. Мартинюк, О.І. Суприган, Д.І. Клімкіна. – Вінниця: УНІВЕРСУМ-Вінниця, 2007. – 126 с. – ISBN 978-966-641-219-8 .

2. Образний відеокомп'ютер око-процесорного типу : [Монографія] / В.П. Кожем'яко, Г.Л. Лисенко, А.А. Яровий, А.В. Кожем'яко. – Вінниця: УНІВЕРСУМ-Вінниця, 2008. – 215с. – ISBN 978-966-641-261-7.

3. Кожем'яко В.П. Погляд на природу штучного інтелекту / В. П. Кожем'яко // Вісник Вінницького політехнічного інституту.– 1997. – №1. – С.26-30. – ISSN 1997-9266.

4. Кожем'яко В. П. Око-процесорна обробка та розпізнавання образної інформації за геометричними ознаками : [Монографія] / С.І. Кормановський, В.П. Кожем'яко. – Вінниця: УНІВЕРСУМ-Вінниця, 2008. – 160С. – ISBN 978-966-641-240-2 .

5. Колесниченко О.К. Области применения операции параллельного сравнения изображений и устройства для ее осуществления (обзор) / О.К. Колесниченко, В.П. Кожемяко, Сами Важих Хейреддин // Оптико-електронні інформаційно-енергетичні технології. – 2006. –№2(12). – с.15-25.

6. Горелик Л. А. Методы распознавания / Л.А. Горелик В. А. Скрипкин. – М.: Высшая школа, 1989. – 230 с.

7. Кожем'яко В. П., Ліщинська Л. Б., Понура О. І. Особливості формування систем ознак при розпізнаванні зображень в системах технічного зору / В.П. Кожем'яко, Л.Б. Ліщинська, О.І. Понура // Вісник ВПІ. – 1998. – № 1. – С.47–56.

8. Кориков А. М. Корреляционные зрительные системы роботов / А. М. Кориков, В.И. Сырякин.– Томск: Радио и связь, 1990. – 184 с.

9. Гупал А.М. Оптимальные процедуры распознавания / А.М. Гупал, И.В. Сергиенко // Кибернетика и системный анализ. – 2003. – №1. – С.21 – 29.

10. Меденников П.А., Павлов Н.И. Адаптивный алгоритм и система признакового распознавания / П.А. Меденников, Н.И. Павлов // Оптический журнал. – 2000. – №1. С. 46–51.

11. Кожемяко В.П. Квазиимпульсно-потенциальные оптоэлектронные элементы и устройства логико-временного типа / С.В. Свечников, В.П. Кожемяко, Л.И. Тимченко. – Киев: Наукова Думка, 1987. – 256 с.

12. Кожемяко В.П. Оптоэлектронные логико-временные информационно-вычислительные среды / В.П. Кожемяко. – Тбилиси: Ганатлеба, 1984. – 357 с.

13. Кожемяко В.П. Организация оптоэлектронных некогерентных процессоров ЦВМ / В.П. Кожемяко, О.Г.Натрошвили, Д.О.Саникидзе. – Тбилиси: Ганатлеба, 1989. – 345 с.

14. Сачанюк-Кавецька Н.В. Елементи око-процесорної обробки зображень у логіко-часовому середовищі / Н.В. Сачанюк-Кавецька // Вісник ВПІ. – 2004. – № 6. – С. 58-63.

15. Кожемяко В.П., Мартинюк Т.Б., Кожем'яко А.В. Класифікаційна модель створення прототипу образного комп'ютера // Оптико-електронні інформаційно-енергетичні технології: – 2006. – №2(12). – С. 129-141.

16. Кожем'яко В.П. Паралельно-ієрархічні мережі як структурно-функціональний базис для побудови спеціалізованих моделей образного комп'ютера : [Монографія] / В.П. Кожем'яко, Л.І. Тимченко, А.А. Яровий. – Вінниця: УНІВЕРСУМ, 2005. – 160 с.

17. Методологічні аспекти принципів паралельності та ієрархічності в нейронній обробці інформації / В.П. Кожем'яко, Л.І. Тимченко, А.А. Яровий, Р.М. Новицький // Оптико-електронні інформаційно-енергетичні технології. – 2006. – № 2(12). – С. 98-109.

18. Паралельно-ієрархічне перетворення і Q-обробка інформації для систем реального часу : [Монографія] / М.О. Ковзель, Л.І. Тимченко, Ю.Ф. Кутаєв та ін. – Київ: „КУЕТТ”, 2006. – 492 с.

19. Паралельно-ієрархічне перетворення як системна модель оптико-електронних засобів штучного інтелекту : [Монографія] / В.П. Кожем'яко, Ю.Ф. Кутаєв, С.В. Свечніков, Л.І. Тимченко, А.А. Яровий. – Вінниця: УНІВЕРСУМ-Вінниця, 2003. – 324 с.

20. Око-процесорне розпізнавання образів за ознакою осі орієнтації для геоінформаційно-енергетичної системи / В.П. Кожем'яко, Р.М. Новицький, В.І. Маліновський, Я.М. Бондарчук // Штучний інтелект. – 2009. – №3. – С.556-563.

Відомості про авторів

Кожем'яко Володимир Прокопович – академік АІНУ, д.т.н., професор, завідуючий кафедрою лазерної і оптоелектронної техніки Вінницького національного технічного університету, Хмельницьке шосе 95, м Вінниця, 21021, e-mail: oeipt@ukr.net.

Новицький Руслан Михайлович – пошукач кафедри лазерної і оптоелектронної техніки, Вінницького національного технічного університету, Хмельницьке шосе 95, м Вінниця, 21021, e-mail: miha-lych_13@mail.ru.

Кожем'яко Костянтин Володимирович – пошукач кафедри лазерної і оптоелектронної техніки Вінницького національного технічного університету, Хмельницьке шосе 95, м Вінниця, 21021.

УДК 004.25; 004.27: 004.382.2

Ю. С. ЯКОВЛЕВ

Институт кибернетики имени В.М. Глушкова НАН Украины, Киев

НОВЫЕ ПАРАДИГМЫ ПОСТРОЕНИЯ КОМПЬЮТЕРНЫХ СИСТЕМ

Аннотация. Предложен фрагмент схемы онтологии перехода к новым парадигмам построения компьютерных систем, в которой выявлена взаимосвязь новых парадигм линий развития архитектур, технологии создания элементной базы компьютерных систем и парадигм развития информационных технологий. Кратко рассмотрены основные направления развития архитектур ЭВМ и элементной базы, принципы работы которой основаны на различных физических процессах и явлениях в областях нанотехнологий.

Ключевые слова: онтология, вычислительная парадигма, нанотехнологии, архитектура компьютерных систем, элементная база.

Анотація. Запропоновано фрагмент схеми онтології переходу до нових парадигм побудови комп'ютерних систем, в якій виявлено взаємозв'язок нових парадигм ліній розвитку архітектури, технології створення елементної бази комп'ютерних систем і парадигми розвитку інформаційних технологій. Стисло розглянуті основні напрями розвитку архітектури ЕОМ і елементної бази, принципи роботи якої засновані на різних фізичних процесах і явищах в областях нанотехнологій.

Ключові слова: онтологія, обчислювальна парадигма, нанотехнології, архітектура комп'ютерних систем, елементна база.

The summary. The scheme fragment ontology transition to new paradigms of construction of computer systems, in which the interrelation of new paradigms of lines of development of architecture is revealed, is offered technology of creation of element base of computer systems and a paradigm of development of information technologies. The basic directions of development of architecture of the computer and the element base which principles of job are based on various physical processes and the phenomena in areas nanotechnology are is short considered.

Keywords: ontology, a computing paradigm, nanotechnology, architecture of computer systems, element base.

Введение

Определяющими факторами развития информационных технологий, являются социально-экономические потребности общества. На каждом этапе развития человеческого общества рост трудоемкости, сложности и количества задач, подлежащих решению с помощью средств вычислительной техники, как правило, опережают возможности имеющейся средств, что приводит к необходимости поиска новых архитектурно-структурных решений и новых парадигм организации вычислительного процесса. К таким задачам относятся задачи государственного стратегического значения, необходимость решения которых в текущий момент времени не вызывает сомнений.

Существенный вклад в развитие вычислительных средств обычно вносили технологические решения, при этом основополагающей характеристикой нового поколения вычислительных систем являлась элементная база, так как переход на новую элементную базу всегда способствовал улучшению показателей производительности и надежности вычислительных систем.

Актуальность

В настоящее время практически во всех средствах вычислительной техники используется твердотельная элементная база, выполненная в виде больших интегральных схем (БИС), технологический размер элементов при создании которых достиг порядка 0,1 мкм за счет использования современных технологических решений. Однако дальнейшая миниатюризация элементов ограничена физическим барьером. Проектная норма 0,05 – 0,1 мкм (50-100 нм) - это нижний предел твердотельной микроэлектроники, основанной на классических принципах синтеза схем, так как при меньших размерах начинают проявляться эффекты квантовой связи, в результате чего твердотельное устройство становится системой, действие которой основано на коллективных электронных процессах [1 – 3]. Современные полупроводниковые компьютеры скоро исчерпают свой потенциал, и даже при условии перехода к трехмерной архитектуре микросхем их быстродействие будет ограничено (преполагается до значения 1015 операций в секунду). Независимо от указанных выше факторов, а также от уровня потребляемой мощности вычислительной техникой, выполненной на базе кремния, и достигнутого технологического предела уменьшения размеров элементов на кристалле БИС, кремниевая элементная база все еще может быть использована в течение нескольких лет. В предверии окончания кремниевой эпохи, весьма актуальной становится проблема поиска новых парадигм построения элементной базы, вычислительных средств и систем на её основе. В этом плане проводятся активные исследования в области нанотехнологий в следующих направлениях: молекулярная электроника; биохимические и органические элементы; квазимеханические элементы на основе нанотрубок; квантовые компьютеры, нейрокомпьютеры и др.

Цель

На основе анализа современного состояния и перспектив развития компьютерных систем, элементной базы и информационных технологий определить их взаимосвязь и стратегию перехода к новым вычислительным парадигмам, обеспечивающим существенное повышение производительности компьютерных систем за счет перехода от микротехнологий к нанотехнологиям при создании элементной базы и компьютерной системы в целом.

Постановка задач

1. Исследовать перспективы развития архитектур современных ЭВМ и компьютерных систем.
2. Исследовать особенности элементной базы, принципы работы которой основаны на
3. применении молекулярной электроники, биохимических и органических элементов,
4. квантовых процессов, нейросетей и др. Определить научно-технические проблемы, которые
5. необходимо решить при создании нанокomпьютеров на перспективной элементной базе.
6. Определить взаимосвязь между процессами поиска новых парадигм создания перспективной элементной базы и организации перспективных архитектур ЭВМ и систем, используя при этом онтологический подход.

Решение задачи

Нанотехнологии – это технологии, оперирующие величинами порядка нанометра. Это технологии манипуляции отдельными атомами и молекулами, в результате которых создаются структуры сложных спецификаций. Переход от “микро” к “нано” – это уже не количественный, а качественный переход, скачок от манипуляции веществом к манипуляции отдельными атомами.

На рис.1 представлен фрагмент схемы онтологии поиска новых парадигм для обеспечения требуемых характеристик ВТ, необходимых для решения пользовательской задачи. При этом символы “О”, “&”, вписанные в малые кружки, соответственно означают: “&” – реализация функции “И” между всеми объектами (событиями) в точке их соединения или переход одновременно ко всем объектам (событиям) в точке их разъединения; символ “О” обозначает функцию “ИЛИ” при переходе от общей линии в точке их разъединения одновременно к одному или нескольким объектам (событиям) или в точке соединения объектов – переход от одного или нескольких объектов к общей линии.

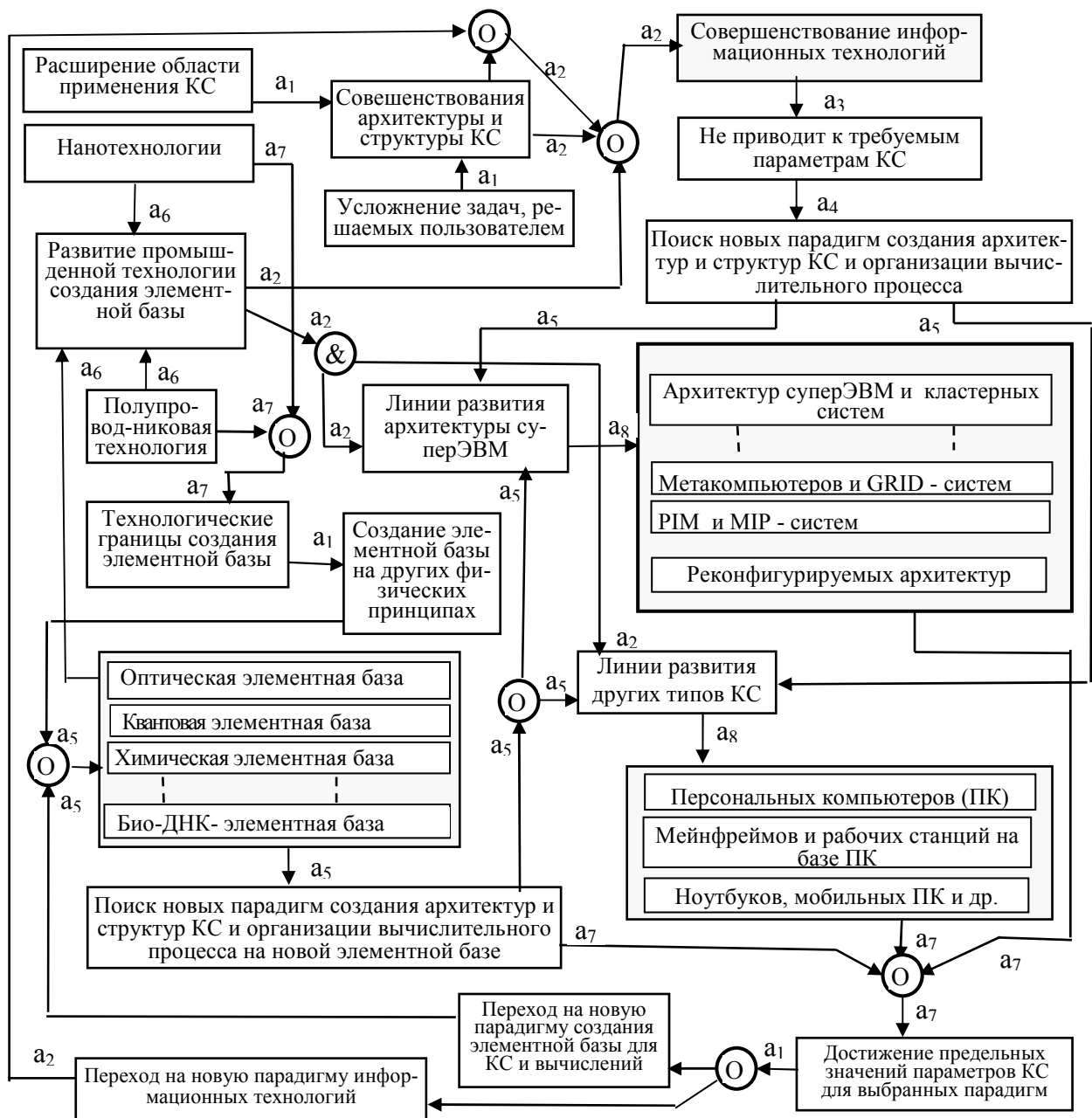
Если же при выбранной парадигме организации вычислительного процесса не удастся обеспечить требуемые для решения задачи характеристики ВТ, то переходят на поиск новой парадигмы создания элементной базы, отличающейся от полупроводниковой другими физическими принципами и явлениями. Речь идет об использовании особенностей оптических, квантовых, химических, биологических и других явлений, приспособляемых для организации вычислительных процессов.

Если для выбранной новой парадигмы создания элементной базы все научно-технические проблемы создания ВТ решены, то так же, как и в предыдущем случае, анализируют линии развития средств, выполненных уже на новой элементной базе, с целью определения соответствующей парадигмы архитектурно - структурной организации, которая могла бы обеспечить требуемые технические характеристики ВТ. Если и при этом выходят на предельные значения параметров ВТ, которые не удовлетворяют требованиям со стороны задач пользователя, то изменяют парадигму информационной технологии, и точно также проходят цикл анализа линий развития ВТ, но уже с учетом новой элементной базы, пока не будут обеспечены необходимые технические параметры.

Анализ архитектур различных классов показал, что основные перспективные научно-технические решения, как правило, апробируют при создании супер-ЭВМ и, спустя некоторое время, затем их переносят на другие классы ЭВМ и систем. Например, персональные компьютеры (ПК) в настоящее время имеют те же характеристики, что и супер-ЭВМ примерно 13 лет назад. С другой стороны, переход в область нанотехнологий позволяет не только повысить производительность супер-ЭВМ, но и существенно уменьшить её размеры, что как раз согласуется с тенденцией развития ПК. Поэтому, в данной работе мы остановимся на линии развития супер-ЭВМ и ПК, выделяя при этом особенности построения элементной базы, основанной на новой парадигме, которая и по назначению, и по своим особым свойствам как раз и соответствует линии развития этих классов машин. Что касается линий развития машин других классов (серверов, майнфреймов и др.), то они обычно строятся на базе ПК или супер-ЭВМ, и поэтому здесь не рассматриваются.

Линии развития ЭВ и систем. Можно выделить следующие основные линии их развития: кластерные системы; метакомпьютеры и GRID-системы; персональные ЭВМ, а также мобильные компьютеры типа ноутбуков и карманных персональных компьютеров; супер-ЭВМ, основанные на использовании новой машинной алгебры (например, алгебры матриц, рядов Фурье и др.); знание-ориентированные и онтологоуправляемые компьютерные системы; распределенные компьютерные системы типа “Процессор-в-памяти” (“Processor-in-Memory”(PIM-системы)), “Память-в-процессоре (“Memory-in-processor” (MIP - системы)); линии развития образного компьютера; искусственного интеллекта; реконфигурируемых компьютерных систем; линия развития специализированных средств обработки информации и др. [4]

Достигнутая степень интеграции позволяет строить параллельные системы, в которых число процессоров может достигать десятки и сотни тысяч. Отмечено, что в области повышения производительности вычислительных систем резерв технологических решений ограничивается одним порядком. Освоение же массового параллелизма и новых архитектурных решений содержит резерв повышения производительности на несколько порядков.



Обозначения: a_1 – определяет; a_2 – обеспечивает; a_3 – не приводит к требуемым результатам; a_4 – является причиной; a_5 – определяет объект поиска; a_6 – входит в состав; a_7 – достигает; a_8 – включает.

Рисунок 1 – Фрагмент схемы онтологии перехода к новым парадигмам при создании КС

Одним из факторов, влияющих на архитектуру высокопроизводительных вычислительных систем, является взаимозависимость архитектуры и алгоритмов решаемых задач. Такая взаимозависимость выступает стимулом для поиска алгоритмов, наилучшим образом соответствующих возможным формам параллелизма на уровне аппаратуры.

Одним из способов дальнейшего повышения производительности вычислительной системы является объединение суперкомпьютеров в кластеры при помощи оптоволоконных соединений.

Линия развития персональных компьютеров. Персональный компьютер (ПК) представляет собой наиболее развитый вид микропроцессорных систем. Применительно к ПК можно выделить следующие семейства, отличающиеся, в основном, назначением использования: офисные (настольные), рабочие станции, майнфреймы, серверы и мобильные (переносные) типа ноутбук, карманных ПК (КПК).

В среднем по уровню производительности сегодняшние профессиональные ПК практически соответствуют суперкомпьютерам 13-летней давности. Именно поэтому исследование рынка высокопроизводительных вычислений — хороший способ оценить направление развития массовых компьютеров будущего.

Сейчас происходит трансформация парадигмы компьютерных вычислений: раньше ее центром был компьютер, а теперь им становится человек, при этом вычисления становятся по-настоящему персональными. Пользователя сегодня окружают самые разные электронные устройства и приложения. Ему необходимо обеспечить доступ к нужным сервисам и приложениям везде, в любой момент времени и с любых устройств.

Существенные моменты новой компьютерной парадигмы — это вовлечение пользователей в работу с компьютером, делая ее простой и удобной; согласованная работа различных формфакторов и гарантированная защита всех устройств. Возрастает спрос на интернет-планшеты. Они меньше ноутбука, обеспечивают функции сотового телефона, Wi-Fi-доступ, имеют полноцветный сенсорный экран, а также обширный набор приложений и смогут удовлетворить потребности даже самого привередливого пользователя. Беспроводные домашние сети окажутся в выигрышном положении: для быстрого просмотра интернет-ресурсов, интернет-планшеты дома могут оказаться даже удобнее, чем телефон или компьютер.

Мнение других утверждает, что будущее ПК лежит, прежде всего, в расширении его коммуникационных возможностей. А поскольку они ассоциируются с беспроводной связью и мобильностью, то конечным результатом трансформаций настольного компьютера может стать и такое устройство, как «портативный мобильный коммуникатор», которое может комбинировать в себе бытовые электронные приборы: КПК и беспроводной телефон, видео-конференц-связь и аудио MP3-плеер, терминал для Интернет-доступа и электронной почты, цифровую камеру и др.

Сегодня карманные компьютеры именуется суперпортативными компьютерами или персональными электронными помощниками (Personal Digital Assistant, PDA). PDA имеют существенные положительные свойства, которые в будущем можно будет использовать. Среди них следует выделить [5]: наличие практически во всех современных карманных компьютерах цветного экрана; достаточно высокая для такого класса компьютеров производительность; приближение клавиатуры к оптимальному для быстрого ввода размеру, распознавание рукописного ввода, а также возможность рисовать на экране "электронными чернилами"; возможность поддержки доступа в локальную и глобальную сеть через сетевые карты Ethernet, что дает возможность использовать высокоскоростной доступ к информационной сети; поддержка некоторыми моделями карманных компьютеров универсального порта USB, что позволяет подключать к КПК многие периферийные устройства.

Прогнозируя обозримое будущее персональных компьютеров, невозможно обойти вниманием и тот многообещающий потенциал, что несут в себе биомолекулярные, квантовые и другие вычислители, построенные на новых физических принципах и явлениях.

Переход на новую парадигму создания элементной базы для перспективных компьютерных систем. Одновременно с совершенствованием полупроводниковой технологии, которая все же будет занимать доминирующее положение еще несколько десятков лет, выполняется комплекс научных исследований, инженерно-технических и технологических решений в области создания элементной базы, использующей (в отличие от полупроводников) другие физические среды, процессы и явления: оптические, квантовые, биологические, химические и др. Основные свойства таких компьютеров приведены в таблице 1 [1 – 3].

Молекулярная схемотехника и молекулярные компьютеры. Это устройство, в котором вместо кремниевых чипов, применяемых в современных компьютерах, работают молекулы и молекулярные ансамбли. На основе уникальных свойств органических молекул разработано множество вариантов схем для гипотетического молекулярного компьютера. В качестве триггеров удобнее всего использовать молекулы, которые обладают одинаковой молекулярной массой и составом, но различаются строением или расположением атомов в пространстве. Некоторые из них можно переводить из одной формы в другую путем внешнего воздействия, например, с помощью ультрафиолетового излучения либо с помощью света видимого диапазона. На основе такого триггера можно строить как устройства оперативной памяти, так и элементы, выполняющие логические функции.

Набор необходимых базовых элементов молекулярного компьютера можно объединить в единый вычислительный комплекс с помощью методов химического синтеза, химически соединяя "выходы" одних логических элементов с "входами" других. Теоретически можно соединить отдельные молекулярные компоненты "проводами", например, из так называемых углеродных нанотрубок — цилиндрических структур диаметром несколько нанометров — или из токопроводящих полимеров, называемых иногда «органическими металлами».

Таблица 1 – Основные свойства нанокomпьютеров

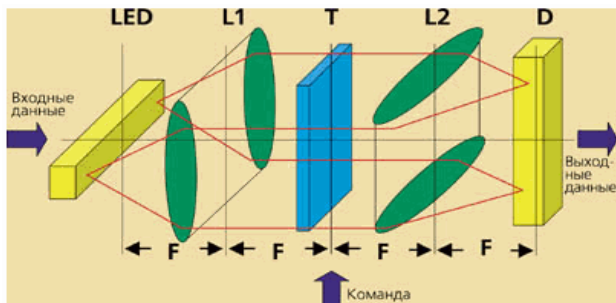
Тип компю-тера	Носитель информа-ции:	Основной принцип организации вычислительного процесса
Молекуляр-ные компью-теры	молекулы и молеку-лярные ансамбли	Это наноразмерная двухбитовая система, воспроизводящая на молекулярном уровне функцию триггера. Переключение моле-кулы из одного состояния в другое осуществляют с помощью электрического, магнитного полей, света, химических агентов и др.
Квантовые компьютеры	квантовые частицы, образующие кванто-вый бит – qubit (ку-бит),	Кубит имеет два базовых состояния – 0 или 1 и состояние суперпо-зиции. Кубиты формируют в полупроводниковых структурах, в органических веществах и др. Смена состояний – согласно зако-нам квантовой механики под действием электромагнитного поля.
Оптические компьютеры	оптическое излуче-ние на всех этапах обработки и переда-чи информации	Два состояния световой волны, отличающиеся амплитудой и (или) параметрами поляризации, соответствуют изменению интенсивно-сти луча света, прошедшего через вещество: увеличение интен-сивности падающего луча приводит к резкому возрастанию интен-сивности выходного луча; при уменьшении – интенсивность про-шедшего луча резко падает.
Нейроком-пьютеры	искусственные ней-роны (нейро-чипы) и нейро-ноподобные связи	Алгоритм решения задачи реализуют с помощью нейронных сетей, моделирование которых выполняют как программно на ЭВМ, так и программно-аппаратно на цифровых СБИС. Обладают свойствами параллельной обработки информации, способностью к обучению и высокой помехо - и отказоустойчивостью.
ДНК-компьютеры	нуклеотиды ДНК, выстраиваемые в строго определен-ной последователь-ности. Машинный код представляют с помощью различных состояний этих нук-леотидов	Информация, закодированная на ДНК, используется для выполнения молекулярных процессов при построении компьютеров. Особен-ностью схем являются масштабируемость и автономность само-стоятельной сборки процессов, связанных с наномасштабом моле-кул ДНК. Метод ДНК позволяет одновременно сгенерировать все возможные варианты решений с помощью известных биохимиче-ских реакций. Далее необходимо выделить молекулу-нить, в кото-рой закодирован нужный ответ.
Химические компьютеры	индивидуальные «клетки», состоящие из липидов, искус-ственно внедренных в органическую среду и способные пере-носить химические сигналы внутри мо-лекул.	Данные кодируются химическими свойствами веществ (концен-трация, цвет, кристаллическая структура и т. д.), а вычислитель-ные процессы протекают в виде химических реакций, в результа-те которых периодически изменяются параметры – концентрация, температура, цвет и другие, что позволяет клеткам «общаться» между собой и передавать информацию. Работает «химический компьютер» от электричества, но все изменения происходят вследствие химических процессов.

Квантовые компьютеры работают на основе квантовой логики. Свойство квантовых частиц одно-временно пребывать в разных состояниях, именуют квантовым параллелизмом, который и используется в квантовых вычислениях. Для квантовых компьютеров, как и для классических, были введены элемен-тарные логические операции: дизъюнкция, конъюнкция и квантовое отрицание, при помощи которых организована вся логика квантового компьютера.

Квантовые компьютеры могут быть созданы путем синтеза новых идей в математике, физике, ин-форматике и нанотехнологии.

Оптические компьютеры. Оптическая бистабильность - это одно из проявлений взаимодействия света с веществом в нелинейных системах с обратной связью, при котором определенной интенсивности и поляризации падающего на вещество излучения соответствуют два (аналог 0 и 1 в полупроводниковых системах) возможных стационарных состояния световой волны, прошедшей через вещество, отличающихся амплитудой и (или) параметрами поляризации. Набор полностью оптических логических уст-ройств для синтеза более сложных блоков оптических компьютеров реализуется на основе пассивных нелинейных резонаторов-интерферометров. В зависимости от начальных условий (начального положен-ния пика пропускания и начальной интенсивности оптического излучения) в пассивном нелинейном резонаторе, нелинейный процесс завершается установлением одного из двух устойчивых состояний про-пускания падающего излучения. А из нескольких нелинейных резонаторов можно собрать любой, более сложный логический элемент (триггер). Элементы памяти оптического компьютера представляют собой полупроводниковые нелинейные оптические интерферометры, в основном, созданными из арсенида галлия (GaAs). Минимальный размер оптического элемента памяти определяется минимально необходи-мым числом атомов, для которого устойчиво наблюдается оптическая бистабильность. Это число состав-ляет ~1000 атомов, что соответствует 1-10 нанометрам.

К настоящему времени уже созданы и оптимизированы отдельные составляющие оптических компьютеров – оптические процессоры, ячейки памяти, однако до полной сборки еще далеко. На рис.2 показан пример построения оптического функционального блока для реализации матричного преобразования входного вектора-строки в выходной вектор-столбец.



Обозначения на схеме : LED - линейка световых излучающих диодов; L1-цилиндрическая линза; T - оптический транспарант с записанной на нем матрицей пропускания $T(i, j)$; L2 – цилиндрическая линза, образующая которой параллельна столбцам матрицы транспаранта; D -линейный фотоприемник

Рисунок 2 – Упрощенная схема оптического функционального блока для реализации матричного преобразования входного вектора-строки в выходной вектор-столбец

В нейрокомпьютерах алгоритм решения задачи представлен логической сетью элементов - нейронов. На рис.3 представлен биологический нейрон, а на рис.4 – его модель. Нейроны имеют состояния 0, 1 и пороговую логику перехода из состояния в состояние. Каждый нейрон в сети определяет взвешенную сумму состояний всех других нейронов и сравнивает ее с порогом, чтобы определить свое собственное состояние.

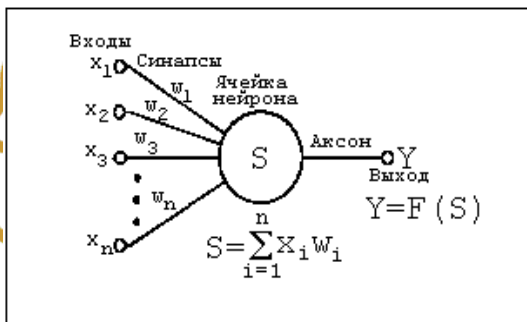
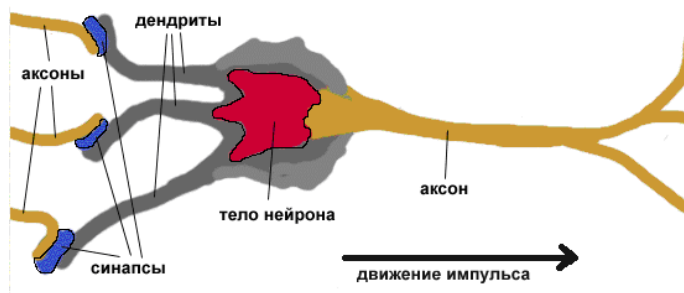


Рисунок 3 – Биологический нейрон

Рисунок 4 – Математическая модель нейрона

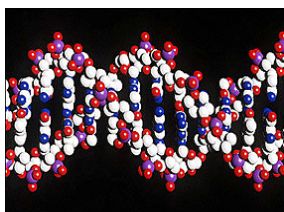


Рисунок 5 – Модель молекулы ДНК

Как следствие этого введены специфические связи между элементами. В отличие от классических КС нейрокомпьютеры реализуют алгоритмы решения задач, представленные в виде нейронных сетей. Такие сети функционально ориентированы на конкретный алгоритм решения конкретной задачи, поэтому для решения задач разного типа требуется нейронная сеть разной топологии. Нейрокомпьютеры можно строить на базе нейрочипов. Возможна эмуляция нейрокомпьютеров (моделирование) - как программно на ЭВМ, так и программно-аппаратно на цифровых супербольших интегральных схемах. Достоинства нейрокомпьютеров - параллельная обработка информации и способность к обучению *Био-ДНК-компьютеры* [1 - 3]. На рис.5 показана модель молекулы ДНК.

Ученым удалось реализовать логические вентили, выполняющие различные логические функции, в частности функцию исключающего ИЛИ. Ключевыми элементами выступали комплементарные нити ДНК – то есть нити определенных последовательностей, которые образуют друг с другом прочное соединение. Например, биологическая информация кодируется нуклеотидами аденином, тимином, цитозином и гуанином (А, Т, С, G), выстраиваемыми в строго определенной последовательности. Идея ДНК-компьютера заключается в представлении двоичного машинного кода как раз с помощью различных состояний этих нуклеотидов. Теоретически развитие этой работы может привести к созданию компьютера из ДНК, способного выполнять элементарные арифметические операции — сложение и вычитание. Исключительное свойство ДНК — способность одиночных цепей к «взаимоузнаванию». Привлекательное для компью-

терной области свойство ДНК – невероятная плотность хранимой информации и низкая потребляемая мощность, а значит перегрев биокомпьютерам не грозит. Некоторые из особенностей такой вычислительной схемы являются масштабируемость и автономность самостоятельной сборки процессов, связанных с нано-масштабом молекул ДНК. Метод ДНК позволяет сразу сгенерировать все возможные варианты решений с помощью известных биохимических реакций. Затем возможно отфильтровать именно ту молекулу-нить, в которой закодирован нужный ответ.

Химические компьютеры. Для создания химического компьютера потребуются 2 платформы. Первая – «индивидуальные клетки» – органические вычислительные системы, состоящие из липидов, способных переносить сигналы внутри молекул. Вторая платформа – внутриклеточный состав, который проводит химические реакции, протекающие в колебательном режиме. В результате этих реакций периодически изменяются параметры – концентрация, температура, цвет и другие, что позволяет клеткам «общаться» между собой и передавать определенную информацию. Работать «химический компьютер» будет от электричества, но все изменения будут происходить вследствие химических процессов.

Схемотехника и архитектура нанокomпьютеров. Примеры первых наиболее успешных экспериментов показали, что основу вычислительной среды будущего нанокomпьютера будет составлять регулярная, для начала – двумерная, матрица, образованная нанотранзисторами размером 2 – 10 нм. [6]. При этом молекулярно-кластерными методами можно будет создавать наиболее мелкие элементы схем, требующие высокого пространственного разрешения порядка 0,5 – 1 нм, недоступного для литографии.

Несмотря на то, что основные рабочие элементы разрабатываемых нанотранзисторов имеют некремниевую основу, уже имеется проработка технологии их изготовления с системной интеграцией на кремниевой подложке. В частности, базовый кремниевый кристалл может быть использован для создания интерфейсного обрамления наночипа в стандарте TTL. К тому времени, когда рынок электроники будет наполнен устройствами мезоэлектроники с разрешением 20–30 нм (примерно через десять лет), должны появиться первые экспериментальные образцы универсальных программируемых молекулярно-кластерных и спинтронных чипов с кремниевым интерфейсным TTL - обрамлением. Все это выглядит вполне реальным, так как базисные логические функции типа ИЛИ-НЕ на основе углеродных нанотрубок уже изготовлены и испытаны.

В свою очередь, на их основе можно будет создавать нанопроцессоры, наночипы памяти и полнофункциональные однокристалльные нанокomпьютеры. По своему разнообразию мир нанокomпьютеров будет необычайно широк [7 – 10]. Нанокomпьютеры минимального размера в несколько микрон смогут содержать сотни тысяч транзисторов. Однокристалльные нанокomпьютерные гиганты с размером кристалла порядка дюйма будут содержать уже триллионы транзисторов. Для обеспечения их работы на предельной частоте порядка 1000 ГГц понадобятся специальные меры по снижению ландауэровского тепловыделения.

Среди разрабатываемых в настоящее время нейроБИС выделяются модели фирмы Adaptive Solutions (США) и Hitachi (Япония). НейроБИС фирмы Adaptive Solutions, вероятно, станет одной из самых быстродействующих: объявленная скорость обработки составляет 1,2 млрд. соединений/с. (НС содержит 64 нейрона и 262144 синапса). НейроБИС фирмы Hitachi позволяет реализовать НС, содержащую до 576 нейронов. Эти нейроБИС, несомненно, станут основой новых нейрокомпьютеров и специализированных многопроцессорных изделий.

На сегодняшний день предложены варианты конструкции квантовых компьютеров на основе квантовых точек и сверхпроводящих переходов. Более того, есть заявления коммерческих организаций о создании реальных процессоров на базе кубитов, например, компьютер «Ogion». Он основан на микросхеме, содержащей 16 «кубитов».

Выполнение унитарных логических операций над элементами в квантовых компьютерах предполагается осуществлять с помощью соответствующих внешних воздействий, которыми управляют классические компьютеры.

Уже предложена новая архитектура будущего суперкомпьютера - гибридно-технологическая многопоточная архитектура (Hybrid Technology Multithreaded Architecture, НТМТ), в основе которой лежит использование нетрадиционных технологий ("гибридность") и расщепление параллельных процессов на более мелкие независимые фрагменты: потоки и нити ("многопоточность"). За новым петафлопсным суперкомпьютером закрепился термин "гиперкомпьютер" [11]. Важнейшей коммуникационной артерией гиперкомпьютера является широкополосная оптическая многоступенчатая пакетная сеть. Эта сеть, соединяет друг с другом все банки динамической и статической полупроводниковой памяти. Еще одна оптическая компонента гиперкомпьютера - голографическая память. Один кристалл голографической памяти будет вмещать 1 Гигабайт информации, которая может быть считана или записана в виде одномегабайтных страниц.

Основные проблемы создания нанокomпьютеров. Если выделить эти проблемы в качестве доминирующих по различным видам нанокomпьютеров, то получим следующую картину.

Если рассматривать оптическую связь между элементами, то в системе одна молекула - один фотон надежность переключения будет невелика из-за относительно малой вероятности перехода молекулы в возбужденное состояние. Кроме того, элемент должен однозначно переключаться в требуемое состояние и находиться в нем достаточно долго – до следующего воздействия. Для сравнительно простых молекул это требование, как правило, не выполняется: если переходом в возбужденное состояние можно управлять, то обратный переход может происходить спонтанно.

В квантовых компьютерах существуют проблемы, связанные с неоднозначностью интерпретации информации, поэтому заявления о создании квантовых компьютеров ставятся в научном сообществе под сомнение. Существующие квантовые системы еще не способны обеспечить надежные вычисления, так как они либо недостаточно управляемы, либо очень подвержены влиянию шумов. Существует также проблема ввода-вывода информации. Необходимость связи компьютера с внешним миром, а также отдельных его частей друг с другом может приводить к существенным потерям в скорости вычислений.

Основной проблемой, стоящей перед учеными, является синхронизация работы отдельных элементов оптического компьютера в единой системе, поскольку существующие элементы характеризуются различными параметрами рабочей волны светового излучения (интенсивность, длина волны), и уменьшение размера компьютера. Имеются и другие факторы, препятствующие оптической технологии построения компьютеров. Наиболее важными из них являются: каскадирование - интегрирование большого числа оптических вентиляей.

Ошибка считывания и копирования информации на ДНК в 1000 раз превышает ошибку магнитных накопителей. Тем не менее, ученые смотрят на ДНК-компьютеры с большим оптимизмом, нежели на квантовые, ведь при производстве некоторых узкоспециализированных расчетов они будут незаменимы.

Объемная плотность транзисторов в разрабатываемых интегральных наносхемах предельно высока. В таких условиях вопросы энергетики перспективного нанокomпьютера оказываются чрезвычайно важными. Существует фундаментальное ограничение плотности упаковки логических элементов, связанное уже не с атомной структурой вещества, а с термодинамикой вычислительного процесса как такового. В настоящее время просматриваются различные пути решения проблемы перегрева процессора.

Следует упомянуть о радиационной опасности, грозящей нанокomпьютерам со стороны обычных материалов, используемых в электронике. Дело в том, что в числе незначительных примесей, всегда присутствующих даже в самых чистых материалах, есть радиоактивные элементы. Особую опасность представляют альфа-активные изотопы тория. Одна альфа-частица с типовой энергией 1 МэВ даже в условиях обычной микроэлектроники при попадании в кристалл способна освободить из связанного состояния миллионы электронов. С этим борются, применяя помехоустойчивое кодирование.

Наночастицы имеют одно свойство, которое очень мешает их использованию. Они могут образовывать агломераты, то есть, слипаться друг с другом. Эту проблему пытаются решить, используя вещества – дисперсанты.

Поиски альтернативных способов изготовления нанотранзисторов и сборки из них компьютеров составляют еще одно важное направление современных исследований в области нанотехнологий.

Проблема финансирования. Борьба с накладываемыми на вычислительный процесс фундаментальными ограничениями – дело чрезвычайно сложное и дорогостоящее, доступное лишь высокоразвитым странам в рамках крупных национальных программ.

Архитектурно-структурные методы повышения производительности КС. Технологические приемы, положенные в основу реализации любой парадигмы обеспечения высокой производительности КС, по различным причинам достигнут своего технологического предела. В настоящее время это особенно наглядно проявляется при создании полупроводниковых БИС, технологический предел для которых может быть достигнут в ближайшее время (10 – 15 лет) при размерах полупроводниковых элементов менее 0.1мкм. В этот период времени, вряд ли уже будут созданы полноценные суперЭВМ, выполненные в соответствии с другими парадигмами организации вычислительного процесса: квантовой, нейронной, химической и др. Анализ зарубежных источников показал, что такие компьютеры, появятся не менее, чем через 50 – 100 лет, хотя уже имеются определенные успехи в части создания отдельных узлов и блоков. Появившийся промежуток времени между технологическими пределами полупроводниковой техники и нанокomпьютерами, выполненными на другой физической основе, может быть заполнен компьютерными системами, выполненными на полупроводниковой элементной базе. При этом требуемое повышение производительности может быть реализовано новыми архитектурными решениями по всем указанным выше линиями их развития, для которых уже на настоящее время существуют определенные заделы. [12 – 14].

Архитектура для вычислений пространственной решетки отображает непосредственно на кристалле БИС, используемого в качестве узла больших систем, особенности алгоритмов моделирования, двумерной и трехмерной обработки изображения и представления, и других трудоемких задач обработки данных.

Архитектура КС типа “Memory-in-Processor” (“Память-в-процессоре”) или МІР-система, которая обеспечивает логическую и физическую интеграции памяти в функциональных блоках микропроцессора. МІР – системы и содержит блоки обработки с высокими функциональными возможностями (HLF – Higher Layer Function), каждый из которых связан со своим функциональным блоком и может работать на 128-битовых данных. МІР спроектирована на одном кристалле с “компилятором-на-кремнии”, генерирующим команды для модулей HLF.

Архитектура КС типа “Процессор-в-памяти” (“Processor-in-memory”) или РІМ-система, которая отличается от КС с классической архитектурой широкой полосой пропускания по каналам процессор-память, низкой потребляемой мощностью, невысокими габаритами и весом за счет размещения на одном кристалле большого количества банков памяти, присоединённых к ним средств обработки (до нескольких тысяч) и обеспечивающих считывание из этих банков памяти и одновременную обработку нескольких сот и даже тысяч операндов.

Архитектура типа “Циклоп” – архитектура для высокоэффективных параллельных компьютеров, которая основана на трех основных принципах: интеграция логики обработки и памяти на одной подложке кремния; использование массового параллелизма внутри чипа; ячеистый подход к построению больших систем. Большие системы с тысячами чипов могут быть построены, применяя чип Циклопа как стандартный узел распределенной КС.

Архитектура с реконфигурируемым информационным каналом использует методы реконфигурации архитектуры и динамического перераспределения ресурсов с помощью среды реконфигурации, например, программируемых логических схем (ПЛИС) и соответствующих средств их программирования пользователем.

Реконфигурируемая архитектура системы памяти использует блок коммутации информации, с помощью которого осуществляется настройка архитектуры БИС системы памяти на конкретный класс задач. Таким образом, множество устройств памяти имеет соответствующее множество связанных процессоров, и все они вместе размещены на одной БИС.

Архитектура типа “Импульс направлена на ускорение доступа к ячейкам памяти путем применения более совершенного контроллера памяти с расширенными функциональными возможностями, не требуя изменений в организации процессора, памяти типа КЭШ или типа DRAM. Импульс поддерживает предвыборку памяти, которая может скомпенсировать (скрыть) часть задержки доступа к DRAM, а также задержку переадресации адресной трансляции Импульса.

Архитектура псевдо интеллектуальной памяти (СПІМ) ориентирована на интенсивное использование данных, при решении рекурсивных и высокопараллельных задач, например, задач обработки графических изображений.

Процессорные ускорители (акселераторы) используются как правило в качестве приставки к процессору, существенно повышая его производительность при обработке графических изображений. Кроме того, они могут быть также самостоятельно использованы в качестве процессоров для построения высокопроизводительных компьютерных систем, ориентированных на выполнение задач, содержащих массовые операции, характерные для обработки и отображения графических объектов.

Другие разновидности архитектур различных линий развития.

В институте кибернетики имени В.М. Глушкова НАН Украины разработаны новые типы архитектур компьютерных систем и устройств вычислительной техники, которые по сравнению с известными имеют более высокое быстродействие, широкие функциональные возможности, а также улучшенные значения некоторых других параметров, в частности, – коэффициента загрузки процессоров КС полезной (вычислительной) работой [15 – 20].

Выводы

1. Дальнейший прогресс компьютерной техники, бесспорно, возможен. Он будет двигаться в направлении миниатюризации ЭВМ с одновременным увеличением ее быстродействия. Возможности создания логических элементов микросхем путем манипулирования отдельными атомами, а также новые архитектурно-структурные решения продлят прогресс кремниевых чипов еще, по меньшей мере, на несколько десятилетий.

2. Компьютеры будущего будут основаны на применении главным образом широкого спектра научных дисциплин: молекулярной электроники, молекулярной биологии, робототехники, а также квантовой механики, органической химии и др. Поэтому нанокomпьютеры будут развиваться одновременно по нескольким направлениям, реализующим различные способы представления информации: на основе квантовой логики, классической логики, нейробиологии и др. При этом скорость компьютерных вычислений достигнет значения 10^{51} операций в секунду.

3. С переходом на уровень нанотехнологий станет возможным снижение минимально допустимых размеров компьютера до субклеточного уровня. Способы представления информации в

системах, созданных человеком, почти достигли физических пределов, установленных фундаментальными законами природы.

4. На данном этапе приоритет отдается направлению, в основе которого лежит применение электронных нанотранзисторов (в том числе одноэлектронных). В подобных транзисторах уже реализован квантовомеханический предел передачи информации, налагаемый физическими законами. Достигнут также и предельный уровень тепловыделения. Разработка разнообразных архитектурных вариантов нанокomпьютеров на этой основе ведется очень активно, и технологии будущего уже имеют вполне реальные очертания.

5. Известно, что уже существуют системы обработки информации, построенные на объединении оптических и нейронных компьютеров, - это так называемые нейроно-оптические компьютеры. Можно предположить, что объединение квантовых и нейроно-оптических компьютеров даст миру самую мощную гибридную вычислительную систему.

6. Для создания нано-компьютеров будут необходимы значительные экономические затраты, в несколько десятков раз превышающие затраты на производство современных "классических" полупроводниковых компьютеров.

7. Разнообразие существующих научных разработок в области микроэлектроники, а также обширность накопленных знаний в области других научных дисциплин позволяет надеяться на создание "полноценного суперкомпьютера" на основе нанотехнологий в сроки - примерно 100-150 лет.

Список литературы

1. Beyond silicon: new computing paradigms / Communications of the ACM . – September 2007. – Vol. 50, No. 9. – С. 30 - 41. – Режим доступа: <http://cis.csuohio.edu/~munakata/profess/pdf/907Table.pdf> . – Дата доступа: 15.11.11
2. Рыбалкина М. Нанотехнологии для всех. Большое в малом. Изд-во "Nanotechnology News Network". – 2005. – 444 С. – Режим доступа: http://nano.com.ua/Nanotehnolog_dlya_vseh.pdf. – Дата доступа: 12.12.11.
3. Нанотехнологии – Википедия, – Режим доступа: <http://en.wikipedia.org/wiki/Nanotechnology>. – Дата доступа: 19.12.11.
4. Яковлев Ю.С. Однокристалльные компьютерные системы высокой производительности. Особенности архитектурно-структурной организации и внутренних процессов: монография / Ю.С. Яковлев. – Винница: ВНТУ, 2009. – 294 с.
5. Прогноз тенденций развития сектора высоких технологий на 2010 год . – Режим доступа: http://www.infeconomy.ru/index.php?option=com_content&view=article&id=74:2010-&catid=5:2010-02-06-10-09-02&Itemid=5). – Дата доступа: 15.11.11
6. Минкин В. И. Молекулярные компьютеры / Минкин В. И. . – Минкин В. И. http://wsyachina.narod.ru/technology/molecular_computer.html. – Дата доступа: 15.11.11
7. В европе создают "химический компьютер". –Режим доступа: <http://www.bezpeka.com/ru/news/2010/01/13/chemical-computer.html>. – Дата доступа: 15.11.11
8. Коняев А. Вычисления в колбе: прошлое, настоящее и будущее химических компьютеров / Коняев А.. –Режим доступа: <http://www.nanonewsnet.ru/articles/2011/vychisleniya-v-kolbe-proshloe-nastoyashchee-budushchee-khimicheskikh-kompyuterov>. – Дата доступа: 15.11.11
9. Химический компьютер будет действовать как клетки мозга . – Режим доступа: http://virtual-reality.at.ua/load/zheleznye_novosti/zheleznye_novosti/khimicheskij_kompjuter_budet_dejstvovat_kak_kletki_mozga/24-1-0-135. – Дата доступа: 15.11.11
10. Японцы создали уникальный нанокomпьютер. – Режим доступа: <http://news.studclub.poltava.ua/new/14061>. – Дата доступа: 15.11.11
11. Перспективы развития вычислительной техники. Квантовые вычисления. – Режим доступа: http://paralichka85.px6.ru/11future/glava11_3.htm. – Дата доступа: 15.11.11
12. Яковлев Ю.С. Об оптимизации архитектуры компьютерных систем на современной элементной базе / Ю.С. Яковлев // Управляющие системы и машины I. – 2007. – № 5. – С. 43 – 51.
13. Яковлев Ю.С. Об оптимизации архитектуры компьютерных систем на современной элементной базе / Ю.С. Яковлев // Управляющие системы и машины II. – 2007. – № 6. – С. 28 – 35.
14. Яковлев Ю.С. Оптимизация вычислительного процесса в компьютерных системах архитектурными методами . Праці міжнародної конференції «50 років Інституту кібернетики імені В.М.Глушкова НАН України», Україна, Київ, 24 – 26 грудня 2008 року. – С. 182 – 189.
15. Пат. 57629 Україна, МПК (2006) G06F 15/16, G06F 13/42. Інтелектуальна розподілена система пам'яті з кільцевою шиною/ Палагін О. В., Яковлев Ю. С., Тихонов Б. М., Єлісєєва О. В.; Інститут кібернетики ім. В.М. Глушкова НАН України; заявл. 16.07.2010; опубл. 10.03.2011, Бюл. № 5/2011 . – 41с.

16. Яковлев Ю.С. Компьютерная система типа “Процессор-в-памяти” с модифицированной кольцевой шиной / Ю.С. Яковлев, Б.М. Тихонов, Е.В. Елисеева // Управляющие системы и машины . – 2011 . – № 3 . – С. 54-61, 88.
17. Яковлев Ю.С. Реконфигурируемые PIM-системы: методология построения, примеры моделей / Ю.С. Яковлев // Математичні машини і системи. – 2007. – № 2.– С. 27 – 42.
18. Пат. 6259 України, G06F13/00, G06F12/00. Система пам’яті з інтеграцією функцій зберігання та обробки інформації на одному кристалі / Сергієнко І.В., Кривонос Ю.Г., Палагін О.В. , Коваль В.М., Яковлев Ю.С., Тихонов Б.М. ; Інститут кібернетики Імені В.М.Глушкова НАН України. опубл. 15.04.2005, Бюл. № 4.– 24 с.
19. Яковлев Ю.С. Разработка и моделирование процессорного элемента операционной среды PIM-системы / Яковлев Ю.С, Тихонов Б.М. //Электронное моделирование. – 2009. – Т.31, № 2. – С.65 – 79.
20. Пат. №88666 України, G06F 7/38, G06F 7/52. Операційний пристрій PIM-системи / Яковлев Ю.С, Тихонов Б.М., Інститут кібернетики Імені В.М.Глушкова НАН України, Бюл. № 21. – 26 с.

Інформація про автора:

Яковлев Юрий Сергеевич – докт. техн. наук, зав. отделом Інститута кібернетики імені В. М. Глушкова НАН України, проспект Академіка Глушкова 40, г. Київ, 03680, тел. 526-32-07, e-mail: jakus@bigmir.net.

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ТА ОБЧИСЛЮВАЛЬНІ МЕТОДИ

УДК 519.6

В.О. БОГАЄНКО, Ю.Ю. ДАНИЛЕНКО, Г.С. ФІНІН

Інститут водних проблем і меліорації НААН України, Київ

НАБЛИЖЕННЯ ГЕОІНФОРМАЦІЙНИХ ДАНИХ З ВИКОРИСТАННЯМ МЕТОДИКИ ПСЕВДООБЕРНЕННЯ

Анотація. Розглядається задача наближення геоінформаційних даних у випадку їх некоректності та наявності обмежень, що накладаються фізичною сутністю задачі. Пропонується шукати наближення, відштовхуючись від спрощеної математичної моделі у вигляді диференціального рівняння в частинних похідних еліптичного типу, розв'язуючи обернені задачі щодо нього методом функцій Гріна з використанням методики псевдообернення, що дозволяє отримувати множини можливих інтерполяцій. Пропонуються на основі методики почергового проектування процедури вибору з цієї множини, розв'язків, які задовольняють обмеженням (з інтерполюючим чи апроксимуючим вихідні дані результатом), й враховують похибки вимірювання вихідних даних. Приводяться результати обчислювального експерименту, які демонструють адекватність отриманих інтерполяцій фізичним процесам.

Ключові слова: інтерполяція, апроксимація, псевдообернення, функція Гріна.

Аннотация. Рассматривается задача приближения геоинформационных данных в случае их некорректности и наличия ограничений, которые накладываются физической сутью задачи. Предлагается искать приближения, отталкиваясь от упрощенной математической модели в виде дифференциального уравнения в частных производных эллиптического типа, решая обратные задачи для него методом функций Грина с использованием методики псевдообращения, что позволяет получать множества возможных интерполяций. Предлагаются, на основе методики попеременного проектирования, процедуры выбора из этого множества, решений, которые удовлетворяют ограничениям (с интерполирующим или аппроксимирующим исходные данные результатом), и учитывают погрешности измерения исходных данных. Приводятся результаты вычислительного эксперимента, которые демонстрируют адекватность полученных интерполяций физическим процессам.

Ключевые слова: интерполяция, аппроксимация, псевдообращение, функция Грина.

Abstract. Geoinformational data approximation problem solution in the case of their inaccuracy and presence of restrictions which are imposed by physical problem substance has been studied. Approximations finding, based on simplified mathematical model in the form of elliptical partial differential equation, by solving inverse problems for it using Green's functions method and pseudoinversion methodic, that allows to obtain sets of possible interpolations, has been suggested. On the base of alternating projection method, procedures for choosing solutions from this sets which satisfy restrictions (with result, interpolating or approximating initial data) and take into account initial data measurement inaccuracy, have been proposed. Computational experiment results which show adequacy of obtained interpolation to physical processes have been given.

Key words: interpolation, approximation, pseudoinversion, Green's function.

Вступ

Такі властивості геоінформаційних даних, як просторовий розподіл та великі похибки вимірювань створюють складності при їх аналізі та візуалізації. Задля врахування похибок, замість методів інтерполяції таких, як лінійна інтерполяція чи інтерполяція методом обернених зважених відстаней [1], застосовуються методи апроксимації аналізованої функції, зокрема, крігінг [2].

Крігінг – це метод, базований на використанні методів математичної статистики, який інтерполює значення випадкової величини $Z(\vec{x})$ у неспостережуваній точці \vec{x} відштовхуючись від спостережень у близьких до \vec{x} точках \vec{x}_i . Інтерполяція виконується за допомогою методу найменших квадратів на основі незміщеної лінійної оцінки з мінімальною дисперсією на стохастичній моделі просторової залежності, що описується варіограмою або математичним очікуванням та коваріаційною функцією.

Альтернативою таким методам інтерполяції є методи, у яких використовується інформація стосовно фізичної суті аналізованих даних (такі, як підхід, описаний у [3]). Зокрема, для інтерполяції даних, пов'язаних з процесами у ґрунтових водах, доцільно розв'язувати обернені задачі щодо еліптичних диференціальних рівнянь у частинних похідних. Для розв'язання таких задач використовується символічний підхід Лур'є [4,5], метод функцій Гріна та методика псевдообернення [6,7]. При застосуванні такого підходу, за наявності некоректностей у вихідних даних, результуюча інтерпольована функція в деяких випадках може приймати від'ємні значення в ситуації, коли її фізична сутність не дозволяє цього. Адекватне розв'язання задачі тоді вимагає застосування додаткових процедур, що враховують особливості вихідних даних.

Актуальність

Більшість методів інтерполяції та апроксимації геоінформаційних даних не враховують обмеження, що їх накладає фізична сутність даних. Це зумовлює необхідність розробки нових методів наближення геоінформаційних даних, які дозволили б отримувати більш адекватні апроксимації.

Мета

Підвищення ефективності обробки геоінформаційних даних з використанням методики псевдообернення.

Задачі

Розробка процедур інтерполяції, що враховують обмеження, які накладаються фізичною сутністю інтерпольованих даних, та некоректності у них.

Розв’язання задач

Розглянемо крайову задачу для еліптичного рівняння виду

$$\operatorname{div}(\vec{k} \operatorname{grad} y(\vec{x}) - \vec{w} y(\vec{x})) = q(\vec{x}), \quad \vec{x} = (x_1, x_2) \in \Omega, \tag{1}$$

де компоненти коефіцієнта $\vec{k} = (k_1, k_2)$ – задані додатні числа, вектор $\vec{w} = (w_1, w_2)$, $q(\vec{x})$ – інтегрована за Ріманом в області моделювання Ω та обмежена функція, $y(\vec{x})$ – шукана функція.

Будемо розглядати рівняння (1) за відсутності крайових умов, проте з умовою керування щодо правої частини рівняння. Функція $q(\vec{x})$ є невідомою й знаходиться з умови:

$$y(\vec{x}) = Y_1(\vec{x})|_{D^C}, \tag{2}$$

де $Y_1(\vec{x})$ – задана функція, D^C – задана підобласть області моделювання Ω .

Розв’язок задачі (1) – (2) шукається у вигляді

$$y(s) = \int_{\Omega} G(s - s^*) q(s^*) ds^*, \quad s = (x_1, x_2),$$

де $G(s)$ – функція Гріна рівняння (1) у необмеженій області зміни координат $s = (x_1, x_2)$.

З урахуванням умови (2) отримуємо інтегральне рівняння

$$\int_{\Omega} G(s - s^*) q(s^*) ds^* = Y_1(s), \quad s \in D^C. \tag{3}$$

Шукаючи розв’язки рівняння (3) на класі функцій виду

$$q(s) = \sum_{i=1}^n q_i \delta(s - s_i), \quad \delta(\vec{x}) = \begin{cases} 1, & \vec{x} = (0, \dots, 0) \\ 0, & \vec{x} \neq (0, \dots, 0) \end{cases},$$

де S_i – точки дискретизації області Ω , а n – кількість точок дискретизації, отримуємо систему лінійних алгебраїчних рівнянь (СЛАР)

$$\sum_{i=1}^n G(s'_j - s_i) q_i = Y_1(s'_j), \quad j = 1, \dots, m. \tag{4}$$

Тут s'_j – точки дискретизації підобласті D^C , а m – їх кількість. У випадку задачі інтерполяції, s'_j та $Y_1(s'_j)$ є вихідними даними.

З використанням методики псевдообернення можна побудувати множину розв’язків СЛАР (4), або середньоквадратичних наближень до них у вигляді

$$Q = A^+ Y + (I - A^+ A) v, \tag{5}$$

де $Q = (q_1, \dots, q_n)^T$, $A = (\bar{G}_1, \dots, \bar{G}_n)$, $\bar{G}_i = (G(s'_1 - s_i), \dots, G(s'_m - s_i))^T$, $Y = (Y_1(s'_1), \dots, Y_1(s'_m))^T$, A^+ - матриця, псевдообернена до A , v , $\dim v = [1 \times n]$, - довільний вектор.

Розв'язок задачі інтерполяції знаходимо у вигляді

$$y(s) = \sum_{i=1}^n G(s - s_i) q_i \quad (6)$$

Якщо фізична сутність інтерпольованої функції визначає її невід'ємність, проте розрахована за формулою (6) при $v = 0$ інтерполяція умові $y(s) \geq 0$ не задовольняє, можливо не порушуючи дискретного аналогу умови (2) знайти необхідний розв'язок, або наближення до нього, шляхом вибору параметру v , враховуючи, що при $G(s) \geq 0$, $y(s) \geq 0$ якщо $Q \geq 0$.

Розв'язок системи лінійних нерівностей

$$Q = A^+ Y + (I - A^+ A) v \geq 0 \quad (7)$$

пропонується шукати шляхом ітеративного проектування вектору $Q^{(i)}$ спочатку на множину $A^+ Y + (I - A^+ A) v$, а потім на множину $Q \geq 0$.

Таким чином, якщо $Q^{(i)}$ - розв'язок на i -тій ітерації, $Q^{(0)} = A^+ Y$,

$$\bar{Q}^{(i)} = \begin{cases} Q^{(i)}, & Q^{(i)} \geq 0 \\ 0, & Q^{(i)} < 0 \end{cases} -$$

проекція вектору $Q^{(i)}$ на множину $Q \geq 0$, а завдяки проєктивним властивостям оператора $(I - A^+ A)$,

$$Q^{(i+1)} = A^+ Y + (I - A^+ A) \bar{Q}^{(i)} -$$

проекція вектору $\bar{Q}^{(i)}$ на множину $A^+ Y + (I - A^+ A) v$.

Оскільки розглядувані множини є опуклими, $\|Q^{(i)} - \bar{Q}^{(i)}\| \geq \|Q^{(i+1)} - \bar{Q}^{(i)}\|$. Ітераційний процес сходиться на $i + 1$ -ій ітерації до розв'язку системи (7) коли $\|Q^{(i+1)} - \bar{Q}^{(i)}\| = 0$, або дозволяє отримати наближення до розв'язку при $\|Q^{(i)} - \bar{Q}^{(i)}\| = \|Q^{(i+1)} - \bar{Q}^{(i)}\|$.

У другому випадку, при виборі $Q^{(i+1)}$ у якості розв'язку задачі отримуємо інтерполяцію, що задовольняє дискретному аналогові умові (2), проте не задовольняє умові невід'ємності (процедура 1), при виборі ж $\bar{Q}^{(i)}$, отримуємо невід'ємну апроксимацію вихідної функції (процедура 2).

Вищеописаний алгоритм дозволяє врахувати й похибку \mathcal{E} вимірювання вихідних даних. Для цього замість вектора $Y = (Y_1(s'_0), \dots, Y_1(s'_m))^T$ вводиться паралелепіпед

$$\bar{Y} = Y + \sum_{i=1}^m \alpha_i \bar{Y}_i, \quad \bar{Y}_i = (0, \dots, \varepsilon Y_1(s'_i), \dots, 0)^T, \quad \alpha_i \in [-1, 1].$$

Оскільки розглядувані задачі є лінійними, то множина у цьому випадку (5) прийме вигляд

$$\hat{Q}(\alpha) = Q + \tilde{Q}(\alpha), \quad \tilde{Q}(\alpha) = \{\tilde{q}_i(\alpha)\}_{i=1}^n = \sum_{i=1}^m \alpha_i A^+ \bar{Y}_i, \quad \alpha = (\alpha_0, \dots, \alpha_m),$$

а множина інтерполяцій (6) –

$$\bar{y}(\alpha, s) = y(s) + \sum_{i=1}^n G(s - s_i) \tilde{q}_i(\alpha) \quad (8)$$

Невід’ємний розв’язок, або наближення до нього, будемо знаходити як

$$y_m(s) = \bar{y}(\alpha_m, s), \quad \alpha_m : \|\tilde{y}(\alpha_m, s)\| = \min_{\alpha} \|\tilde{y}(\alpha, s)\|,$$

$$\|\tilde{y}(\alpha_m, s)\| = \int_{\Omega} \tilde{y}^2(\alpha, s) ds, \quad \tilde{y}(\alpha, s) = \begin{cases} \bar{y}(\alpha_m, s), & \bar{y}(\alpha_m, s) < 0 \\ 0, & \bar{y}(\alpha_m, s) \geq 0 \end{cases} \quad (9)$$

Задачу мінімізації (9) можна розв’язувати, зокрема, градієнтними методами, а отриманий розв’язок буде невід’ємною апроксимацією, або наближенням до неї, вихідних даних з врахуванням похибки їх вимірювання (процедура 3).

Запропоновані процедури знаходження невід’ємної інтерполяції були програмно реалізовані та тестовані на наборі даних з 55 точок. Результати інтерполяції без використання процедур знаходження невід’ємного розв’язку зображено на рис.1, а з використанням трьох вищеописаних процедур на рис.2-4.

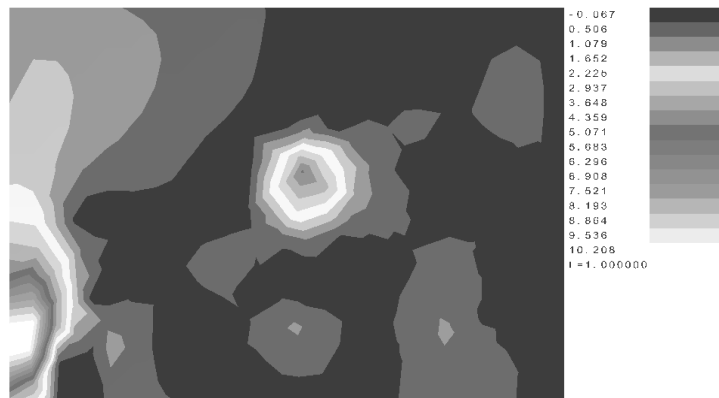


Рисунок 1 – Результат інтерполяції без використання процедур знаходження невід’ємного розв’язку

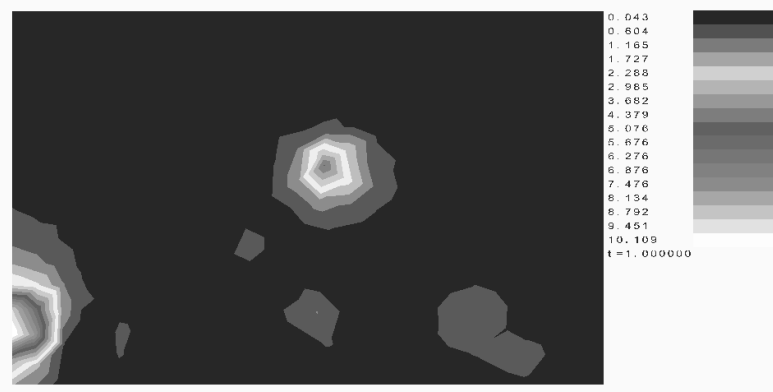


Рисунок 2 – Результат інтерполяції з використанням процедури 1

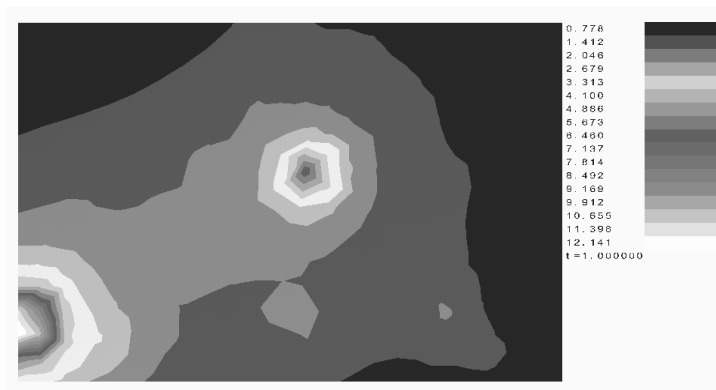


Рисунок 3 – Апроксимація з використанням процедури 2

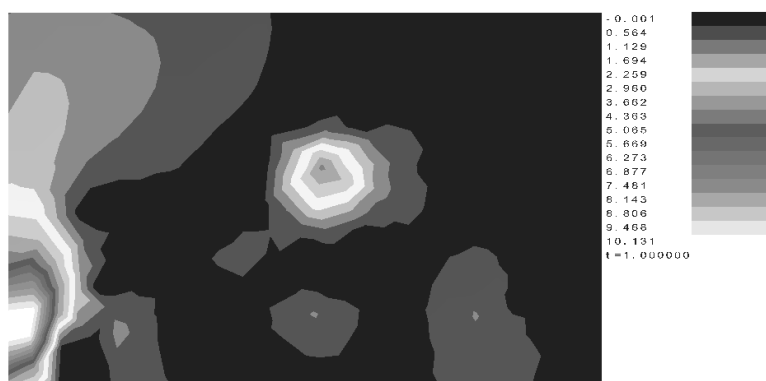


Рисунок 4 – Апроксимація з використанням процедури 3 з врахуванням 30% похибки вимірювання

Висновки

1. Алгоритм інтерполяції геоінформаційних даних, базований на символічному підході Лур'є, методі функцій Гріна та методикі псевдообернення дозволяє отримувати множини інтерполяцій щодо яких були розроблені процедури вибору невід'ємного розв'язку та врахування похибок вимірювання вихідних даних.
2. Аналіз розрахованих інтерполяцій показав, що у певних випадках розроблені процедури дозволяють отримати фізично коректні розв'язки, які неможливо отримати базовим алгоритмом.

Список літератури

1. Li, J., A. D. Heap A review of spatial interpolation methods for environmental scientists, Geoscience Australia, Canberra, 2008, 137pp.
2. Matheron, G., Principles of geostatistics // Economic Geology, 58, pp 1246–1266, 1963
3. Ковальчук П.І., Шевчук С.А. Математичне моделювання в системі моніторингу затоплення і підтоплення сільгоспугідь і сільських населених пунктів // Вісник УДУВГП. Вип. 4(40). Ч. 1. – Рівне. – 2007. – С. 285 – 290.
4. Лурье А.И. К теории толстых плит // Прикл. Математика и механика – 1942. – 6, вып. 23. – С. 12-36.
5. Лурье А.И. Пространственные задачи теории упругости. – М.: Гостехиздат, 1955. – 370 с.
6. Скопецкий В.В., Стоян В.А., Кривonos Ю.Г. Математичне моделювання прямих та обернених задач динаміки систем з розподіленими параметрами - К.:Наукова Думка, 2002.- 361 с.
7. Стоян В.А. Псевдообращение интегральных операторов в задачах наблюдения, терминального управления и моделирования динамики систем с распределенными параметрами // Проблемы управления и информатики. –1998. –№4. С. 112–120.

Відомості про авторів

Богасенко Всеволод Олександрович, к. т. н., старший науковий співробітник, Інститут кібернетики ім. В.М. Глушкова НАНУ, sevab@ukr.net.

Даниленко Юлія Юрївна, к. т. н., зав. лабораторії, Інститут водних проблем і меліорації НААН України.

Финин Георгий Семенович, д. ф.-м. н., главный научный сотрудник, Институт водных проблем и меліорації НААН України.

УДК 681.32

Р. Н. КВЕТНИЙ, О. О. ДРУЖИНИНА

Вінницький національний технічний університет, Вінниця

МЕТОДИ ФОРМУВАННЯ НАВЧАЛЬНОЇ ВИБІРКИ В ЗАДАЧАХ МОДЕЛЮВАННЯ ЧАСОВИХ РЯДІВ НЕЙРОННИМИ МЕРЕЖАМИ

Анотація. Здійснено огляд існуючих підходів до формування навчальної вибірки в задачах моделювання часових рядів радіально-базисними нейронними мережами. Розглядаються методи попередньої обробки кількісних ознак навчальної вибірки. Проведені експериментальні дослідження підтверджують доцільність використання додаткових вхідних компонент та попередньої обробки вхідних даних.

Анотация. Приведен обзор существующих подходов к формированию обучающей выборки в задачах моделирования временных рядов радиально-базисными нейронными сетями. Рассматриваются методы предварительной обработки признаков обучающей выборки. Проведенные экспериментальные исследования подтверждают целесообразность использования дополнительных входных компонент и предварительной обработки входных данных.

Abstract. Existing approaches to the training set generating issue for time series modeling using radial-basis neural networks tasks were analyzed. Methods of training set features preprocessing are considered. Experimental results confirm the reasonability of additional input components using and input data preprocessing.

Вступ

Задача аналізу часових рядів (ЧР) являє собою широку область досліджень у різних областях науки, яка зазнає найбільш стрімкого розвитку останнім часом. Така задача є актуальною для часових рядів, що описують об'єкти різної природи: біржових котирувань акцій, медицини, сейсмології та багато інших.

Класичні методи ідентифікації часових рядів, стають все менш придатними для моделювання складних нелінійних систем. Адекватним апаратом для побудови моделей практично будь-яких нелінійних структур можуть слугувати методи, побудовані на основі штучного інтелекту, а саме штучні нейронні мережі, які мають здатність до моделювання нелінійних процесів, адаптації та дозволяють працювати з зашумленими даними.

Саме таким інструментом є радіально-базисні нейронні мережі (RBF NN – Radial Basis Function Neural Network), які на фоні інших інтелектуальних засобів відрізняються особливо високою швидкістю навчання, і сьогодні широко використовуються для розв'язання задач моделювання часових рядів [1, 2].

Актуальність

В нейромережевому підході задача прогнозування часових рядів може бути сформульована як задача розпізнавання образів. Радіально-базисні нейронні мережі передбачають використання парадигми навчання «з вчителем». Під навчальною вибіркою будемо розуміти сукупність прецедентів – пар «об'єкт, клас», яка подається на вхід нейронної мережі.

Аналіз літературних джерел показав, що на сьогодні питанню формування навчальної вибірки приділено недостатньо уваги, хоча ефективність моделювання часових рядів у значній мірі визначається якістю формування навчальної вибірки [1]. Серед праць науковців які присвячені питанням формування навчальної вибірки слід виділити роботи В.А.Крісілова, Н.В. Пєскова, Р.О. Тарасенко, Д. Н. Олешко, О. А. Блажко, Р. І. Франка, Н. Деві, С. П. Ханта, Х. Мурвейта, М. Вайнтрауба, М. Коена [3-11]. Роботи цих вчених, в переважній мірі, присвячені дослідженню і розробці комплексних критеріїв якості навчальної вибірки. Технології, методики та методи, присвячені безпосередньо вирішенню задачі формування набору прецедентів, а не оцінці його якості в наукових працях освітлені не в достатній мірі. Це і визначило актуальність і задачі даного дослідження.

Мета

Метою даного дослідження є підвищення ефективності моделювання часових рядів засобами радіально-базисних нейронних мереж за рахунок підвищення якості навчальної вибірки.

Постановка задачі

Для ідентифікації часових рядів, оцінки поточного моменту і перспектив розвитку методами розпізнавання образів висувається наступна основна гіпотеза. Нехай в момент часу $t = t_0$ існує деякий набір (вектор) факторів процесу, що досліджується:

$$x(t_0) = \{x_1(t_0), x_2(t_0), x_3(t_0), \dots, x_N(t_0)\}. \quad (1)$$

Припускається, що внаслідок існування $x(t_0)$ в момент часу $(t_0 + T)$ реалізується набір векторів, які оцінюють розвиток процесу:

$$Q(t_0 + T) = \{q_1(t_0 + T), q_2(t_0 + T), q_3(t_0 + T), \dots, q_m(t_0 + T)\}. \quad (2)$$

Тобто існує деяке відображення простору X в простір Q . Це відображення може бути інтерпретоване як функціональна або як кореляційна залежність.

Компоненти $q_j (j = \overline{1, m})$ можуть задовольняти або не задовольняти нормативні умови. Таким чином, за результатами кожного експерименту, фіксуються значення компонент вектора $Q(t_0 + T)_i$ визначається до якого з класів $B_i (i = 1 \dots k)$ відноситься отриманий результат (табл. 1).

Таблиця 1 - Таблиця відповідності вхідних компонент заданим класам

<i>N</i> експер.	$q_1(t_0)$	$q_2(t_0)$...	$q_n(t_0)$	$q_1(t_0+T)$...	$q_m(t_0+T)$	Клас
1	x_{11}	x_{21}	...	x_{n1}	q_{11}	...	q_{m1}	$B_{i1}(i=1 \dots k)$
2	x_{12}	x_{22}	...	x_{n2}	q_{12}	...	q_{m2}	$B_{i2}(i=1 \dots k)$
...
<i>N</i>	x_{1N}	x_{2N}	...	x_{nN}	q_{1N}	...	q_{mN}	$B_{iN}(i=1 \dots k)$

Дані наведеної таблиці можуть використовуватись як навчальна вибірка без попередньої обробки. Але використання «чистих» вхідних даних для навчальної вибірки веде до збільшення розмірності вхідного вектора та часу для навчання нейромережевого класифікатора. У вирішенні цієї проблеми допомагає використання вхідних компонент, які являють собою виділені ознаки. Наявність надлишкових і непотрібних ознак обтяжуватиме навчальний процес, в той час як замала кількість ознак може бути недостатньою для представлення важливих ознак образу.

Задача полягає у аналізі існуючих методів формування навчальної вибірки та дослідженні залежності ефективності ідентифікації часових рядів засобами радіально-базисних нейронних мереж від методу формування навчальної вибірки.

Розв'язання задачі

Аналіз літературних джерел показав, що зазвичай для опису ситуації вибирається однакова глибина занурення – використовується метод ковзного вікна (moving/sliding window) зі стаціонарною шириною. Схематично даний метод зображений на рисунку 1.

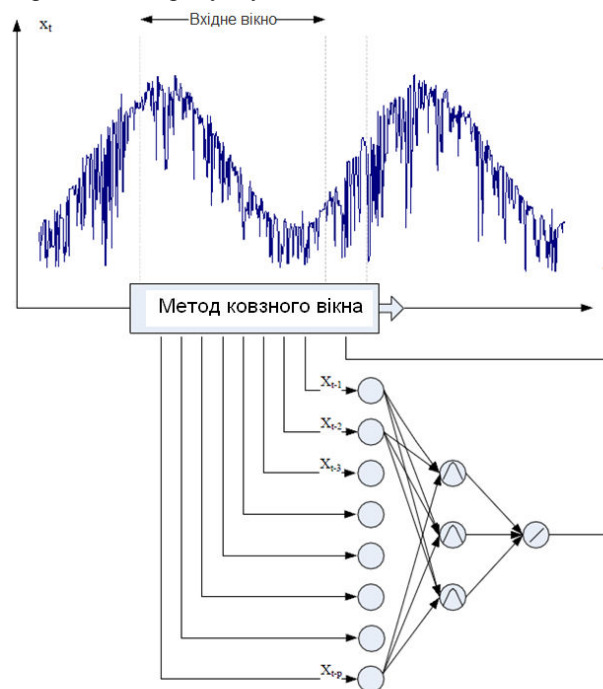


Рисунок 1 – Прогнозування часових рядів методом ковзного вікна

У випадку застосування даного методу необроблені дані спостережень, як правило, перетворюються шляхом центрування і нормалізації або ж подаються на вхід нейронної мережі без попередньої обробки [12-14].

Дані процедури являють собою процедуру лінійного перетворення зразків (X_t) до нормальних стандартних змінних (Z_t). Це може бути здійснено шляхом використання наступного рівняння:

$$Z_t = \frac{X_t - \mu}{\sigma}, \quad (3)$$

де μ – середнє відхилення від статистично стабільного процесу, σ – середнє квадратичне відхилення випадкового процесу X_t . За умови стабільного процесу, Z_t задовольняє нормальний розподіл в межах $[-3, 3]$ з нульовим середнім і одиничним середньоквадратичним відхиленням. В роботі [15] зазначається, що навчання НМ буде більш ефективним, якщо дані знаходяться в межах певного діапазону.

Нормалізація зразків (Z_t) до компактного інтервалу $[0, 1]$ чи $[-1, 1]$ дає можливість мінімізувати вплив випадкових шумів.

Процедури центрування і нормалізації є корисними у випадках, коли дані змінюються у широкому діапазоні. Процедура нормалізації також часто застосовується для ремасштабування виділених ознак в діапазон між $[-1, 1]$ для представлення в класифікаторах НМ.

Нормалізація до діапазону між $[0, 1]$ може бути виконана шляхом застосування наступного перетворення [16]:

$$Y_t = \frac{Z_t - Z_{\min}}{Z_{\max} - Z_{\min}}, \quad (4)$$

де Z_{\min} – мінімальне значення даних чи виділених ознак; Z_{\max} – максимальне значення даних чи виділених ознак.

Нормалізація до діапазону між $[-1, 1]$ може бути виконана шляхом застосування наступного перетворення [17, 18]:

$$Y_t = \frac{2(Z_t - Z_{\min})}{Z_{\max} - Z_{\min}} - 1. \quad (5)$$

В роботах [13, 14] для вхідного представлення даних використовувалось бінарне кодування. Дане кодування являло собою перетворення стандартизованих зразків (Z_t) закодованої форми, тоді як зонування являла собою процедуру ре-масштабування і поділу графіку стандартизованих даних на 7 зон ($zone+3, zone+2, zone+1, zone0, zone-1, zone-2, zone-3$). Наприклад, якщо зразок розташований в зоні $zone+2$, то бінарне кодування буде представлятись як '0100000'. Один зразок потребує сім вхідних нейронів. Тим паче, це потребує відносно великого розміру мережі і збільшення обчислень.

З іншої сторони в роботах [18, 19, 20] було опубліковано іншу технологію обробки даних. Зразки були лінійно перетворені в діапазон між $[-7.625, 7.625]$, який відрізняється від звичайного діапазону стандартизації $[-3, 3]$. Далі перетворені зразки були поділені на 61 зону з інтервалом в 0,25 стандартного відхилення. Вони зазначили, що великий діапазон ранжування перетворених зразків і зонування дозволить ідентифікувати процеси з високими відхиленнями, які досягають 4 стандартних відхилень.

Дослідники в області аналізу часових рядів розглядали множини ознак для стиснення часових рядів. Широко використовувались перетворення Фур'є, що дозволяли підвищити швидкість даної процедури [21-24], однак цей підхід має декілька недоліків. На практиці, ці перетворення згладжують локальні мінімуми та максимуми, що може призвести до втрати важливої інформації.

В роботі [25] для розв'язання даної задачі пропонується використовувати вейвлети і показані переваги цієї технології над перетвореннями Фур'є.

Однак, аналізуючи нестационарні ЧР об'єктів різної природи легко бачити, що для різних ділянок ЧР ця необхідна глибина занурення різна. Таким чином, з однієї сторони на різних ділянках ЧР необхідно формувати образи з різним розміром ситуації, а з іншої сторони розмір вхідних векторів збудження навчальних наборів навчальної вибірки має бути однаковим. Саме це протиріччя створює проблему вибору розміру опису ситуації для нестационарних ЧР. Лише в небагатьох джерелах на це звертається увага. Однак, наприклад, в роботі [3] запропоновано модифікований метод формування НВ в

задачах прогнозування ЧР, що відрізняється від методу «вікон» тим, що вхідне «вікно», що задає розпізнаваний образ навчального набору, має перемінний розмір, що залежить від складності ділянки ряду, на якій формується навчальний набір. Для досягнення цього пропонується розділити поняття розміру опису образу і розміру вікна W_i , що переміщається по ЧР. Ділянка ЧР, що розглядається через вікно W_i , розбивається на ділянки, складність яких однакова. Кожна з цих ділянок, узагалі говорячи, має різну довжину. Для кожної з ділянок формується стиснутий опис за рахунок завдання її не переліком значень, а коефіцієнтами апроксимуючої функції. Таким чином розпізнаваний образ має фіксовану довжину і задається вектором стиснутих описів ділянок однакової складності.

Погоджуючись з твердженням, що використання однакового розміру опису ситуації не дозволяє отримати інформативну навчальну вибірку пропонується використовувати сегментований часовий ряд для виділення найбільш важливої інформації з масиву початкових даних, шляхом ідентифікації важливих точок мінімуму та максимуму і видалення інших незначних коливань [3]. Ступінь сегментації визначається показником R , який завжди має перевищувати одиницю, тобто збільшення даного показника буду призводити до зменшення точок (рис 2).

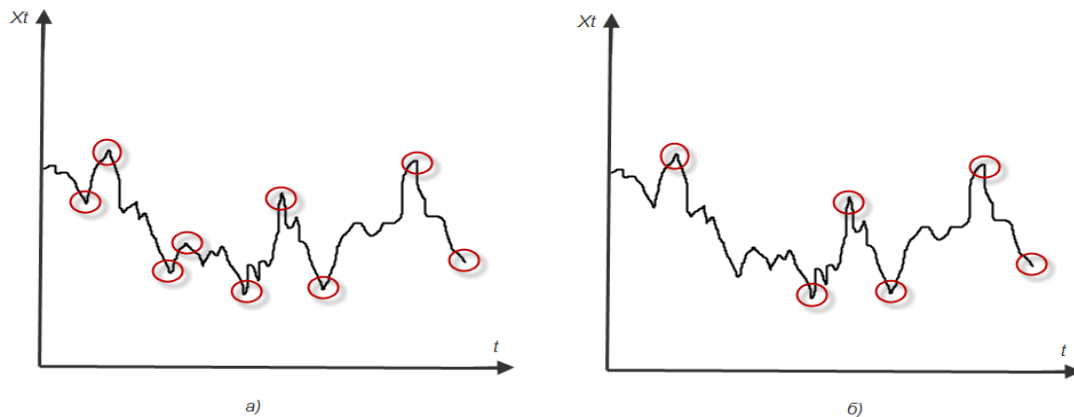


Рисунок 2 – Сегментований часовий ряд: а) $R=90\%$, б) $R = 95\%$

Точка x_{min} ряду x_1, \dots, x_n , буде важливими мінімумом (рис. 4 – а), якщо є такі індекси t_1 та t_2 , де $t_1 < min < t_2$, такі, що:

$$\begin{cases} x_{min} = \min[x(t_1), x(t_2)] \\ \frac{x(t_1)}{x_{min}} \geq R, \\ \frac{x(t_2)}{x_{min}} \geq R. \end{cases} \quad (6)$$

Аналогічно точка a_m ряду a_1, \dots, a_n , буде важливими максимумом (рис. 4 – б), якщо є такі індекси t_1 та t_2 , $t_1 < max < t_2$, такі, що:

$$\begin{cases} x_{max} = \max[x(t_1), x(t_2)], \\ \frac{x_{max}}{x(t_1)} \geq R, \\ \frac{x_{max}}{x(t_2)} \geq R. \end{cases} \quad (7)$$

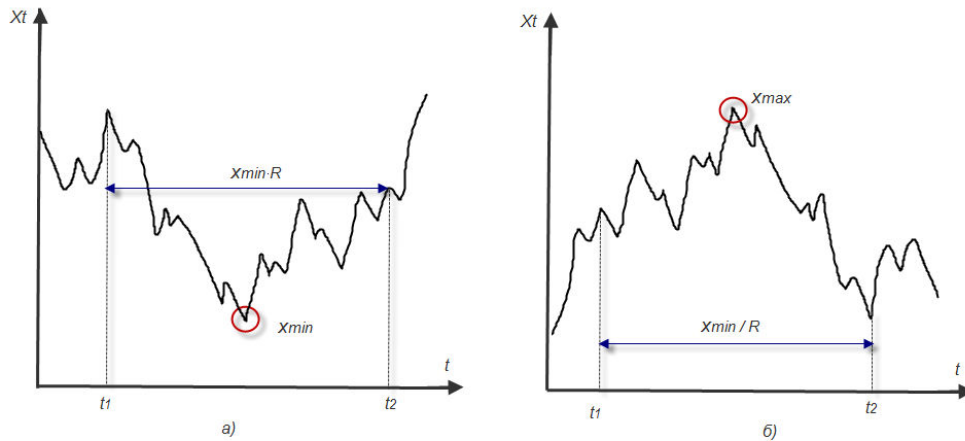


Рисунок 3 – приклади важливого мінімуму (а) і важливого максимуму (б)

Результати експериментальних досліджень

Для аналізу розглянутих підходів та методів щодо підготовки навчальної вибірки було проведено комп’ютерне моделювання. Досліджувалась ефективність ідентифікації часових рядів на основі ймовірнісної нейронної мережі (PNN) з використанням необроблених вхідних даних у навчальній вибірці та навчальних зразків отриманих шляхом попередньої обробки даних та з використанням запропонованого алгоритму сегментації. В результаті застосування запропонованої технології для ідентифікації часових рядів валютних курсів було отримано сигнальний індикатор, значення якого використовувались в режимі реального часу. Сигнальний індикатор у випадку класифікації на 2 класи може приймав 3 можливих значення: 1 – сигнал на купівлю фінансового інструменту, -1 – сигнал на продаж фінансового інструменту, 0 – у випадку відсутності сигналу щодо зміни тенденції.

Достовірність ідентифікації була оцінена за допомогою наступних показників: FRR (помилка першого роду), FAR (помилка другого роду) та Detection Rate (DR). Для задачі ідентифікації ЧР, у випадку класифікації поточної ринкової ситуації на два класи, помилка FRR визначалась як середнє значення пропусків зміни тенденції для класів Buy і Sell, а помилка FAR як середнє значення хибних виявлень зміни тенденції для цих же класів. Показник Detection Rate (DR) визначався за формулою

$$DR = \frac{\text{кількість зразків, коректно класифікованих}}{\text{загальна кількість зразків в множині даних}} \cdot 100\%. \quad (8)$$

В результаті проведення досліджень була оцінена ефективність технології ідентифікації часових рядів з застосуванням різних моделей НМ на різних часових рядах і інтервалах (табл. 2).

Таблиця 2 – Результати експериментальних досліджень

Часовий ряд	Навчальна вибірка	FRR (%)	FAR (%)	DR (%)
Часовий ряд валютного курсу GBPUSD. Часовий інтервал – 1 год.	побудована на основі тільки вхідних нормалізованих даних	31	18	58
	побудована на основі нормалізованих вхідних даних; додатково враховувались компонента екстремумів та низка технічних індикаторів (Stochastic, MA, RSI).	23	14	72
Часовий ряд споживання мережевого трафіку. Часовий інтервал – 0.5 год	побудована на основі нормалізованих даних	34	15	61
	побудована на основі вхідних нормалізованих даних; видалена сезонна компонента, додатково враховувались компонента екстремумів	25	16	63

Висновки

1. В даній роботі був проведений аналіз існуючих методів формування навчальної вибірки. Аналіз показав, що дослідження методів формування навчальної вибірки та їх впливу на результати нейромережових моделей є актуальним і необхідним для можливості розширення та уточнення правил організації процесу формування навчальної вибірки.

2. Результати комп'ютерного моделювання підтвердили ефективність використання попередньо оброблених вхідних даних на основі запропонованого алгоритму. Для досліджених в роботі прикладних задач, запропонований підхід дозволив підвищити якість ідентифікації на 11 – 14% для валютних часових рядів і незначне підвищення (2%) було отримано моделюванні часового ряду споживання мережевого трафіку. На основі отриманих даних можна також зробити висновок, що вхідні компоненти, які подаються на вхід нейронної мережі мають бути вибрані в залежності від природи об'єкту, який описує часовий ряд, що досліджується.

Список літератури

1. Кветний Р. Н. Імовірнісні нейронні мережі в задачах ідентифікації часових рядів [Електронний ресурс] / Р. Н. Кветний, В. В. Кабачій, О. О. Чумаченко // Наукові праці Вінницького національного технічного університету. – 2010. – № 3. – 6 с. – Режим доступу до журн.: <http://www.nbu.gov.ua/e-journals/VNTU/2010-3/2010-3.htm>.

2. Чумаченко О. О. Проблема формування навчальної вибірки в нейромережевому моделюванні нелінійних часових рядів : (Тези доповідей XIX міжнародної науково-практичної конференції «Інформаційні технології: наука, техніка, технологія, освіта, здоров'я») / О.О. Чумаченко. – Харків, 2011 – С. 375.

3. Тарасенко Р. А. Выбор размера описания ситуации при формировании обучающей выборки для нейронных сетей в задачах прогнозирования временных рядов / Р. А. Тарасенко, В. А. Крисиллов // Труды Одесского политехнического университета. – 2001. – № 2. – С.25 – 28.

4. Тарасенко Р.А. Предварительная оценка качества обучающей выборки для нейронных сетей в задачах прогнозирования временных рядов / Р. А. Тарасенко, В. А. Красиллов // Труды Одесского политехнического университета. – 2001. – № 1. – С. 90 – 93.

5. Крисиллов В. А. Использование гипотезы λ -компактности при построении обучающей выборки для прогнозирующих нейросетевых моделей [электронный ресурс] / В. А. Крисиллов, С. А. Юдин, Д. Н. Олешко. Режим доступу до файлу: http://journal.iasa.kpi.ua/zm456st/2006/No-3/Krisilov_Judin_Ol_N3_06.doc

6. Дюкова Е. В. Поиск информативных фрагментов описаний объектов в дискретных процедурах распознавания [Электронный ресурс] / Е. В. Дюкова, Н. В. Песков // Журнал вычислительной математики и математической физики. – 2002. – т.42, №5. – С.741-753. Режим доступу до журн.: <http://www.ccas.ru/frc/papers/djukova02poisk.pdf>

7. Колесникова С. И. Методы анализа информативности разнотипных признаков [Электронный ресурс] / С. И. Колесникова. Режим доступу до файлу: <http://www.lib.tsu.ru/mminfo/000063105/inf/06/image/06-069.pdf>.

8. Востров Г. Н. Сегментация экономических временных рядов с использованием вейвлет-анализа [Электронный ресурс] / Г. Н. Востров, М. В. Полякова, В. В. Любченко // Труды Одесского политехнического университета. – 2003. – № 1(19). Режим доступу до журн.: [http://www.library.ospu.odessa.ua/online/periodic/opu_2003_1\(19\)/3/3-6.pdf](http://www.library.ospu.odessa.ua/online/periodic/opu_2003_1(19)/3/3-6.pdf)

9. Frank R. J. Input Window Size and Neural Network Predictors [Электронный ресурс] / N. Davey, S. P. Hunt. Режим доступу до файлу: <http://homepages.feis.herts.ac.uk/~nngroup/pubs/papers/frank-ijcnn00.pdf>

10. Murveit H. Training Set Issues in SRI's DECIPHER Speech Recognition System [Электронный ресурс] / Murveit H., Weintraub M., Cohen M. Режим доступу до файлу: <http://acl.ldc.upenn.edu/H/H90/H90-1065.pdf>

11. Preprocessing Time Series Data for Classification with Application to CRM / [Yiming Yang, Qiang Yang, Wei Lu, Jialin Pan, Rong Pan, Chenhui Lu, Lei Li, and oth]. Режим доступу до файлу: [http://www1.i2r.a-star.edu.sg/~jspan/publications/\[AI05\]Preprocessing%20Time%20Series%20Data%20for%20Classification%20with%20Application%20to%20CRM.pdf](http://www1.i2r.a-star.edu.sg/~jspan/publications/[AI05]Preprocessing%20Time%20Series%20Data%20for%20Classification%20with%20Application%20to%20CRM.pdf)

12. Christophe Paoli Solar radiation forecasting using ad-hoc time series preprocessing and neural networks [Электронный ресурс] / Christophe Paoli, Cyril Voyant, Marc Muselli, Marie-Laure Nivet. Режим доступу до файлу: <http://arxiv.org/ftp/arxiv/papers/0906/0906.0311.pdf>

13. Hwang H. B. Detecting process non-randomness through a fast and cumulative learning ART-based pattern recognizer / H. B. Hwang, C.W. Chong // International Journal of Production Research. – 1995. – № 33. – PP. 1817-1833.

14. Hwang H. B. X-Bar Chart pattern recognition using neural nets / H. B. Hwang, N. F. Hubele // ASQC quality congress transactions. – 1991. – PP. 884-888.
15. Barghash M. A. Pattern Recognition of Control Charts Using Artificial Neural Networks - Analysis the Effects of the Training Parameters / M. A. Barghash, N. S. Santarisi // Journal of Intelligent Manufacturing. – 2004. – №15. – PP. 635-644.
16. Gauri S. K. Improved recognition of control chart patterns using artificial neural networks / S. K. Gauri, S. Chakraborty// International Journal of Advanced Manufacturing Technology. –2008. –№ 36. – PP. 1191-1201.
17. Eamonn Keogh Segmenting time series: a survey and novel approach [Електронний ресурс] / Eamonn Keogh, Selina Chu, David Hart, Michael Pazzani. Режим доступу до файлу: <http://www-scf.usc.edu/~selinach/segmentation-slides.pdf>
18. Guh R. S. On-Line identification of control chart patterns using selforganizing approaches / Guh R.S., and Shiue, Y.R. // International Journal of Production Research. – 2005. №43. – PP. 1225-1254.
19. Guh R.S. On-Line identification and quantification of mean shifts in bivariate processes using a NN-Based approach/ R. S. Guh // Quality and Reliability Engineering International.– 2007. – № 23. – PP. 367-385.
20. Sameer Singh Dynamic time-series forecasting using local approximation / Sameer Singh, Paul McAtackney // In Proceedings of the Tenth IEEE International Conference on Tools with Artificial Intelligence. – 1998. – PP. 392 –399.
21. David S. Stoffer Detecting common signals in multiple time series using the spectral envelope / David S. Stoffer // Journal of the American Statistical Association. – 1999. – PP. 1341–1356.
22. Online data mining for co-evolving time series / [Byoung-Kee Yi, Nikolaos D. Sidiropoulos, Theodore Johnson, H. V. Jagadish, Christos Faloutsos, and Alexadros Biliris] // In Proceedings of the Sixteenth International Conference on Data Engineering. – 2000. – PP. 13-22.
23. Hara K. A training data selection in online-training for multilayer neural networks / K. Hara, K. Nakayama, A. A. Kharaf // Proc. IEEE IJCNN’1998. –1998. – PP.2247-2252.
24. Klevecka I. Leli Pre-Processing of input data of neural networks: the case of forecasting telecommunication network traffic [Електронний ресурс] / I. Klevecka, Leli Janis. Режим доступу до файлу: http://www.telenor.com/en/resources/images/168-178_Pre-processingInput Data-ver1_tcm28-36193.pdf
25. Wang C.H. An Integrated Approach for process monitoring using wavelet analysis and competitive neural network / C.H.Wang, W. Kuo, H.Qi // International journal of production research. –2007. – №45. – PP.227-244.

Відомості про авторів

Квстний Роман Наумович – д.т.н., проф., завідувач кафедри АІВТ. Вінницький національний технічний університет, м. Вінниця, вул. Хмельницьке шосе 95, (0432) 598243.

Дружиніна Ольга Олегівна – аспірант кафедри АІВТ. Вінницький національний технічний університет, м. Вінниця, вул. Хмельницьке шосе 95, (0432) 598243, oo.druzhinina@gmail.com.

ПРИЛАДИ І МЕТОДИ КОНТРОЛЮ ТА ВИЗНАЧЕННЯ СКЛАДУ РЕЧОВИНИ

УДК 693.54

В. Ю. КУЧЕРУК, І. А. ДУДАТЬСВ

Вінницький національний технічний університет, Вінниця

МЕТОД КОМПЕНСАЦІЇ ВОЛОГОСТІ ДИМОВИХ ГАЗІВ КОТЕЛЬНИХ УСТАНОВОК З ПОКРАЩЕНИМИ МЕТРОЛОГІЧНИМИ ХАРАКТЕРИСТИКАМИ

Анотація: Синтезована функціональна схема системи вимірювання концентрації димових газів котельних установок з додатковим вимірювальним каналом визначення вологості газу. Представлена математична модель аналізатора вологості газових сумішей, в основу якої покладено метод абсорбційної спектроскопії, в якому лежить явище вибіркового поглинання багатоатомними молекулами випромінювання при його проходженні через газове середовище. Наведено структурну схему двоканального аналізатора вологості, яка побудована на основі оптичних схем введення ви-промінювання вимірювального та опорного каналів, світлодіодів, які випромінюють монохроматичне світло відповідної частоти, оптичних схем виведення випромінювання вимірювального та опорного каналів і фотоприймачів. Отримані нові, науково-обгрунтовані теоретичні та практичні результати, які є істотними для підвищення достовірності при необхідній швидкодії процесу контролю складу димових газів котельних установок на основі оптико-абсорбційного інфрачервоного методу з компенсацією дестабілізуючих факторів інфрачервоного перетворювача, за рахунок введення додаткового аналізатора вологості.

Ключові слова: статична характеристика, динамічна характеристика, перехідний процес, димовий газ, газоаналітична система, дестабілізуючі фактори, компенсаційний вимірювальний канал.

Аннотация: Разработана функциональная схема системы измерения концентрации димовых газов котельных установок с дополнительным измерительным каналом определения влажности газа. Представленная математическая модель анализатора влажности газовых смесей, в основу которой положен метод абсорбционной спектроскопии, в котором лежит явление выборочного поглощения многоатомными молекулами излучения при его прохождении через газовую среду. Приведена структурная схема двухканального анализатора влажности, которая построена на основе оптических схем введения излучения измерительного и опорного каналов, светодиодов, которые излучают монохроматический свет соответствующей частоты, оптических схем выведения излучения измерительного и опорного каналов и фотоприемников. Получены новые, научно-обоснованные теоретические и практические результаты, которые являются существенными для повышения точности при необходимом быстродействии процесса контроля состава димовых газов котельных установок на основе оптико-абсорбционного инфракрасного метода с компенсацией дестабилизирующих факторов инфракрасного преобразователя, за счет введения дополнительного анализатора влажности.

Ключевые слова: статическая характеристика, динамическая характеристика, переходный процесс, димовой газ, газоаналитическая система, дестабилизирующие факторы, компенсационный измерительный канал.

Annotation: The functional diagram of the system of measuring of concentration of smoke gases of caldrion options is worked out with the additional measuring channel of determination of gas humidity. Presented mathematical model of analyzer of humidity of gas mixtures, the method of absorption spectroscopy, in which the phenomenon of selective absorption lies by the polyatomic molecules of radiation at his passing through an environment, is fixed in basis of which. A flow diagram over of twochannel analyzer of humidity, which it is built on the basis of optical charts of introduction of radiation of measuring and supporting channels, light-emitting diodes, which radiate monochromatic light of corresponding frequency, optical charts of leading out of radiation of measuring and supporting channels and приемчів, is brought. New, scientifically-reasonable theoretical and practical results which are substantial for the increase of exactness at the necessary fast-acting of process of control of composition of smoke gases of caldrion options on the basis of optic-absorption of infra-red method with indemnification of destabilizing factors of infra-red transformer are got, due to introduction of additional analyzer of humidity.

Key words: static description, dynamic description, transient, smoke gas, газоаналітична system, destabilizing factors, compensative measuring channel.

Вступ

Ефективність роботи котельних установок (КУ) прямо залежить від наявності достовірної інформації про хід технологічних процесів. Відсутність контролю-вимірювальної апаратури може спричинити неефективну роботу установок, зокрема неякісне згорання палива.

Під час аналізу вологості технологічних газів найбільш інформативною величиною, яка характеризує безпосередньо кількість води в заданому об'ємі, є абсолютна вологість або відносна вологість, що визначає споживчі властивості газів [1]. Контроль вологості димових газів котельних установок дає змогу оцінити: ступінь завершеності процесу згорання палива (втрат від хімічної неповноти горіння), умови згорання палива (коефіцієнт надлишку повітря), характер згорання палива в окремих зонах котла (наявність локальних низькотемпературних зон), динаміку процесу горіння. В свою чергу, визначення даних факторів дає змогу максимально точно корегувати співвідношення паливо-повітря на вході об'єкта.

Мета роботи

Метою роботи є підвищення швидкодії при необхідній достовірності контролю складу димових газів котельних установок на основі оптико-абсорбційного інфрачервоного методу вимірювання.

Аналіз попередніх досліджень та публікацій

У роботі [2] представлено систему автоматичного керування котельною установкою у якій співвідношення "паливо-повітря" регулюється підсистемою контролю складу димових газів котельних установок. Підсистема контролю реалізована з закритим оптичним каналом, тобто проба, що аналізується, проходить через блок стабілізації входних параметрів (БСВП), яка очищує пробу від механічних домішок, зменшує до необхідного значення вологість суміші, стабілізує температуру та тис-

ку. На наступному етапі проба газу проходить через первинний та вторинний вимірювальні перетворювачі (ПВП та ВВП). Потім сигнал підсилюється до уніфікованого значення в основному підсилювачі і поступає в пристрій обробки інформації (ПОІ). У ПОІ здійснюється перетворення змінного електричного сигналу у значення концентрації компонента димового газу у %, відповідно функціональній залежності, в основі якої лежить закон поглинання Ламберта-Бера:

$$I_{вих} = a \cdot [1 - \exp(-b \cdot \varphi)], \tag{1}$$

де $I_{вих}$ – вихідний електричний сигнал (струм), a, b – постійні, φ - масова концентрація компонента газової суміші.

Одже, завдяки введеному у склад газоаналітичної системи (ГС) БСВП компенсуються вище перераховані дестабілізуючі фактори. Але існує неоднозначність: при зменшенні статичної складової похибки збільшується динамічна складова (при проходженні проби через БСВП (1)), і навпаки при зменшенні динамічної складової збільшується статична складова (з відкритим оптичним каналом (2)).



Рисунок 1 – Шляхи подачі димового газу на ГС, для його аналізу

Дійсно, похибка в усталеному режимі є статичною похибкою. Наявність перехідного процесу зумовлює динамічну похибку. Тривалість перехідного процесу при заданій динамічній похибці визначають з рівняння [3]:

$$\Delta_{дн} = k \cdot e^{-\left(\frac{t_{пн}}{\tau}\right)}, \tag{2}$$

де $t_{пн}$ – тривалість перехідного процесу, τ – стала часу, $\Delta_{дн}$ – динамічна похибка засобу вимірювання, k – коефіцієнт передачі.

З останнього рівняння тривалість перехідного процесу при заданому значенні динамічної похибки визначається за формулою:

$$t_{пн} = \tau \cdot \ln\left(\frac{k}{\Delta_{дн}}\right). \tag{3}$$

Таким чином, у засобах вимірювань з аперіодичним перехідним процесом, який описується перехідною першого порядку, динамічна похибка зменшується разом із зменшенням тривалості перехідного процесу. Протікання перехідного процесу у ВП після зміни концентрації газової суміші представлено на рис. 2 та рис. 3. Перехідна характеристика $h(t)$ побудована в середовищі символічної математики Maple VI.

При $t \approx$ від 0 до 4,5 с (рис.2) – це час, затрачений на транспортування проби та обумовлений інерційністю елементів ГС. При $t \approx$ від 0 до 3 с (рис. 3) – це час, обумовлений інерційністю елементів ГС.

Отже, існує проблема $\Delta_{дн1} > \Delta_{дн2}$ при $\Delta_{ст1} < \Delta_{ст2}$, і, навпаки, $\Delta_{дн1} < \Delta_{дн2}$ при $\Delta_{ст1} > \Delta_{ст2}$.

Розв’язок даної проблеми розпочато у [4], де у ГС введено додатковий канал компенсації температурної похибки. В результаті з БСВП було видалено блок стабілізації температури, а вихідний сигнал з ГС обраховується як $U_{вих} = U(\varphi) + U(T) + U'(T)$, де $U_{вих}$ – загальний вихідний сигнал з вимірювального каналу, $U(\varphi)$ – частка вихідного сигналу від концентрації компонента газової суміші, $U(T)$ – температурна похибка оптичного перетворення виражена через паразитну напругу, $U'(T)$ – сигнал з компенсаційно-

го температурного датчика. Таким чином, температурна похибка знищувалась програмним шляхом при забезпеченні умови $U(T) = -U'(T)$.

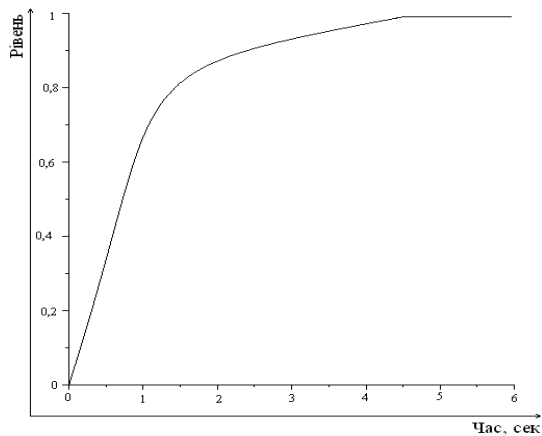


Рисунок 2 – Перехідний процес у ВП з закритим оптичним каналом

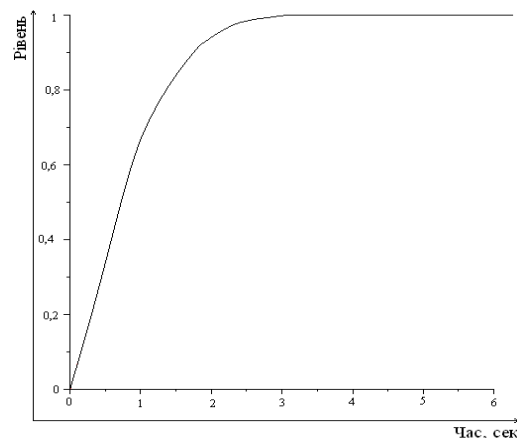


Рисунок 3 – Перехідний процес у ВП з відкритим оптичним каналом

Основна частина

Інфрачервону радіацію поглинають всі гази, за винятком O_2 , N_2 , H_2O , Cl_2 і одноатомних газів. Спектр поглинання одноатомних газів або пари металів відрізняється від інфрачервоних спектрів поглинання молекул своєю відносною простотою і складається не із смуг, а з окремих ліній, у багатьох випадках розташованих тільки в ультрафіолетовій області спектру. Смуги поглинання димового газу інфрачервоною радіацією представлено на рис. 4.

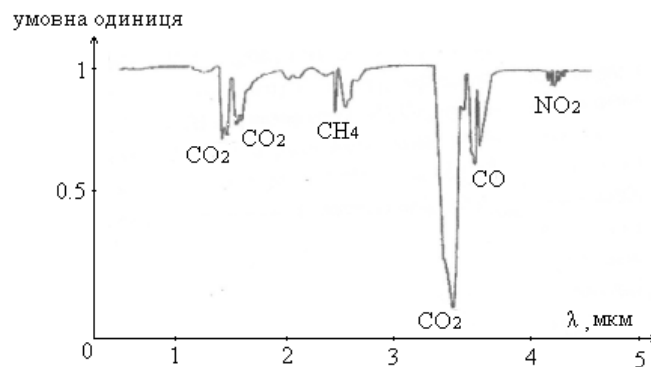


Рисунок 4 – Коливально-обертальні смуги поглинання димового газу ІЧ випромінюванням

Сума основних компонентів димових газів визначається по формулі 5 (решта компонентів димових газів є на рівні мікроконцентрацій і практично не впливають на визначення втрат тепла)

$$\mu(O_2) + \mu(NO_2) + \mu(CO_2) + \mu(CH_4) + \mu(CO) + \mu(H_2O) + \text{сажа} = 100\%, \quad (5)$$

де $\mu(X)$ – масова концентрація газової суміші у відсотках.

Оптимізацію процесу спалювання пропонується вести за компонентом – O_2 . Це дає змогу максимально точно корегувати співвідношення паливо-повітря на вході об'єкта. Але фізично кисень не поглинається ІЧ випромінюванням (це видно з рис. 5, де ліній поглинання кисню взагалі немає), тому пропонується визначати CO , CO_2 , NO_2 практично, а O_2 – аналітично (при вилученні з БСВП блока стабілізації температури проби та блока осушки проби)

$$\mu(O_2) = 100\% - [\mu(NO_2) + \mu(CO_2) + \mu(CO) + \mu(CH_4) + \mu(H_2O)]. \quad (6)$$

У роботі [1] розроблено математичну модель аналізатора вологості газових сумішей. На основі експериментальних досліджень визначено, що аналізатор відповідає таким вимогам: швидкий відгук, відсутність впливу агресивних домішок, а також перевертання результатів вимірювання газу від впливу спиртових з’єднань, низька похибка та висока відтворюваність результатів вимірювання.

В основу аналізатора покладено метод абсорбційної спектроскопії, в якому лежить явище вибіркового поглинання багатоатомними молекулами випромінювання при його проходженні через середовище.

Запропонований аналізатор вологості, який містить два рівноцінних канали: перший – вимірювальний, налаштований на робочу довжину хвилі (ν_1), другий – на опорну (ν_2). Абсолютна вологість газу визначається за формулою

$$\varphi = \frac{\mu_{e.n.} p_c T Z \cdot \left(K_{(\nu_2)}^{C.G.} \ln \frac{f_1 U_{01} K_1 U_{on1}}{F_{01}} - K_{(\nu_1)}^{C.G.} \ln \frac{f_2 U_{02} K_2 U_{on2}}{F_{02}} \right)}{dT_c p_{e.n.(max)} \left(K_{(\nu_1)}^{B.П.} \cdot K_{(\nu_2)}^{C.G.} - K_{(\nu_1)}^{C.G.} \cdot K_{(\nu_2)}^{B.П.} \right) \cdot \rho_{e.n.(c)}}, \quad (7)$$

де φ – абсолютна вологість газу; $\mu_{e.n.}$ – молекулярна маса водяної пари; p_c , T_c – тиск і температура газу за стандартних умов; T – температура газу; Z – коефіцієнт стисливості; $K_{(\nu)}^{C.G.}$, $K_{(\nu)}^{B.П.}$ – питомі молярні показники поглинання природного сухого газу та водяної пари на частоті ν відповідно; f_1 , f_2 – частота вимірювального і опорного каналів; U_{on1} , U_{on2} – опорна напруга першого і другого каналів, яка виставляється як початкова освітленість фотоприймача; F_{01} , F_{02} – початкова (нульова) частота першого і другого каналів; K_1 , K_2 – коефіцієнти перетворення першого і другого каналів; d – товщина шару газу, через який проходить пучок інтенсивності світла I_0 ; $p_{e.n.(max)}$ – максимальний тиск водяної пари; $\rho_{e.n.}^{B.П.}$ – густина водяної пари за стандартних умов.

Проведено дослідження, які показали, що залежність вологості газу від зміни частоти носить лінійний характер завдяки використанню оберненого зв’язку. Перевагою розробленої математичної моделі є те, що за її допомогою можна визначити абсолютну вологість газу без урахування густини суміші газів, його молекулярної маси, тиску.

На рис. 5 наведено структурну схему двоканального аналізатора вологості. Сенсор складається з корпусу, оптичних схем виведення випромінювання вимірювального та опорного каналів, світлодіодів, які випромінюють монохроматичне світло відповідної частоти, оптичних схем виведення випромінювання вимірювального та опорного каналів і фотоприймачів.

Корпус аналізатора представляє собою трубу, діаметр якої зумовлений габаритними розмірами оптичної системи, що забезпечує оптимальні параметри світлового потоку.

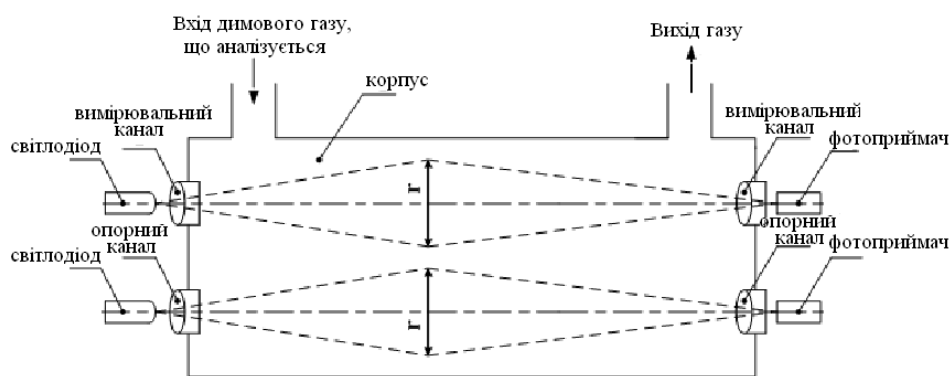


Рисунок 5 – Аналізатор вологості

Використаємо даний аналізатор, як окремий вимірювальний канал у системі контролю складу димових газів котельних установок. Синтезована функціональна схема системи контролю представлена на рис. 6.

За формулою (7) обчислюється кількість водяної пари, яка фактично міститься в 1 м^3 повітря (газової суміші), та обчислюється у одиницях $[\text{г}/\text{м}^3]$. Сумарний об’єм кювет (для представленої ГС), що використовуються для визначення компонента газової суміші дорівнює, $0,02 [\text{г}/\text{м}^3]$. На наступному

етапі об'єм кювет приймаємо як 100%, а вимірне $0,02\varphi - x\%$. Розв'язавши пропорцію, отримаємо значення φ_{H_2O} яке можна підставляти у основні формули (5 та 6).

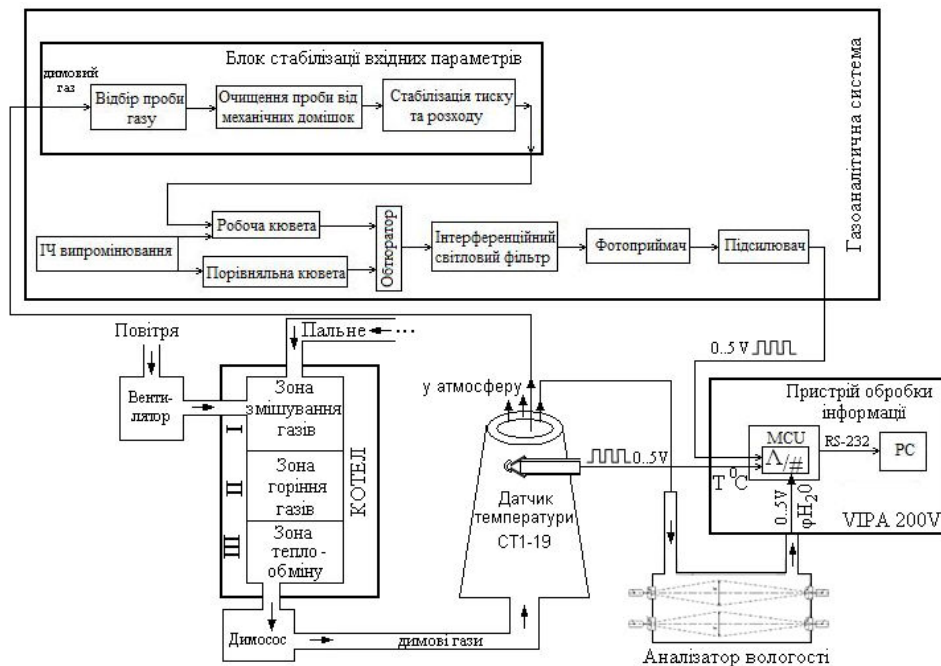


Рисунок 6 – Газоаналітична система

Для компенсації температурної похибки використовуємо напругу (В), а вологості безпосередньо кількість у заданому об'ємі (%). У подальшому планується дослідити вплив механічних домішок на результати вимірювання концентрації та розробити метод компенсації даних факторів підвищивши швидкість при необхідній точності контролю складу димових газів котельних установок.

Висновок

Отримані нові, науково обгрунтовані теоретичні та практичні результати, які є істотними для підвищення достовірності при необхідній швидкодії процесу контролю складу димових газів котельних установок на основі оптико-абсорбційного інфрачервоного методу з компенсацією дестабілізуючих факторів інфрачервоного перетворювача, за рахунок введення додаткового аналізатора вологості.

Список літератури

1. Білинський Й.Й., Онушко В.В. Математична модель аналізатора вологості газу. / Й.Й. Білинський, В.В. Онушко // Наукові праці ВНТУ. – м. Вінниця, 2010. – №4.
2. Кучерук В.Ю., Дудатьєв І.А. Ресурсоенергозбережна система автоматичного керування котельною установкою з контролем складу димових газів / В.Ю. Кучерук, І.А. Дудатьєв // Вісник інженерної академії України. – м. Київ, 2010. – №3-4. – С. 98-103.
3. Володарський Є.Т., Кухарчук В.В., Поджаренко В.О., Сердюк Г.Б. Метрологічне забезпечення вимірювань і контролю. Навчальний посібник. – Вінниця ВДТУ, 2001.
4. Кучерук В.Ю., Дудатьєв І.А. Метод компенсації температурної похибки оптико-абсорбційних температурних перетворювачів концентрації газової суміші / В.Ю. Кучерук, І.А. Дудатьєв // Оптико-електронні інформаційно-енергетичні технології, 2011. – №1 (21). – С. 163-167.

Відомості про авторів

Кучерук Володимир Юрійович – д.т.н., професор, зав. кафедри метрології та промислової авто-матики. Вінницький національний технічний університет; 21021 м.Вінниця, Хмельницьке шосе, 95; тел. (0432)598672; e-mail: kucheruk@mail.ru.

Дудатьєв Ігор Андрійович – аспірант кафедри метрології та промислової автоматики. Вінницький національний технічний університет; 21021 м.Вінниця, Хмельницьке шосе, 95; тел. (0432)598672; e-mail: dudatiev@mail.ru.