

БІОЛОГІЧНІ ТА МЕДИЧНІ ПРИЛАДИ І СИСТЕМИ

УДК 616.24-085.816 (045)

Л.О.КОШЕВА, І.В.ФЕДОРЕНКО

Національний авіаційний університет, Київ

ЗАБЕЗПЕЧЕННЯ ОПТИМАЛЬНИХ ПОКАЗНИКІВ ШТУЧНОЇ ВЕНТИЛЯЦІЇ ЛЕГЕНІВ

Анотація. Запропоновано підхід до реалізації режиму штучної вентиляції легень, що полягає у досягненні оптимальних показників дихання. Наведено структурну схему апарату штучної вентиляції легень з реалізацією запропонованого режиму та описано алгоритм його роботи.

Ключові слова: штучна вентиляція легень, адаптивний режим, респіратор, контрольована та допоміжна вентиляція, хвилинна вентиляція, оптимальні показники.

Аннотация. Предложен подход к реализации режима искусственной вентиляции легких, которая заключается в достижении оптимальных показателей дыхания. Приведена структурная схема аппарата искусственной вентиляции легких с реализацией предложенного режима и описаны алгоритм его работы.

Ключевые слова: искусственная вентиляция легких, адаптивный режим, респиратор, контролируемая и вспомогательная вентиляция, минутная вентиляция, оптимальные показатели.

Abstract. An approach to the implementation mode of mechanical ventilation, which is to achieve the optimal performance of breathing. Shows the block diagram of the ventilator with the implementation of the proposed regime and describes the algorithm of its work.

Key words: ventilator adaptive mode, respirator, controlled and auxiliary ventilation, trigger, supporting pressure, minute ventilation.

Вступ

Штучна вентиляція легень (ШВЛ) здійснюється за допомогою спеціального апарату – респіратора, який призначено для індивідуального захисту органів дихання та контролю дотримання заданих лікарем параметрів. Режим ШВЛ – це алгоритм управління потоком киснево-повітряної суміші у дихальному контурі. Запорукою раціональної респіраторної підтримки є уважне спостереження за хворим, врахування особливостей патологічного процесу і чітке розуміння технічних деталей реалізації режимів і алгоритмів ШВЛ.

У теперішній час лікувальні установи різного рангу мають на оснащенні відділень інтенсивної терапії апарати ШВЛ, парк яких невпинно збільшується. Сучасний апарат ШВЛ є складним пристроєм, що потребує спеціальних знань лікарів та обслуговуючого персоналу і пропонує лікарю не альтернативні варіанти проведення ШВЛ, а безперервну гамму режимів. Мета використання різних режимів і алгоритмів ШВЛ – здійснення індивідуального підходу до конкретної клінічної ситуації. Різноманіття режимів та контрольованих параметрів ускладнює роботу лікаря, коли необхідно швидко визначити оптимальний режим дихання на поточний момент стану та потреби пацієнта. У зв'язку з цим виникає необхідність у реалізації такого режиму, який забезпечить оптимальні показники штучної вентиляції легень з урахуванням поточного стану пацієнта. Для проведення раціональної респіраторної підтримки необхідне розуміння не лише фізіологічних особливостей хворого, але і деталей реалізації режимів ШВЛ в апаратах різних класів і моделей.

Сучасний стан питання

В даний час спостерігається принципова зміна поглядів на респіраторну підтримку [1]. Так, спостерігається тенденція до повної відмови від нетригерованої вентиляції з максимальною увагою до збереження спонтанного дихання пацієнта; особлива увага приділяється попередженню пошкодження легень через нераціональний вибір параметрів ШВЛ; відмова від прагнення до нормалізації газообміну та інших показників гомеостазу на користь так званих стрес-норм. Сучасні апарати штучної вентиляції легень відрізняються орієнтацією на допоміжні режими вентиляції, наявністю мікропроцесорного управління всіма параметрами вентилятора, розширеними можливостями моніторингу параметрів респіраторної механіки пацієнта, а також розвинутою системою тривоги для відстеження небезпечних відхилень. Таким чином, сучасні апарати забезпечують підтримку дихальної функції пацієнта протягом довгого часу [1,2].

Постановка проблеми

Метою ШВЛ, як відомо, є забезпечення нормальної вентиляції легень і оксигенації артеріальної крові, що досягається шляхом вдування газових сумішей з високим вмістом кисню. Звичним підходом до режимів ШВЛ є поділ вентиляції легень на контрольовану та допоміжну. Допоміжна вентиляція передбачає включення тригера і подачу механічних вдихів з урахуванням самостійного дихання пацієнта [3]. При контрольованій вентиляції відсутня можливість аналізу респіратором дихальних спроб хворого. Якщо поставлено завдання подати в легені певний дихальний об'єм газів (киснево-повітряної суміші), то вентиляція є контрольованою за об'ємом. Якщо респіратор створює тиск в дихальних шляхах, то вентиляція є контрольованою за тиском. Очевидно, що в обох випадках в дихальні шляхи подається потік киснево-повітряної суміші, що має певний об'єм, і створює в легенях тиск залежно від механічних влас-

тивостей легень. Сучасні респіратори можуть забезпечувати суміщення контрольованих вентиляцій як за тиском, так і за об'ємом [4,5].

При формуванні вдиху з контролем за об'ємом респіратор подає потік в легені пацієнта і перемикається на видих при досягненні заданого лікарем об'єму вдиху (дыхального об'єму). Недоліками такого режиму є небезпека баротравми (пошкодження органів, що містять повітря або вуглекислий газ, викликані різкими змінами атмосферного тиску); нерівномірність вентиляції різних відділів легень; неможливість адекватної вентиляції без частого контролю заданих показників. При формуванні вдиху з контролем за тиском респіратор подає потік в легені пацієнта і перемикається на видих при досягненні заданого лікарем тиску. Недоліками такого режиму є відсутність гарантованого дыхального об'єму, необхідність повного моніторингу вентиляції та частого контролю заданих показників. Запропонований режим надасть можливість зменшення цих недоліків шляхом контролю необхідних та достатніх для цього параметрів: маси тіла пацієнта (M_t), відсоток обов'язкової (примусової) хвилинної вентиляції (V_{\min} , %), максимально допустимий тиск в легенях (P_{\max}).

Опис структурної схеми

Вибір режиму ШВЛ залежить від цілей (ШВЛ при анестезії, при патології легень або механічного апарату дихання тощо), від стану хворого та можливостей лікарні. Контрольована та допоміжна вентиляції реалізуються за класичною схемою апарату ШВЛ. Для впровадження адаптивного режиму штучної вентиляції легень пропонується ввести до класичної структурної схеми апарату ШВЛ блок автоматичної адаптації (БАА) до пацієнта, що наведено на рисунку 1.

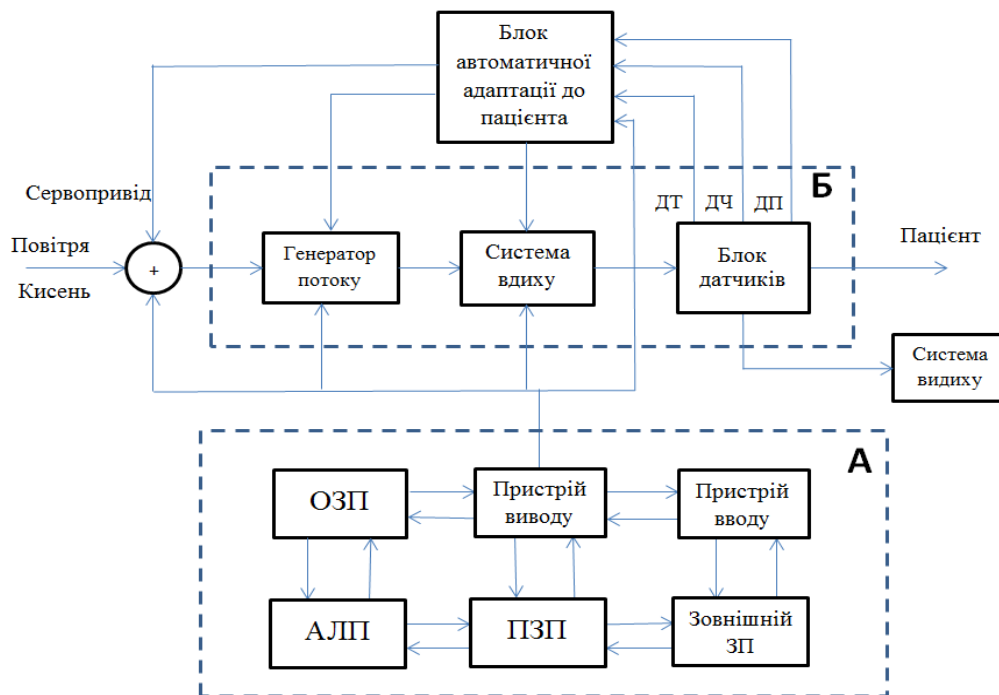


Рисунок 1– Спрощена структурна схема апарату ШВЛ, що забезпечує адаптивний режим

Вхідні сигнали до БАА поступають від датчиків об'єму потоку, тиску та частоти дихання, порівнюються з вхідними параметрами, які задаються лікарем і виробляють керуючі сигнали, що подаються на відповідні вузли приладу. На відміну від традиційних режимів ШВЛ запропонований режим працює тільки з заданням трьох параметрів: M_t (за 100% приймається «фізіологічна» хвилинна вентиляція, рівна 100 мл / кг / хв у дорослих і 200 мл / кг / хв у дітей); V_{\min} , % ; P_{\max} . Після цього апарат автоматично вибирає тип вентиляції (контрольована або допоміжна) і розраховує необхідні оптимальні параметри ШВЛ.

Особливістю адаптивного режиму вентиляції є постійна інтерактивна взаємодія в системі «режими ШВЛ- апарат-хворий» з автоматичним регулюванням трьох параметрів вентиляції для досягнення тієї чи

іншої заданої мети (наприклад, заданого дихального або хвилинного об'єму дихання). Режим може бути реалізований у сучасних апаратах ШВЛ, що дасть змогу з мінімальним втручанням лікаря контролювати та покращувати стан пацієнта [6]. У такий спосіб блок автоматичної адаптації надасть змогу:

- оцінити легеневу механіку пацієнта;
- оптимізувати співвідношення дихальний об'єм / частота дихання з урахуванням легеневої механіки;
- забезпечити оптимальне поєднання дихального об'єму / частоти дихання за рахунок автоматичної підтримки цільового обсягу хвилинної вентиляції при дотриманні правил захисту легенів;
- автоматично контролювати стан, і корегувати показники дихання в залежності від стану пацієнта.

Алгоритм адаптивного режиму

Алгоритм, реалізація якого забезпечує автоматичний розрахунок обов'язкової апаратної хвилинної вентиляції, самостійне визначення найбільш оптимальних показників: інспіраторного тиску (керованого або підтримуючого) та частоти дихання для досягнення заданої мети, наведено на рисунку 2.

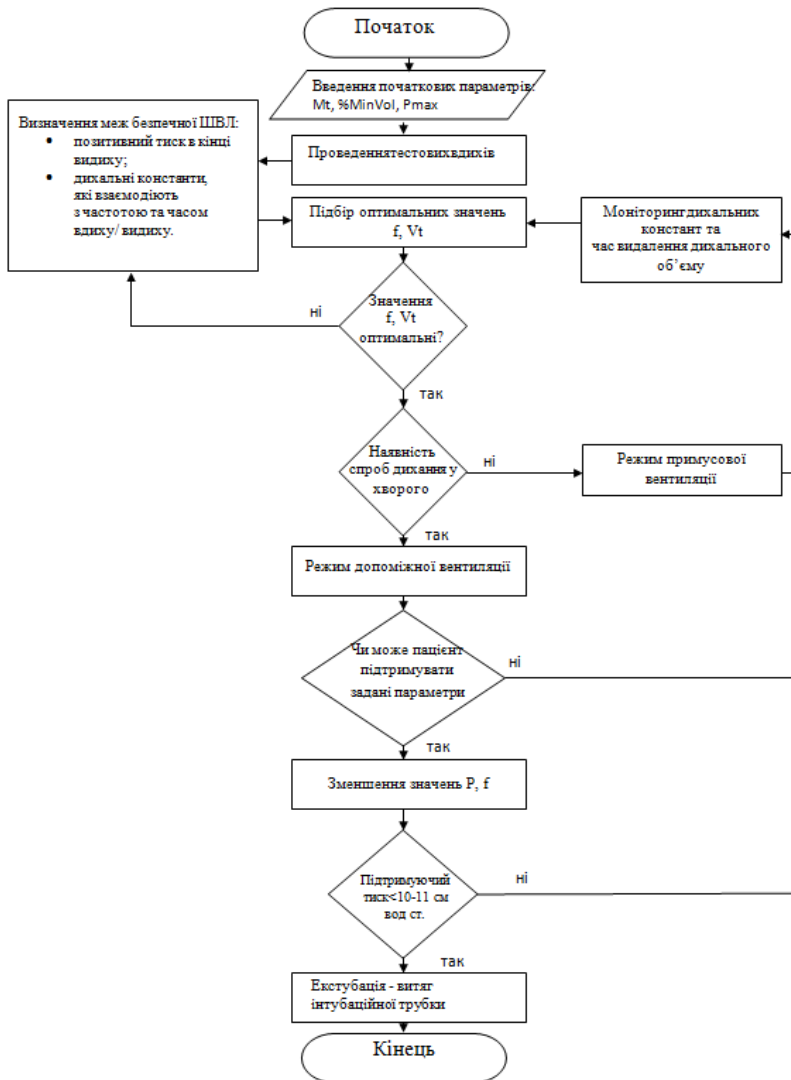


Рисунок 2 –Алгоритм адаптивного режиму

Даний алгоритм працює із заданням наведених показників: M , V_{\min} , P_{\max} у блоці автоматизації. На початку роботи апарат, підключений до хворого, реалізує програму тестових вдихів, за якими визначаються такі межі безпечної вентиляції як позитивний тиск в кінці видиху та дихальні константи, які пов'язані з частотою та часом вдиху/ видиху. За результатами тестових вдихів визначається стан пацієнта, відповідно до якого підбираються оптимальні значення частоти та об'єму дихання. Якщо показники частоти та об'єму дихання не набули оптимальних значень, то БАА, що постійно проводить моніторинг дихальних констант, повертається до підбору оптимальних значень, коригуючи їх, і знову визначаються межі безпечної вентиляції. При досягненні оптимальних значень встановлених показників, апарат підтримує наявні спроби дихання у хворого. Якщо відсутні спроби хворого самостійно дихати, або вони недостатні, то вмикається режим примусової вентиляції і проводиться моніторинг додаткового параметру – часу, протягом якого з легких видаляється дихальний об'єм. Всі параметри, які підлягають моніторингу, перевіряються на відповідність оптимальним значенням, і у разі їх невідповідності БАА повертається до підбору оптимальних значень. Якщо показники відповідають оптимальним значенням, тобто присутні спроби хворого самостійно дихати, тоді включається режим допоміжної вентиляції легень, який стимулює дихання хворого. Якщо пацієнт підтримує задані параметри, то апарат зменшує тиск в легенях та частоту вдихів. Підтримуючий тиск має бути меншим за 10-11 см.вод.ст [7]. Якщо ці показники задовольняються, то відбувається витяг респіраторної трубки лікарем. У разі, перевищення оптимального значення тиску, то БАА, проводячи постійний моніторинг дихальних констант, повертається до підбору оптимальних значень.

У такий спосіб запропонований алгоритм дозволяє постійно проводити моніторинг (контролювати та проводити оцінку) стану пацієнта, розраховуючи та підтримуючи оптимальні показники вентиляції легень для досягнення розрахункових (оптимальних) параметрів, та в максимально можливий короткий термін переводити пацієнта на самостійне дихання та відключати від апарату ШВЛ.

Висновки

Запропонований адаптивний режим штучної вентиляції легень забезпечує мінімальне втручання лікаря при здійсненні штучної вентиляції легень, автоматично підтримуючи оптимальні значення найважливіших показників дихання, що полегшує роботу лікаря з апаратом ШВЛ та забезпечує потреби дихання пацієнта шляхом постійного моніторингу поточних показників дихання.

Список використаної літератури

1. Clinical Application of Mechanical Ventilation by David W Chang D. W. Edition 2006. [Електронний ресурс]. Режим доступу: <http://www.amazon.com>
2. Mechanical Ventilation, MacIntyre N. & Branson R. by Neil R. MacIntyre and Richard D. Branson 2nd Edition, 2008. [Електронний ресурс]. Режим доступу: <http://www.amazon.com/Mechanical-Ventilation-Neil-R>
3. Сатишур О.Е. Механическая вентиляция легких / О.Е. Сатишур.- М.: Мед. лит., 2006.- 352 с.
4. Респіраторна підтримка: Керівництво по штучної та допоміжної вентиляції легень в анестезіології та інтенсивній терапії / [Кассиль В. Л., Лескінен Г. С., Вижігіна М. А.] - М.: Медицина, 1997. – 324 с.
5. Горячев А.С. Основы ШВЛ. Часть третья [Електронний ресурс]. Режим доступу: <http://anest.ugansk.ru>
6. Царенко С. В. Практичний курс штучної вентиляції легень/С. В.Царенко.- М., 2007.- 97 с.
7. Оценка местных принципов искусственной вентиляции легких [Електронний ресурс]. Режим доступу: <http://www.euroanaesthesia.org>

Відомості про авторів

Кошева Лариса Олександрівна – д-р техн. наук, доцент, професор кафедри біокібернетики та аерокосмічної медицини, Національний авіаційний університет, т.4067442, amis@ukrpost.net, Космонавта Комарова, 1, Київ 03058.

Федоренко Іванна Володимирівна – магістр, Національний авіаційний університет, т.4067442, fedorenko_ivanna@ukr.net, Космонавта Комарова, 1, Київ 03058.

УДК 615.471:616

PAVLOV S.V., AZAROV O.D., SALDAN I.R., ROZMAN A.O., BABYUK N.P.

Vinnitsya National Technical University, Vinnitsya
Vinnitsya National Medical University, Vinnitsya**FUZZY EXPERT OPTO-ELECTRONIC SYSTEM FOR THE ANALYSIS OF BIOMEDICAL IMAGES
(FOR EXAMPLE DIAGNOSING GLAUCOMA)**

Анотація. В статті отримано математичні моделі для оцінювання біомедичних зображень з використанням методів нечітких множин на основі експертних баз знань. Це дозволило проводити комплексне якісне діагностування та підвищити достовірність і оперативність діагностування. Розроблено методику аналізу біомедичних зображень на основі апарату нечітких множин, що дозволило повніше діагностувати захворювання пр. глаукомі.

Розроблено архітектуру інформаційної технології для підтримки прийняття рішень при аналізі біомедичної інформації, що дозволило підвищити достовірність оброблення біомедичних зображень шляхом врахування структурних особливостей біооб'єктів.

Ключові слова: опто-електронна система, біомедичні зображення, біооб'єкти.

Аннотация. В статье получены математические модели для оценки биомедицинских изображений с использованием методов нечетких множественных чисел основанных на экспертных базах знаний. Это дало возможность проводить комплексное качественное диагностирование и повысить достоверность и оперативность диагностирования. Разработана методика анализа биомедицинских изображений на основании аппарата нечетких множественных чисел, что позволило полнее диагностировать заболевания при глаукоме.

Разработана архитектура информационной технологии для поддержки принятия решений при анализе биомедицинской информации, что позволило повысить достоверность обработки биомедицинских изображений путем учета структурных особенностей биообъектов.

Ключевые слова: опто-электронная система, биомедицинские изображения, биообъекты.

Abstract. In the article mathematical models are got for the estimation of biomedical images with the using of methods of unclear plurals on the basis of expert bases of knowledges. It allowed to conduct complex high-quality diagnosticating and promote authenticity and operationability of diagnosticating. The method of analysis of biomedical images is developed on the basis of vehicle of unclear plurals, that allowed completer of diagnosing glaucoma.

Architecture of information technology is developed for support of making decision at the analysis of biomedical information, that allowed to promote authenticity of processing of biomedical images by the account of structural features of bioobjects.

Key words: opto-electronic system, biomedical images, bioobjects.

Introduction

In general, the input variables used in the expert system may be presented in qualitative or quantitative manner. Using expert system provided entry input variables, the transformation of quantitative variables in quality. This expert system based on fuzzy logic input implements the functions of collection, storage, correlation analysis and using the knowledge obtained by experts for diagnosing glaucoma.

Method

Input function knowledge base is presented in tabular form. For our expert system as a table of knowledge is presented in Table 1.

Database value. When entered into the table of terms used interface. Or at any point add and delete individual rows in the table. For knowledge representation in an expert system uses an object-oriented approach, which provides storage as objects. Using the same facilities for recording tree structure and knowledge base provides a uniform flow of information and simplify processing of results.

Storage unit membership functions allows for time for different input data more correct decisions that take into account previous medical research. This information from the library allows you to monitor the reliability of the results of the expert system and, if necessary, make recommendations on whether a more precise adjustment of the expert system.

User data are processed based on the matrix of knowledge through fuzzy inference algorithms. Given the fact that during clinical studies quite often need to use not only crisp digital criteria, but also certain linguistic characteristics change parameters (terms), we conducted an analysis of some of them by using mathematical fuzzy logic. This approach allows to obtain explicit numerical expression for the criteria that are descriptive characteristics and, consequently, quality content, such as the following features: L – low, BA – below average, M – medium, IM – intermediate, H – high [1,2].

Each of these terms is a fuzzy set, which is given with special membership functions, and can be made configurable interval that has digital degrees from 0 to 1. On absolutely not belonging to the set indicates 0, and the absolute identity - 1.

Fuzzy sets may be supplemented, unite, intersect that simplifies the formalization process to digital (clear) expression.

Table 1 – Function knowledge base

Severity of pathology	Intraocular pressure, mmHg	Field view	The ratio of the diameter to the diameter of the optic disc excavation	The ratio of the area of excavation to the area OD	Ratio of the volume of excavation to OD	Area of excavation	Volume of excavation	Area neyrorretinal belt	Volume neyrorretinal belt	The average thickness of the nerve fibers in limb	Cross-sectional area nerve fiber layer on the edge of the disc
d1 - standard	15-21	60 ⁰	0,009 – 0,635	0 – 0,42	0 – 1,1	0 – 0,95	0 – 0,295	1,097 – 2,14	0,165 – 0,7	0,25 – 0,52	1,203 – 2,63
d2 – I stage glaucoma	15-21	50 ⁰	0,55 – 0,8	0,3 – 0,62	0,2 – 2,6	0,55 – 1,77	0,04 – 0,52	0,79 – 2,04	0,12 – 0,53	0,18 – 0,46	0,86 – 2,1
d3 – II stage glaucoma	31	45 ⁰ -35 ⁰	0,5 – 0,92	0,25 – 0,82	0,2 - 8	0,7 – 2,04	0,1 – 0,74	0,32 – 2,04	0,07 – 0,49	0,14 – 0,4	0,7 – 1,86
d4 – III stage glaucoma	33	<35 ⁰	0,65 – 1,0	0,4 – 0,97	1,0 - 35	0,82 – 2,47	0,15 – 1,35	0,217 – 2,09	0,025 – 0,262	0 – 0,38	0 – 1,71

Table 2 – Database for diagnosing glaucoma

Severity of pathology	Intraocular pressure, mmHg	Field view	The ratio of the diameter to the diameter of the optic disc excavation	The ratio of the area of excavation to the area OD	Ratio of the volume of excavation to OD	Area of excavation	Volume of excavation	Area neyrorretinal belt	Volume neyrorretinal belt	The average thickness of the nerve fibers in limb	Cross-sectional area nerve fiber layer on the edge of the disc
d1	L	H	L	L	L	L	L	IM	BA	M	M
	BA	H	BA	BA	L	BA	BA	H	M	IM	IM
	BA	H	M	M	L	M	BA	H	IM	H	H
	BA	H	IM	M	L	M	BA	H	H	H	H
d2	L	M	M	BA	L	BA	L	M	BA	BA	BA
	BA	M	IM	M	L	M	BA	IM	M	M	M
	BA	M	IM	IM	L	IM	M	H	IM	IM	IM
	BA	M	IM	IM	L	IM	M	H	IM	H	IM
d3	H	L	M	BA	L	BA	L	L	BA	BA	BA
	H	BA	IM	M	BA	M	BA	BA	M	M	M
	H	M	H	IM	BA	IM	M	M	IM	IM	IM
	H	M	H	IM	BA	IM	IM	IM	IM	IM	IM
	H	M	H	IM	BA	IM	IM	H	IM	IM	IM
d4	H	L	IM	M	L	BA	L	L	L	L	L
	H	L	H	IM	BA	M	BA	BA	BA	BA	BA
	H	L	H	H	M	IM	M	M	M	M	M
	H	L	H	H	IM	H	IM	IM	M	IM	IM
	H	L	H	H	H	H	H	H	M	IM	IM

Application of mathematical fuzzy logic is appropriate in cases of necessity diagnosing glaucoma.

To diagnose glaucoma take into account such factors:

X₁ - Intra-ocular pressure, mmHg

X₂ - Field of view

X₃ - The ratio of the diameter to the diameter of the optic disc excavation

X₄ - The ratio of the area of excavation to the area OD (optic disc)

X₅ - The ratio of volume to the volume of excavation OD

X₆ - The area of excavation

X₇ - The volume of excavation

X₈ - The area neuroretinal belt

X₉ - The volume neuroretinal belt

X₁₀ - The average thickness of the nerve fibers in limb

X_{11} – The cross-sectional area nerve fiber layer on the edge of the disc.

An assessment database for diagnosing glaucoma. Each degree of change is presented as a scale qualitative terms: low (L), below average (BA), medium (M), intermediate (IM), high (H) (Table 2).

For each database in order to formalize the parameters determined by the appropriate membership function [1,2].

To construct the equations necessary to determine the membership function $\mu^j(x_i)$ of all fuzzy terms j (H, IM, M, BA, L) for all factors x_i (in this case j - coefficient symmetry, i - space exploration $i = \overline{1,4}$). If we consider the high level of alternative rules, it is necessary to build equations for the five fuzzy terms (H, IM, M, BA, L).

Each factor x_i must comply with its own five membership functions.

To simplify the modeling must take some action:

Let \underline{x}_i and \overline{x}_i - lower and upper bounds alterations x_i . Expression interval $[\underline{x}_i, \overline{x}_i]$ on the interval $U = [0, 4]$, which set membership function, $\tilde{\mu}^j(u)$, $u \in U$ to fuzzy terms $j = H, IM, M, BA, L$.

Graphical view of membership functions shown in Fig. 1. Choosing these curves due to the fact that they are piecewise linear approximations expert membership functions $\mu^j(x_i)$ obtained for the factors $x_1 \div x_4$ method of paired comparisons.

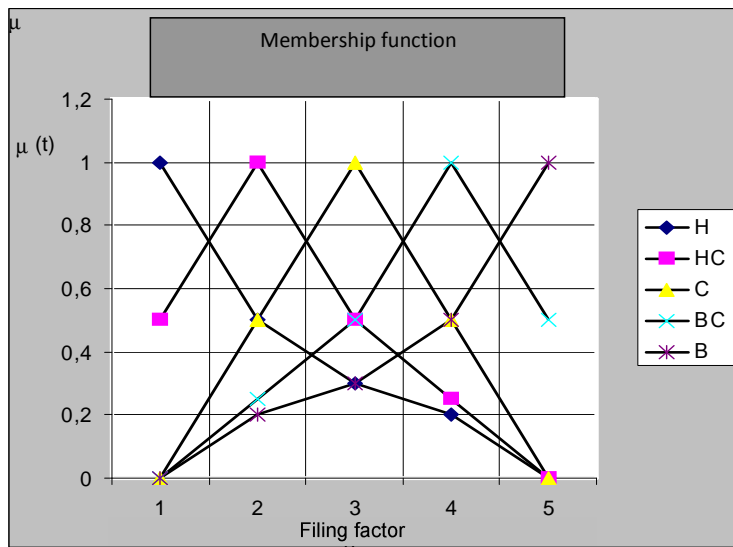


Figure 1 – The functions of fuzzy terms

The transition from function $\tilde{\mu}^j(u)$ to desired functions $\mu^j(x_i)$ is as follows:

$$u_i = 4 \frac{x_n - x_n}{x_n - x_n}, \tilde{\mu}^j(u_n) = \mu^j(x_n). \quad (1)$$

Deciding on the severity of the disease can be done by the following algorithm [1-3]:

Step 1: fix the value factors for a particular patient $x_n (n = \overline{1,4})$;

Step 2: The formulas (3) - (6) define the value of membership functions $\mu^j(x_n)$ for fixed values of factors x_n ;

Step 3: Using a logical equations compute the membership function $\mu^{d,n}(x_1, x_2, \dots, x_n)$ for all degrees of disease severity $d_n, n = \overline{1,6}$. In this transaction I (•) and OR (V) over membership functions $\mu(a)$ and $\mu(b)$ are replaced by operations min and max:

$$\mu(a) \bullet \mu(b) = \min[\mu(a), \mu(b)]; \mu(a) \vee \mu(b) = \max[\mu(a), \mu(b)];$$

Step 4: Determine the decision d_0 , for which:

$$\mu^{d_0}(x_1, x_2, \dots, x_n) = \max[\mu^{d_n}(x_1, x_2, \dots, x_n)]. \quad (2)$$

This decision will meet the desired range, which indicates the severity of the disease.

For example, the patient received the following information:

X_1 - Intra-ocular pressure, mmHg – 31,4

X_2 - Field of view – 32,1⁰

X_3 - The ratio of the diameter to the diameter of the optic disc excavation – 0,83

X_4 - The ratio of the area of excavation to the area OD (optic disc) – 0,54

X_5 - The ratio of volume to the volume of excavation OD – 6,42

X_6 - The area of excavation – 1,42

X_7 - The volume of excavation – 0,53

X_8 - The area neuroretinal belt – 0,74

X_9 - The volume neuroretinal belt – 0,1

X_{10} - The average thickness of the nerve fibers in limb – 0,09

X_{11} - The cross-sectional area nerve fiber layer on the edge of the disc – 0,49

Calculate the severity of disease:

for d_1

$$\begin{aligned} \mu^{d_1}(x_1 \div x_{11}) &= \mu^L(x_1) \cdot \mu^H(x_2) \cdot \mu^L(x_3) \cdot \mu^L(x_4) \cdot \mu^L(x_5) \cdot \\ &\cdot \mu^L(x_6) \cdot \mu^L(x_7) \cdot \mu^{IM}(x_8) \cdot \mu^{BA}(x_9) \cdot \mu^M(x_{10}) \cdot \mu^M(x_{11}) \vee \\ &\vee \mu^{BA}(x_1) \cdot \mu^H(x_2) \cdot \mu^{BA}(x_3) \cdot \mu^{BA}(x_4) \cdot \mu^L(x_5) \cdot \\ &\cdot \mu^{BA}(x_6) \cdot \mu^{BA}(x_7) \cdot \mu^H(x_8) \cdot \mu^M(x_9) \cdot \mu^{IM}(x_{10}) \cdot \mu^{IM}(x_{11}) \vee \\ &\vee \mu^{BA}(x_1) \cdot \mu^H(x_2) \cdot \mu^M(x_3) \cdot \mu^M(x_4) \cdot \mu^L(x_5) \cdot \\ &\cdot \mu^M(x_6) \cdot \mu^{BA}(x_7) \cdot \mu^H(x_8) \cdot \mu^{IM}(x_9) \cdot \mu^H(x_{10}) \cdot \mu^H(x_{11}) \vee \\ &\vee \mu^{BA}(x_1) \cdot \mu^H(x_2) \cdot \mu^{IM}(x_3) \cdot \mu^M(x_4) \cdot \mu^L(x_5) \cdot \\ &\cdot \mu^M(x_6) \cdot \mu^{BA}(x_7) \cdot \mu^H(x_8) \cdot \mu^H(x_9) \cdot \mu^H(x_{10}) \cdot \mu^H(x_{11}) \end{aligned} \quad (3)$$

for d_2

$$\begin{aligned} \mu^{d_2}(x_1 \div x_{11}) &= \mu^L(x_1) \cdot \mu^M(x_2) \cdot \mu^M(x_3) \cdot \mu^{BA}(x_4) \cdot \mu^L(x_5) \cdot \\ &\cdot \mu^{BA}(x_6) \cdot \mu^L(x_7) \cdot \mu^M(x_8) \cdot \mu^{BA}(x_9) \cdot \mu^{BA}(x_{10}) \cdot \mu^{BA}(x_{11}) \vee \\ &\vee \mu^{BA}(x_1) \cdot \mu^M(x_2) \cdot \mu^{IM}(x_3) \cdot \mu^M(x_4) \cdot \mu^L(x_5) \cdot \\ &\cdot \mu^M(x_6) \cdot \mu^{BA}(x_7) \cdot \mu^{IM}(x_8) \cdot \mu^M(x_9) \cdot \mu^M(x_{10}) \cdot \mu^M(x_{11}) \vee \\ &\vee \mu^{BA}(x_1) \cdot \mu^M(x_2) \cdot \mu^{IM}(x_3) \cdot \mu^{IM}(x_4) \cdot \mu^L(x_5) \cdot \\ &\cdot \mu^{IM}(x_6) \cdot \mu^M(x_7) \cdot \mu^H(x_8) \cdot \mu^{IM}(x_9) \cdot \mu^{IM}(x_{10}) \cdot \mu^{IM}(x_{11}) \vee \\ &\vee \mu^{BA}(x_1) \cdot \mu^M(x_2) \cdot \mu^{IM}(x_3) \cdot \mu^{IM}(x_4) \cdot \mu^L(x_5) \cdot \\ &\cdot \mu^{IM}(x_6) \cdot \mu^M(x_7) \cdot \mu^H(x_8) \cdot \mu^{IM}(x_9) \cdot \mu^H(x_{10}) \cdot \mu^{IM}(x_{11}) \end{aligned} \quad (4)$$

for d_3

$$\begin{aligned} \mu^{d_3}(x_1 \div x_{11}) &= \mu^H(x_1) \cdot \mu^L(x_2) \cdot \mu^M(x_3) \cdot \mu^{BA}(x_4) \cdot \mu^L(x_5) \cdot \\ &\cdot \mu^{BA}(x_6) \cdot \mu^L(x_7) \cdot \mu^L(x_8) \cdot \mu^{BA}(x_9) \cdot \mu^{BA}(x_{10}) \cdot \mu^{BA}(x_{11}) \vee \\ &\vee \mu^H(x_1) \cdot \mu^{BA}(x_2) \cdot \mu^{IM}(x_3) \cdot \mu^M(x_4) \cdot \mu^{BA}(x_5) \cdot \\ &\cdot \mu^M(x_6) \cdot \mu^{BA}(x_7) \cdot \mu^{BA}(x_8) \cdot \mu^M(x_9) \cdot \mu^M(x_{10}) \cdot \mu^M(x_{11}) \vee \\ &\vee \mu^H(x_1) \cdot \mu^M(x_2) \cdot \mu^H(x_3) \cdot \mu^{IM}(x_4) \cdot \mu^{BA}(x_5) \cdot \end{aligned}$$

$$\begin{aligned}
& \cdot \mu^{IM}(x_6) \cdot \mu^M(x_7) \cdot \mu^M(x_8) \cdot \mu^{IM}(x_9) \cdot \mu^{IM}(x_{10}) \cdot \mu^{IM}(x_{11}) \vee \\
& \quad \vee \mu^H(x_1) \cdot \mu^M(x_2) \cdot \mu^H(x_3) \cdot \mu^{IM}(x_4) \cdot \mu^{BA}(x_5) \cdot \\
& \cdot \mu^{IM}(x_6) \cdot \mu^{IM}(x_7) \cdot \mu^{IM}(x_8) \cdot \mu^{IM}(x_9) \cdot \mu^{IM}(x_{10}) \cdot \mu^{IM}(x_{11}) \vee \\
& \quad \vee \mu^H(x_1) \cdot \mu^M(x_2) \cdot \mu^H(x_3) \cdot \mu^{IM}(x_4) \cdot \mu^{BA}(x_5) \cdot \\
& \cdot \mu^{IM}(x_6) \cdot \mu^{IM}(x_7) \cdot \mu^H(x_8) \cdot \mu^{IM}(x_9) \cdot \mu^{IM}(x_{10}) \cdot \mu^{IM}(x_{11})
\end{aligned} \tag{5}$$

for d4

$$\begin{aligned}
\mu^{d_4}(x_1 \div x_{11}) &= \mu^H(x_1) \cdot \mu^L(x_2) \cdot \mu^{IM}(x_3) \cdot \mu^M(x_4) \cdot \mu^L(x_5) \cdot \\
& \cdot \mu^{BA}(x_6) \cdot \mu^L(x_7) \cdot \mu^L(x_8) \cdot \mu^L(x_9) \cdot \mu^L(x_{10}) \cdot \mu^L(x_{11}) \vee \\
& \quad \vee \mu^H(x_1) \cdot \mu^L(x_2) \cdot \mu^H(x_3) \cdot \mu^{IM}(x_4) \cdot \mu^{BA}(x_5) \cdot \\
& \cdot \mu^M(x_6) \cdot \mu^{BA}(x_7) \cdot \mu^{BA}(x_8) \cdot \mu^{BA}(x_9) \cdot \mu^{BA}(x_{10}) \cdot \mu^{BA}(x_{11}) \vee \\
& \quad \vee \mu^H(x_1) \cdot \mu^L(x_2) \cdot \mu^H(x_3) \cdot \mu^H(x_4) \cdot \mu^M(x_5) \cdot \\
& \cdot \mu^{IM}(x_6) \cdot \mu^M(x_7) \cdot \mu^M(x_8) \cdot \mu^M(x_9) \cdot \mu^M(x_{10}) \cdot \mu^M(x_{11}) \vee \\
& \quad \vee \mu^H(x_1) \cdot \mu^L(x_2) \cdot \mu^H(x_3) \cdot \mu^H(x_4) \cdot \mu^{IM}(x_5) \cdot \\
& \cdot \mu^H(x_6) \cdot \mu^{IM}(x_7) \cdot \mu^{IM}(x_8) \cdot \mu^M(x_9) \cdot \mu^{IM}(x_{10}) \cdot \mu^{IM}(x_{11}) \vee \\
& \quad \vee \mu^H(x_1) \cdot \mu^L(x_2) \cdot \mu^H(x_3) \cdot \mu^H(x_4) \cdot \mu^H(x_5) \cdot \\
& \cdot \mu^H(x_6) \cdot \mu^H(x_7) \cdot \mu^H(x_8) \cdot \mu^M(x_9) \cdot \mu^{IM}(x_{10}) \cdot \mu^{IM}(x_{11})
\end{aligned} \tag{6}$$

Results. Substituting our values:

$$\begin{aligned}
\mu^{d_1} &= 0,05 \cdot 0,08 \cdot 0,18 \cdot 0,28 \cdot 0,55 \cdot 0,3 \cdot 0,42 \cdot 0,51 \cdot 0,83 \cdot 0,44 \cdot 0,4 \vee \\
& \vee 0,14 \cdot 0,08 \cdot 0,25 \cdot 0,45 \cdot 0,55 \cdot 0,48 \cdot 0,7 \cdot 0,32 \cdot 0,46 \cdot 0,31 \cdot 0,2 \vee \\
& \vee 0,14 \cdot 0,08 \cdot 0,32 \cdot 0,7 \cdot 0,55 \cdot 0,8 \cdot 0,7 \cdot 0,32 \cdot 0,29 \cdot 0,19 \cdot 0,13 \vee \\
& \vee 0,14 \cdot 0,08 \cdot 0,5 \cdot 0,7 \cdot 0,55 \cdot 0,8 \cdot 0,7 \cdot 0,32 \cdot 0,19 \cdot 0,19 \cdot 0,13 = 0,05 \wedge 0,08 \wedge 0,08 \wedge 0,08 = 0,08
\end{aligned}$$

$$\begin{aligned}
\mu^{d_2} &= 0,05 \cdot 0,21 \cdot 0,32 \cdot 0,45 \cdot 0,55 \cdot 0,48 \cdot 0,42 \cdot 0,8 \cdot 0,83 \cdot 0,82 \cdot 0,7 \vee \\
& \vee 0,14 \cdot 0,21 \cdot 0,5 \cdot 0,7 \cdot 0,55 \cdot 0,8 \cdot 0,7 \cdot 0,51 \cdot 0,46 \cdot 0,44 \cdot 0,4 \vee \\
& \vee 0,14 \cdot 0,21 \cdot 0,5 \cdot 0,63 \cdot 0,55 \cdot 0,68 \cdot 0,8 \cdot 0,32 \cdot 0,29 \cdot 0,31 \cdot 0,2 \vee \\
& \vee 0,14 \cdot 0,21 \cdot 0,5 \cdot 0,63 \cdot 0,55 \cdot 0,68 \cdot 0,8 \cdot 0,32 \cdot 0,29 \cdot 0,19 \cdot 0,2 = 0,05 \wedge 0,14 \wedge 0,14 \wedge 0,14 = 0,14
\end{aligned}$$

$$\begin{aligned}
\mu^{d_3} &= 0,85 \cdot 0,9 \cdot 0,32 \cdot 0,45 \cdot 0,55 \cdot 0,48 \cdot 0,42 \cdot 0,38 \cdot 0,83 \cdot 0,82 \cdot 0,7 \vee \\
& \vee 0,85 \cdot 0,28 \cdot 0,5 \cdot 0,7 \cdot 0,8 \cdot 0,8 \cdot 0,7 \cdot 0,64 \cdot 0,46 \cdot 0,44 \cdot 0,4 \vee \\
& \vee 0,85 \cdot 0,21 \cdot 0,65 \cdot 0,63 \cdot 0,8 \cdot 0,68 \cdot 0,8 \cdot 0,8 \cdot 0,29 \cdot 0,31 \cdot 0,2 \vee \\
& \vee 0,85 \cdot 0,21 \cdot 0,65 \cdot 0,63 \cdot 0,8 \cdot 0,68 \cdot 0,5 \cdot 0,51 \cdot 0,29 \cdot 0,31 \cdot 0,2 \vee \\
& \vee 0,85 \cdot 0,21 \cdot 0,65 \cdot 0,63 \cdot 0,8 \cdot 0,68 \cdot 0,5 \cdot 0,32 \cdot 0,29 \cdot 0,31 \cdot 0,2 = \\
& = 0,32 \wedge 0,28 \wedge 0,2 \wedge 0,2 \wedge 0,2 = 0,32
\end{aligned}$$

$$\begin{aligned} \mu^{d_4} &= 0,85 \cdot 0,9 \cdot 0,5 \cdot 0,7 \cdot 0,55 \cdot 0,48 \cdot 0,42 \cdot 0,38 \cdot 0,54 \cdot 0,54 \cdot 0,62 \vee \\ &\vee 0,85 \cdot 0,9 \cdot 0,65 \cdot 0,63 \cdot 0,8 \cdot 0,8 \cdot 0,7 \cdot 0,64 \cdot 0,83 \cdot 0,82 \cdot 0,7 \vee \\ &\vee 0,85 \cdot 0,9 \cdot 0,65 \cdot 0,44 \cdot 0,45 \cdot 0,68 \cdot 0,8 \cdot 0,8 \cdot 0,46 \cdot 0,44 \cdot 0,4 \vee \\ &\vee 0,85 \cdot 0,9 \cdot 0,65 \cdot 0,44 \cdot 0,2 \cdot 0,43 \cdot 0,5 \cdot 0,51 \cdot 0,46 \cdot 0,31 \cdot 0,2 \vee \\ &\vee 0,85 \cdot 0,9 \cdot 0,65 \cdot 0,44 \cdot 0,15 \cdot 0,43 \cdot 0,32 \cdot 0,32 \cdot 0,46 \cdot 0,31 \cdot 0,2 = \\ &= 0,38 \wedge 0,63 \wedge 0,4 \wedge 0,2 \wedge 0,15 = 0,63 \end{aligned}$$

From these data we can conclude that in patients with glaucoma 3 degrees. That coincides with the diagnosis, which established doctor.

In the course of biomedical research problem arises tuning neuron fuzzy network (Fig. 2). To configure the network using recurrence relations proposed by prof. Rothstein [1-2]. The essence of the model setting is in the selection of the parameters of membership functions ($b_i^{jp}(t)$, $c_i^{jp}(t)$) and weights of fuzzy rules ($w_{jp}(t)$) that provide the minimum difference between the models and diagnostic results:

$$\sum_{i=1}^M (F_y(\widehat{x}_1^l, \widehat{x}_2^l, \dots, \widehat{x}_{12}^l, W_i) - \widehat{y}_l)^2 = \min_{W_i} \quad (7)$$

where $\langle \widehat{X}_l, \widehat{y}_l \rangle, l = \overline{1, M}$ - the experimental investigations; b - coordinate maximum; c - parameter compression and stretching.

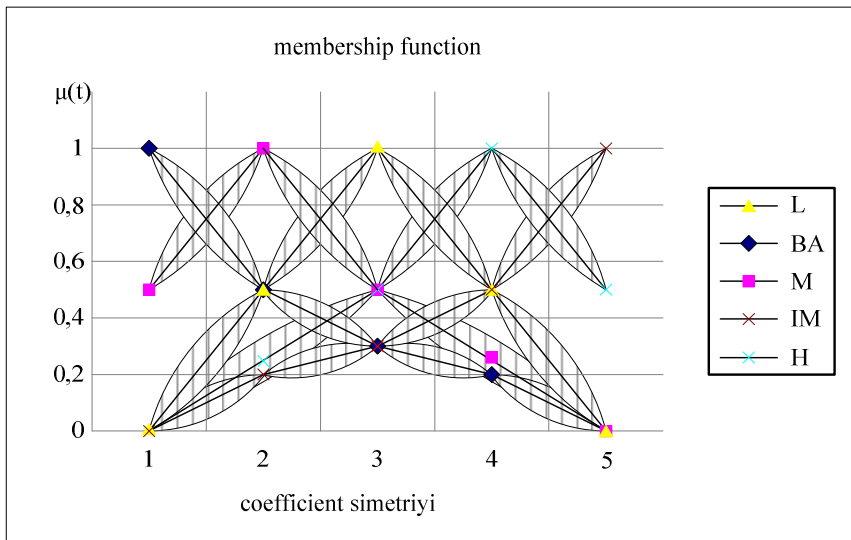


Figure 2 – The functions of fuzzy terms after the procedure set

In our case $\widehat{X}_l = (\widehat{x}_1^l, \widehat{x}_2^l, \dots, \widehat{x}_{12}^l)$ - the results of previous biomedical research, \widehat{y}_l - the result of the l-th trial.

To solve the nonlinear optimization problem (fig. 2) is applied algorithm "coarse" and "fine" setting [1-3] in the off-line and on-line.

To configure neuron fuzzy network used recurrence relations based on correlations [1-3]:

$$w_{jp}(t+1) = w_{jp}(t) - \eta \frac{\partial E_t}{\partial w_{jp}(t)}; \quad c_i^{jp}(t+1) = c_i^{jp}(t) - \eta \frac{\partial E_t}{\partial c_i^{jp}(t)};$$

$$b_i^{jp}(t+1) = b_i^{jp}(t) - \eta \frac{\partial E_t}{\partial b_i^{jp}(t)}, \quad (8)$$

where $w_{jp}(t)$, $c_i^{jp}(t)$, $b_i^{jp}(t)$ - the weight of rules and parameters of membership functions for the t -th step of learning; η - learning option.

This ratio minimizes the criterion $E_t = \frac{1}{2}(\bar{y}_t - y_t)^2$ that is used to configure the network.

Summary

Mathematical models are got for the estimation of biomedical images with the using of methods of unclear plurals on the basis of expert bases of knowledges. It allowed to conduct complex high-quality diagnosticating and promote authenticity and operationability of diagnosticating. The method of analysis of biomedical images is developed on the basis of vehicle of unclear plurals, that allowed completer of diagnosing glaucoma. Architecture of information technology is developed for support of making decision at the analysis of biomedical information, that allowed to promote authenticity of processing of biomedical images by the account of structural features of bioobjects.

Literature

1. Rotshtein A. Design and Tuning of Fussy IF – THEN Vuly for Medical Didicol Diagnosis. In Fussy and Neuro-Fussy Systems in Medicine (Eds: N. Teodovescu, A. Kandel, I. Lain.). – USA. CRC-Press, 1998, pp. 235–295.
2. Ротштейн А.П. Интеллектуальные технологии идентификации: нечеткие множества, генетические алгоритмы, нейронные сети / Ротштейн А.П. – Винница.: Универсум – Винница, 1999. – 320 с.
3. Павлов С.В. Оптико-електронні технології аналізу біомедичних зображень. Павлов С.В., Кожем'яко В.П., Бурденюк І.І. Вінниця: Універсум-ВНТУ, 2012. 166 с.
4. Вуйцік В. Експертні системи : навчальний посібник / Вуйцік В., Готра О., Григор'єв В. – Львів: Ліга-Прес, 2006. –290с.
5. Прэтт У. Цифровая обработка изображений / Прэтт У. – М.: Мир, т.1,2. 1982.
6. Искусственный интеллект. Системы общения и экспертные системы : Справочник : в 3 кн. /под ред. Э.В. Попов и др. – М.: Радио и связь, Кн. 1. – 1990. – 464с.
7. Поспелов Д.А. Искусственный интеллект – основа новой информационной технологии / Поспелов Д.А. – М.: Наука, 1988. – 269с.
8. Продеус А.Н. Экспертные системы в медицине / Продеус А.Н., Захарова Е.Н. – К.: ВЕК, 1998. – 320с.
9. Русин Б.П. Системы синтезу, обробки та розпізнавання складноструктурованих зображень / Русин Б.П. – Львів: Вертикаль, 1997. - 264 с.
10. Заде Л. Понятие лингвистической переменной и его применение к понятию приближенных решений / Заде Л. — М.: Мир, 1976. — 165 с.
11. Мельников В.Г. Медицинская кибернетика / В.Г. Мельников. – Киев: Вища школа, 1978. – 240 с.

Інформація про авторів

Павлов Сергій Володимирович – д.т.н., проф., завідувач кафедри загальної фізики та фотоніки, Вінницький національний технічний університет, (0432) 59-86-25, Хмельницьке шосе, 95, м. Вінниця, Україна, 21021.

Азаров Олексій Дмитрійович – д.т.н., проф., завідувач кафедри обчислювальної техніки, Вінницький національний технічний університет, azarov2@vntu.edu.ua, (0432)598-437, Хмельницьке шосе, 95, м. Вінниця, Україна, 21021.

Салдан Йосип Романович – д.м.н., проф., завідувач кафедри очних хвороб Вінницького національного медичного університету ім. М.І. Пирогова.

Рожман Анна Олександрівна – пошукач кафедри загальної фізики та фотоніки.

Бабюк Наталія Пєтєрївна – пошукач кафедри загальної фізики та фотоніки.

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

УДК 004.896

О.О. ДРУЖИНИНА, Р.Н. КВЕТНИЙ

Вінницький національний технічний університет, Вінниця

ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ФУНКЦІОНУВАННЯ ВЕБ-СЕРВЕРІВ З ВИКОРИСТАННЯМ ТЕХНОЛОГІЇ ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ НА ОСНОВІ НЕЙРОМЕРЕЖ

Анотація. Здійснено огляд та аналіз способів підвищення якості обслуговування веб-серверів. Запропоновано підхід інтелектуального кешування даних. Запропоновано підхід до підвищення ефективності функціонування веб-серверів з використанням технології прогнозування часових рядів на основі нейронних мереж радіально-базисного типу.

Ключові слова: якість обслуговування веб-серверів, прогнозування, нейронні мережі, GRNN, PNN, часові ряди, кешування. **Аннотация.** Приведен обзор и анализ существующих способов повышения качества обслуживания веб-серверов. Предложен подход интеллектуального кеширования данных. Предложен подход к повышению эффективности функционирования веб-серверов с использованием технологии прогнозирования временных рядов на основе нейронных сетей радиально-базисного типа.

Ключевые слова: качество обслуживания веб-серверов, прогнозирование, нейронные сети, GRNN, PNN, временные ряды, кеширование.

Abstract. Web-servers quality of service approaches overviewed. Intellectual data caching approach was proposed. Approach for web-servers efficiency improving, based on time series forecasting with radial-basis neural networks was proposed.

Key words: web-servers Quality of Service, forecasting, neural networks, GRNN, PNN, time series, caching.

Вступ

На ефективність функціонування веб-додатків значним чином впливає якість обслуговування (Quality of Service, QoS) веб-серверів.

Веб-сервер – це апаратно-програмна платформа, підключена до локальної або глобальної мережі, на якій буде розміщено програмний модуль або групу модулів, які здійснюватимуть обробку клієнтських запитів, а також формуватимуть відповіді на ці запити у відповідності до логіки, що реалізована у цих модулях. Апаратну частину веб-сервера становить серверний комп'ютер, програмну – операційна система та набір службових додатків, а також набір модулів, що реалізують логіку обробки клієнтських запитів. Службові модулі – це модулі, що реалізують роботу з мережевими інтерфейсами і надають необхідне програмне оточення для роботи модулів обробки клієнтських запитів.

Серед найбільш розповсюджених службових додатків веб-серверів можна виділити наступні: Apache, Nginx, Microsoft IIS. Зазвичай, веб-сервери отримують запити через мережу по протоколу HTTP (або HTTPS у випадку використання захищеного каналу передачі) або через інші протоколи, надбудовані над ним, наприклад XML-RPC чи SOAP, тощо. Також веб-сервери можуть використовувати базу даних чи інші репозиторії для зберігання необхідних даних. Зазвичай, база даних розташована на окремому сервері, і обмін даними з веб-сервером здійснюється через локальну чи глобальну мережу за допомогою протоколів, специфічних для систем управління базами даних. Отже, веб-сервер є достатньо складним об'єктом, коректна робота та швидкодія якого залежить від багатьох чинників.

До головних метрик якості обслуговування веб-серверів відносяться: час відгуку та собівартість[1]. Зростання Інтернет обчислень призвело до зростання обчислювальних центрів, внаслідок чого корпоративні центри обробки даних, як правило, складаються з великої кількості серверів, які не використовуються в повній мірі. Це в свою чергу стало причиною збільшення вартості обслуговування відповідних сервісів, що обумовлено прямою залежністю між кількістю серверів та витратами на їх аренду, електроенергію, ліцензії на програмне забезпечення та адміністрування. Саме тому на сьогодні зростає кількість корпорацій, які використовують пули серверних ресурсів, що стало можливим завдяки розвитку віртуалізації серверів, мереж та пристроїв збереження інформації. Серверний пул являє собою групу серверів в кластері, які об'єднані в деяке логічне одиницю(пул). Ідея серверних пулів полягає у відсутності прив'язки ресурсів в кластері до конкретних фізичних серверів. Негативною особливістю такої системи є складність її адміністрування, оскільки для ефективного використання серверних пулів необхідно здійснювати аналіз навантажень, прогнозування навантаження на окремий ресурс та ефективний перерозподіл ресурсів.

Задачу ефективного управління пулами ресурсів можна розділити на декілька простіших підзадач:

- 1) управління доступом до ресурсів з метою визначення чи є в пулі достатньо ресурсів для надання нового навантаження;
- 2) розміщення робочого навантаження на програмні додатки з метою отримання рекомендації щодо оптимального розміщення навантаження для зменшення кількості використовуваних серверів.
- 3) прогнозування потреб навантаження.

В даній роботі увага зосередиться на останній задачі прогнозування потреб навантаження. Аналіз розглянутих літературних джерел з даної проблематики показав, що наразі існує невідповідність існуючих інформаційних технологій прогнозування навантаження веб-серверів сучасним вимогам щодо спів-

відношення вартість/ефективність, що спричинено використанням недостатньо точних методів та моделей прогнозу, які базуються на спрощеному математичному апараті [2, 3]. Саме тому є доцільним розробити технологію прогнозування серверного навантаження, яка б базувалась на підходах та методах ідентифікації складних нелінійних систем та дослідити її ефективність у порівнянні з висвітленими у літературних джерелах.

За час відгуку веб-сервера відносно клієнтської сторони приймемо час, необхідний для відправки запиту на сервер та одержання і обробку відповіді клієнтом. Часом, що знадобиться на те, щоб попередньо підготувати до відправки клієнтський запит, а також часом, витраченим на обробку відповіді, можна знехтувати, оскільки дані часові характеристики залежать виключно від клієнтської сторони, і жодним чином не стосуються веб-сервера. Час відгуку веб-сервера може залежати від декількох факторів, кожен з яких може впливати істотним чином на величину часового інтервалу між запитом до сервера і відповіддю сервера. До таких факторів можна віднести:

- якість мережевого підключення, від якої залежить безпосередня швидкість передачі даних;
- швидкодія апаратної платформи, на якій розташований веб-сервер;
- характеристики операційної системи на якій розгорнуто веб-сервер;
- архітектура та реалізація службового додатку веб-сервера (швидкодія, ефективність, оптимальність та потреба в системних ресурсах, тощо);
- швидкодія модулів, за допомогою яких веб-сервер здійснює взаємодію із зовнішніми ресурсами, необхідними для його коректного функціонування;
- архітектура та реалізація системи управління базою даних, яка застосовується для зберігання даних, необхідних веб-серверу;
- архітектура та реалізація модулів обробки клієнтських запитів (рис 1).

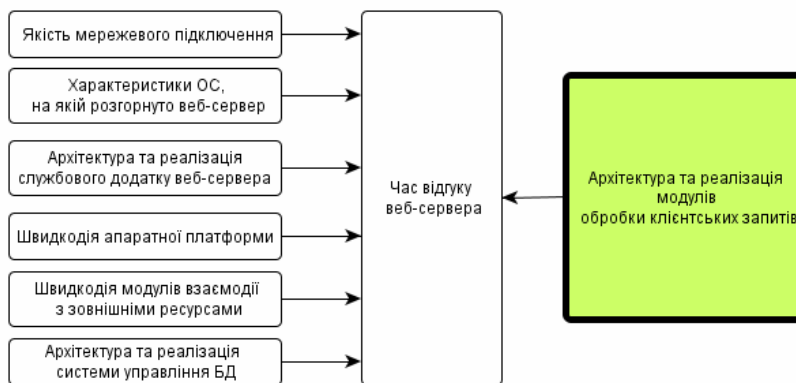


Рисунок 1– Фактори, які впливають на час відгуку веб-сервера

В даній роботі ми не будемо зосереджувати увагу на апаратних і мережевих факторах, а також факторах, пов'язаних із операційною системою, службовим додатком чи системою управління базою даних. Зосередимось на модулі обробки клієнтських запитів.

Модуль обробки клієнтських запитів містить реалізацію роботи сервісу, який вирішує ті чи задачі або надає клієнтам необхідні дані. Як показує практика, при правильному проектуванні і ефективній реалізації сервісного модуля, «вузьким місцем» в ньому стає комунікація з зовнішніми компонентами чи серверами, наприклад, базою даних. Тобто, основна частка часу, що витрачається на обробку клієнтського запиту, йде на встановлення зв'язку між веб-сервером і зовнішніми серверами чи компонентами, а також на передачу даних між ними.

Якщо в якості зовнішнього компонента розглядати базу даних, то значна частина часу витрачається на встановлення з'єднання з базою даних та обробку запитів – їх розбір, компіляцію, визначення і обрання оптимального плану виконання, очікування зняття блокувань при доступі до ізольованих даних, вибірку даних, тощо.

Для вирішення проблеми тривалого очікування з'єднання з базою даних застосовують підхід при якому формують «пули» з'єднань з базою даних, які використовуються повторно. При необхідності з'єднання з базою даних, компонент, який забезпечує це з'єднання, звертається до спеціального контейнера, в якому містяться наявні активні з'єднання з базою даних. Якщо цей контейнер порожній, компонент створює нове з'єднання, повертає його компоненту, який його потребує, і після того як це

з'єднання було використано, не закриває його, а переміщує його у вищевказаний контейнер. Таким чином одне з'єднання з базою даних може бути використано багаторазово без накладних витрат, пов'язаних із його створенням та закриттям.

Для вирішення проблеми тривалої обробки запитів на стороні бази даних використовують технологію кешування даних, які не змінюються на протязі певного інтервалу часу. Наприклад, якщо два запити до бази даних з однаковими умовами повертають однаковий результат, то доцільним є розміщення даного результату у кеші, і виключення звернення до бази даних, що значно прискорює роботу веб-сервера.

Розглянемо кешування більш докладно. Кеш являє собою окремий фізичний сервер, на якому виконується лише додаток, що забезпечує кешування даних. Кешування полягає у розміщенні об'єкта, що буде закешований в оперативній пам'яті кеш-сервера. Закешований об'єкт маркується відповідним ключовим значенням (яке обирає клієнт кеша) і з використанням якого в подальшому відбувається доступ до закешованого об'єкта. Оскільки робота і пошук в оперативній пам'яті є значно швидшими ніж дискові операції – кеш значно прискорює отримання даних, які не змінюються протягом певного часу. Розглянемо роботу кешування на прикладі програмного додатку *Memcached*, який насьогодні є найбільш розповсюдженим і еталонним рішенням для кешування даних [4].

Memcached є кросплатформним додатком, який запускається на цільовому сервері та взаємодіє з зовнішніми клієнтами за допомогою спеціального протоколу. Клієнти, які бажають працювати з додатком *Memcached* повинні містити відповідні клієнти, які можуть взаємодіяти з додатком *Memcached*, наприклад *xmemcached*. Клієнт передає *Memcached* команду *set* і таким чином здійснює розміщення необхідного об'єкта в кеші. Параметрами цього методу виступають ключ, за яким відбуватиметься пошук об'єкта в кеші, сам об'єкт, який необхідно закешувати, а також дата, коли об'єкт «застаріє» і його значення необхідно буде актуалізувати на основі даних з бази даних або інших джерел. Закешовані об'єкти в процесі роботи можуть займати весь обсяг оперативної пам'яті, яка була виділена додатком *memcached* при запуску. Коли об'єкт стає «застарілим», він продовжує зберігатись в пам'яті кешу до тих пір, доки не постане необхідність для «кешування» нових об'єктів, або не відбудеться звернення до даного об'єкта і він буде видалений.

Виходячи з вищенаведеного, ми бачимо, що існуюча концепція кешування має ряд недоліків, зокрема:

- захарачення кешу об'єктами, які ще не «застаріли», проте ніким не використовуються, і як наслідок неможливість закешувати нові об'єкти;
- використання всієї доступної оперативної пам'яті додатком *Memcached* в той час, коли для активних закешованих об'єктів необхідний менший обсяг оперативної пам'яті, неможливість використати цю пам'ять для інших цілей;
- заповнення контейнеру ключів закешованих об'єктів «зайвими» ключами об'єктів, до яких не буде звернення, і відповідно зі зростанням розміру контейнера, збільшенням часу пошуку необхідного ключа;
- чутливість до помилок клієнтів в налаштуванні тривалості кешування об'єктів – при обранні невмотивовано великого строку кешування, захарачення кешу непотрібними об'єктами.

Отже, як видно з вищенаведених недоліків, основним «вузьким місцем» кешу є розмір дозволеної до використання оперативної пам'яті. Дану проблему було вирішено через «приєднання» до кешу, розміщеного в оперативній пам'яті, простору на жорсткому диску. Дане було реалізовано у 2010 році у проєкті *Membase*, який в наш час перейменовано в *Couchbase*. Додаток *Couchbase* є повністю сумісним для клієнтів *Memcached*. *Couchbase* має наступний принцип роботи: при додаванні об'єкта до кешу, він розміщується в оперативній пам'яті, як і у випадку з *Memcached*, а його копія зберігається на диску. При переповненні кешу, розміщеного в оперативній пам'яті, елементи з нього видаляються, при цьому їх копії на жорсткому диску продовжують зберігатись, і в разі потреби клієнта кешу в даних, яких немає в оперативній пам'яті, вони завантажуються з жорсткого диску [5, 6, 7].

Однак, *Couchbase* має ті ж недоліки, які властиві *Memcached* – захарачення кешу об'єктами, які ще не «застаріли», проте ніким не використовуються, заповнення контейнеру ключів закешованих об'єктів «зайвими» ключами об'єктів, тощо.

Мета

Недослідженість питань, висвітлених у вступі визначили основну мету даної роботи, яка полягає у забезпеченні задовільного співвідношення якості обслуговування веб-серверів до собівартості в будь-який момент часу за рахунок оптимізації використання серверних ресурсів, мінімізуючи їх вартість шляхом прогнозування навантаження на сервер, а також у зменшенні часу відгуку веб-серверів за рахунок реалізації інтелектуального кешування даних.

Постановка та розв'язання задач

В даній роботі пропонується підхід, який, на нашу думку, зможе позбавити існуючі підходи до кешування об'єктів недоліків, пов'язаних з неефективним використанням оперативної пам'яті, які були наведені вище.

Основним критерієм ефективності кеша виступає швидкість реакції на запит закешованого об'єкту. Тобто, чим швидше ми одержимо об'єкт з кешу, тим вищою буде ефективність кешу. На нашу думку, підвищення ефективності кешу можна досягти за рахунок зменшення кількості закешованих об'єктів, які зберігатимуться в оперативній пам'яті до меж необхідності. Тобто, ми повинні визначити і залишити в кеші тільки ті об'єкти, в яких дійсно є потреба. Зменшення кількості закешованих об'єктів призведе до зменшення ключів, за якими буде здійснюватись пошук цих об'єктів, і відповідно призведе до підвищення швидкодії пошуку і зменшення часу, що витрачається на пошук ключа. Іншим позитивним ефектом, що досягається зменшенням кількості закешованих об'єктів, є більш ефективне використання оперативної пам'яті, через зменшення потреб в ній з боку кеша.

Ми будемо виходити з тих міркувань, що всі об'єкти в кеші є нерівномірно затребуваними на протязі заданого періоду часу, наприклад, доби. Тобто, одні об'єкти затребувані зранку, інші всередині дня, ті що залишились – ввечері. Крім того, можливо здійснити такий розподіл по годинах. Тому, доцільно буде зберігати закешовані об'єкти в оперативній пам'яті лише на протязі часового інтервалу, коли вони дійсно будуть потрібними, в інший же час, закешовані об'єкти можливо зберігати разом з їхніми ключами на диску.

Розглянемо принцип роботи такого механізму на прикладі одного закешованого об'єкту. При додаванні об'єкту до кеша, він розташовується в оперативній пам'яті, його ключ додається до набору ключів закешованих об'єктів, копія ключа і об'єкту переміщується на жорсткий диск. Далі, спеціальний програмний модуль починає відслідковувати актуальність закешованого об'єкту, яка полягає у кількості звернень до об'єкту в одиницю часу, наприклад, годину. Дані щодо актуальності закешованого об'єкту збираються на протязі керованого контрольного періоду часу. Інший програмний модуль прогнозування аналізує дані актуальності закешованого об'єкту і формує графік актуальності об'єкта в майбутньому. Графік являє собою таблицю, розподілену на часові інтервали і ознаку актуальності об'єкта у вказаний період – буде об'єкт актуальним чи ні. Приклад такої таблиці наведено в табл. 1.

Таблиця 1 – Актуальність «закешованого» об'єкту по часовим інтервалам

Межі часового інтервалу	Ознака актуальності «закешованого» об'єкту
12.00 – 13.00	+
13.00 – 14.00	-
...	...

Перед початком наступного часового інтервалу, спеціальний програмний модуль аналізує графік актуальності закешованих об'єктів і утворює нову множину ключів закешованих об'єктів, а також множину актуальних закешованих об'єктів, які розміщує в оперативній пам'яті на початку наступного часового інтервалу на час цього інтервалу. Попередні значення видаляються з оперативної пам'яті. Якщо для закешованого об'єкта закінчився термін життя в кеші – він видаляється автоматично і з оперативної пам'яті і з жорсткого диску. Кількість, тривалість часових інтервалів можуть бути налаштовані окремо з метою найбільш ефективного пристосування до умов роботи кеша. При збиранні даних, щодо «актуальності» закешованого об'єкта, якщо об'єкт є неактуальним (незатребуваним), навіть, якщо він не застарів, він може бути видалений як з оперативної пам'яті, так і з жорсткого диску до настання того моменту, коли він дійсно застаріє.

Реалізацію даного модуля пропонується здійснити на основі технології ідентифікації часових рядів (ЧР) з використанням нейронних мереж радіально-базисного типу. Метою прогнозування часових рядів є оцінка значень ЧР в майбутньому, що базується на даних, які були отримані в минулому та поточний момент часу. Схематично даний підхід до кешування даних зображений на рисунку 2.

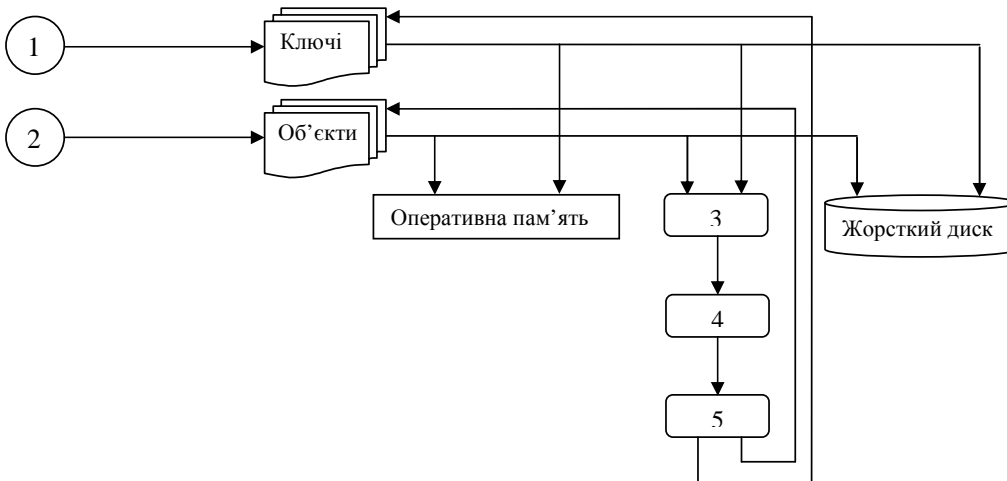
Математично ЧР являє собою ряд динаміки, впорядкований за часом, або сукупність спостережень деякої величини в різні моменти часу:

$$x(t) = \{x(t_1), x(t_2), x(t_3), \dots, x(t_N)\}, \quad (1)$$

де t_1, t_2, \dots, t_N – моменти спостережень, N – довжина часового ряду.

Вибір нейронних мереж радіально-базисного типу для розв'язання даної задачі пояснюється тим, що вони дозволять з легкістю досягти компромісу високої точності прогнозу та малих витрат часу на

його здійснення, що надає можливість застосовувати їх в режимі реального часу для надання коротко-строккових прогнозів.



- 1 – ключ «закешованого» об'єкта,
- 2 – «закешований» об'єкт,
- 3 – модуль, що відслідковує і збирає дані щодо актуальності «закешованого» об'єкта,
- 4 – модуль, що аналізує дані актуальності «закешованого» об'єкта і здійснює прогнозування,
- 5 – модуль, що на основі даних, отриманих від модуля 4, управляє часом життя «закешованого» об'єкта в оперативній пам'яті.

Рисунок 2 – Схема інтелектуального кешування даних

Оскільки результатом модуля прогнозування для окремого є значення бінарного типу, то поставлена задача зводиться до задачі класифікації і пропонується використовувати ймовірнісну нейронну мережу (PNN - Probabilistic Neural Network) [8]. Архітектура ймовірнісної нейронної мережі зображена на рис. 3. Особливістю мереж даного типу є те, що їх структура формується в ході навчання. Розмірність N векторів навчальної вибірки визначає число нейронів і структуру вхідного шару ймовірнісної нейронної мережі. Загальний розмір навчальної вибірки відповідає загальній кількості нейронів шару зразків. Пред'явлення мережі кожного з зразків навчальної вибірки супроводжується вказівкою від вчителя номера j -го класу, якому належить вхідний зразок. Після пред'явлення всіх векторів навчальної вибірки, формується структура мережі, і стають визначеними параметри мережі у вигляді матриці. На цьому процес навчання ймовірнісної нейронної мережі завершується і мережа готова до класифікації невідомих зразків.

В класичній PNN кожний нейрон шару сумування сумує рівні активності всіх нейронів шару зразків свого класу і видає на своєму виході загальний рівень активності даного j -го класу. Далі визначається який нейрон шару сумування має максимальний вихідний сигнал, тим самим (за номером j -го нейрона) визначається номер класу j , до якого з більшою ймовірністю належить пред'явлений образ X .

Функція активності j -го нейрона сумування визначає значення густини розподілу ймовірностей для всього j -го класу:

$$\hat{f}_{n_j}(x) = \frac{1}{n_j} \sum_{i=1}^{n_j} K_{n_j}(x, X_i^{(j)}), \quad (2)$$

де n_j – кількість зразків в класі j ; $j=1, \dots, M$; K – ядерна функція.

Тоді дискримінантна функція оцінки буде обчислюватись за формулою:

$$\hat{d}_{n_j}(x) = \frac{1}{n} \sum_{i=1}^{n_j} K_{n_j}(x, X_i^{(j)}). \quad (3)$$

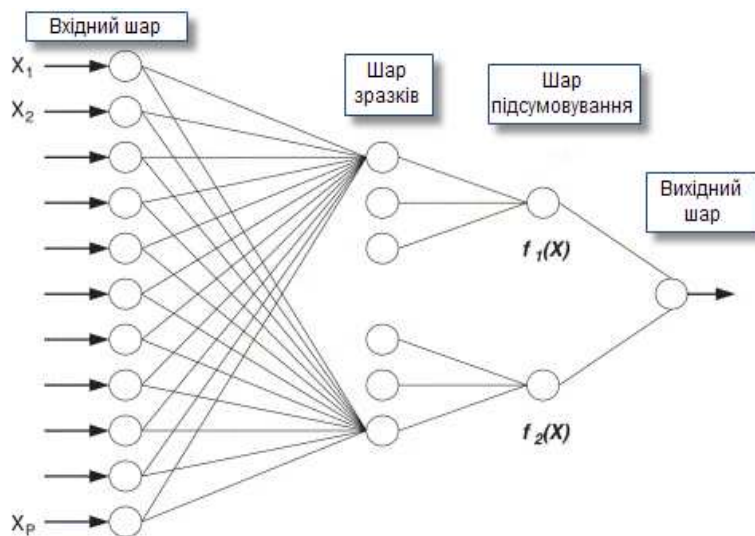


Рисунок 3 – Архітектура PNN

Відповідна процедура класифікації:

$$\hat{\varphi}_n(x) = m, \text{ якщо } \hat{\sum}_{i=1}^{n_m} K_{n_m}(x, X_i^{(m)}) \geq \sum_{i=1}^{n_j} K_{n_j}(x, X_i^{(j)}) \quad (4)$$

інакше $i \neq m, i = 1, \dots, M.$

Модуль прогнозування актуальності об'єкта буде являти собою багатоетапну технологію прогнозування на основі PNN, яка включатиме в себе реалізацію накопичення та попередньої обробки даних, формування навчальних зразків, навчання НМ, пошук розв'язку для визначення актуальності заданого об'єкта.

Отже, запропонований підхід дасть змогу оптимізувати кеш за рахунок зберігання в ньому лише дійсно необхідних об'єктів і їх ключів, що, в свою чергу, підвищить його швидкодію за рахунок зменшення часу на пошук та одержання об'єкта.

Для моделювання ефективності функціонування пулу серверів є необхідним ввести показники, за допомогою яких можна було б здійснити її оцінку. Необхідна потужність – рівень навантаження на пул ресурсів, який відображає максимальну потребу в обчислювальній потужності. Необхідна потужність має завжди задовольняти наступну умову: «сума максимальних навантажень на кожний ресурс не має перевищувати потужність серверу».

Навантаження на веб-сервери може бути визначене величиною параметрів, які відповідають обсягам запитів та з'єднань до них. Практичний досвід та результати вказують на складний вигляд процесу зміни вказаних параметрів. Було доведено, що для веб-серверів типовими є як стаціонарні, так і нестаціонарні режими експлуатації з багатоперіодичним характером зміни величин функціональних параметрів. Нестационарність режимів експлуатації проявляється у нестаціонарності моментів виникнення/зникнення періодичних складових [2, 9].

Для розв'язання задачі короткострокового прогнозування навантаження на сервер пропонується використати узагальнено-регресійну нейронну мережу GRNN (General Regression Neural Network). Архітектура даної мережі подібна до PNN. Прогнозування навантаження на сервер з використанням GRNN буде являти собою багатоетапну технологію, яка включатиме в себе: накопичення даних; попередню

обробку даних; формування навчальних зразків; навчання НМ; пошук розв'язку для визначення рівня навантаження на сервер в майбутньому.

Висновки

В роботі здійснено аналіз підходів до підвищення якості обслуговування веб-серверів. Виділені недоліки існуючих методів підвищення QoS, які базуються на технології кешування даних та прогнозуванні навантаження на сервер.

Запропоновано підхід інтелектуального кешування даних на основі прогнозування актуальності закешованих об'єктів з використанням нейронної мережі радіально-базисного типу, який на відміну від існуючих, реалізує оптимізацію кешу за рахунок зберігання в ньому лише дійсно необхідних об'єктів і їх ключів, що, в свою чергу, підвищує його швидкодію.

Для прогнозування навантаження на веб-сервер запропоновано використання інформаційної технології прогнозування часових рядів на основі узагальнено-регресійної нейронної мережі. Є доцільним продовження даних досліджень і здійснення перевірки ефективності запропонованих підходів на реальних даних.

Список літератури

1. Sravanthi Kalepu, Shonali Krishnaswamy, Seng Wai Loke, Verity: A QoS Metric for Selecting Web Services and Providers [Електронний ресурс] // Proceedings of the Fourth International Conference on Web Information Systems Engineering Workshops (WISEW'03). – 2004. – Режим доступу до файлу: <http://www.acs.org.au/vic/socsig/IEEE-VeritySOA-Krishnaswamy.pdf>
2. Терейковська Л.О. Архітектура марківської моделі зміни навантаження Web-сервера / Л.О. Терейковська // Вісник ДУІКТ. – 2012. – Т.10, №1. – С. 95-100.
3. Tom Vercauteren, Pradeep Aggarwal, Xiaodong Wang Hierarchical Forecasting of Web Server Workload Using Sequential Monte Carlo Training // IEEE transactions on signal processing. – 2007. – VOL. 55. – № 4. – PP. 635-644.
4. Memcached [Електронний ресурс]. – Режим доступу до файлу: <http://en.wikipedia.org/wiki/Memcached>
5. Membase Server is Now Couchbase Server [Електронний ресурс]. – Режим доступу до файлу: <http://www.couchbase.com/membase>.
6. Disk Storage [Електронний ресурс]. – Режим доступу до файлу: <http://www.couchbase.com/docs/couchbase-manual-1.8/couchbase-introduction-architecture-diskstorage.html>.
7. Couchbase features [Електронний ресурс]. – Режим доступу до файлу: <http://www.couchbase.com/couchbase-server/features>.
8. Кветний Р.Н. Імовірнісні нейронні мережі в задачах ідентифікації часових рядів / В. В. Кабачій, О. О. Чумаченко [Електронний ресурс]. – Режим доступу до файлу.: http://www.nbu.gov.ua/e-journals/vntu/2010_3/2010-3.files/uk/10rnktsi_ua.pdf.
9. Agustín C. Caminero, Salvador Ros, Roberto Hernández, Antonio Robles-Gómez, Rafael Pastor Cloud-based e-Learning Infrastructures with Load Forecasting Mechanism Based on Exponential Smoothing: A Use Case [Електронний ресурс]. – Режим доступу до журн.: <http://fie-conference.org/fie2011/papers/1085.pdf>.

Відомості про авторів

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

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

УДК 355.01

ДУДАТЬСВ А.В.

Вінницький національний технічний університет, Вінниця

АКСІОМАТИКА ТЕОРІЇ КОМПЛЕКСНОЇ БЕЗПЕКИ СОЦІОТЕХНІЧНИХ СИСТЕМ

Анотація. У статті запропоновані базові аксиоми, які відносяться до одного з ключових показників стану соціотехнічної системи – стану її захищеності в умовах ведення інформаційної війни. Наведені аксиоми в подальшому дозволять розвинути загальну теорію комплексної безпеки соціотехнічних систем.

Ключові слова: комплексна безпека, соціотехнічна система, інформаційна війна.

Аннотация. В статье предложены базовые аксиомы, которые относятся к одному из ключевых показателей состояния социотехнической системы – состояния её защищённости. Предложенные аксиомы в дальнейшем позволят развить общую теорию комплексной безопасности социотехнических систем.

Ключевые слова: комплексная безопасность, социотехническая система, информационная война.

Abstract: This article contains basic axioms which are one of the key indicators of the socio-technical system - the state of its security. Proposed in the future will axioms express a general theory of complex socio-technical security system.

Keywords: integrated security, sociotechnical system, information warfare.

Вступ

Захищеність сучасних соціотехнічних систем (СТС) залежить від того наскільки ефективно система забезпечення інформаційної безпеки реагує на існуючі загрози. Важливе значення при цьому набуває оцінка стану комплексної безпеки, з урахуванням того, що життєдіяльність СТС, як об'єкта захисту відбувається у внутрішньому і зовнішньому інформаційних середовищах.

Крім того комплексна безпека сучасних СТС складається з декількох пов'язаних між собою складових. З урахуванням того, що порушення тієї чи іншої складової комплексної безпеки СТС може призвести до значних втрат, тобто наслідки виникнення небажаних подій у системі можуть бути значними і навіть критичними щодо подальшого існування самої системи і різних інфраструктур, що забезпечують життєдіяльність СТС, то такі СТС можна віднести до так званих критичних систем.

З урахуванням можливих критичних наслідків, а також того, що СТС, яка включає в себе інформаційні системи, є важливим технологічним ланцюжком щодо отримання, оброблення, зберігання і подальшого передавання інформації, то важливим моментом є забезпечення необхідного рівня захисту її інформаційних ресурсів.

Другим важливим аспектом, який суттєво впливає на функціонування СТС, є те, що вона знаходиться під інформаційним впливом конкуруючих з нею інших об'єктів, які можуть активно використовувати сучасні технології інформаційної війни з метою заволодіння перевагою на тому чи іншому сегменті ринку. Наразі популярними активними діями є проведення конкурентної розвідки, проведення промислової розвідки, проведення «чорного» PR, тощо. Тому наведені вище факти дозволяють навести актуальне твердження сучасності: необхідно захищати власні інформаційні ресурси і захищатись від інформаційного впливу конкурентів.

Актуальність

Ідея розвитку загальної теорії безпеки СТС полягає у побудові системної моделі, яка зв'яже суб'єкти, об'єкти і фактори, що впливають на життєдіяльність системи. Системний підхід дозволить забезпечити комплексний характер захисту і випередити дії конкурентів шляхом аналізу каналів витоку інформації, виконання спеціальних операцій щодо дій конкурентів, проведення відповідних організаційних заходів, тощо.

Життєдіяльність СТС супроводжується взаємним впливом складових: людина(її знання, вміння, психологічний стан) - технологічне середовище. Крім того, цей процес відбувається у конкурентному середовищі, тобто під інформаційним впливом інших об'єктів діяльності. Застосування технологій керованого хаосу, ефективність яких фахівцями порівнюється зі зброєю масового знищення, в останні роки набувають все більшого розповсюдження. Головною метою застосування таких технологій є унеможливлення суб'єктивного розвитку об'єкта захисту, отримання лідерства на відповідному сегменті ринку шляхом дискредитації своїх конкурентів або навіть їх знищення. Тому питання оцінювання та забезпечення комплексної безпеки сучасних СТС є надзвичайно актуальним.

Метою даної роботи є забезпечення необхідного рівня комплексної безпеки сучасних СТС, які функціонують в умовах інформаційної війни шляхом подальшого розвитку загальної теорії безпеки.

Постановка задачі

Формулювання аксіоматики загальної теорії комплексної безпеки сучасних соціотехнічних систем.

Рішення задач

Будь-яка система має свою морфологію, поведінку, самоповедінку, що породжує функціональну діяльність, відповідно до цільових функцій. Опис систем можна виконувати у декількох напрямках: функціональному, морфологічному, інформаційному тощо [1].

Для формулювання аксіом наведемо базові системні визначення, які наведені у роботі [2].

Визначення. Системою називається сукупність універсальних складових одиниць – елементів, які перебувають у певних співвідношеннях і зв'язках між собою, завдяки чому вони й становлять певну цілісність, неподільність, унітарність.

Далі наведемо визначення, які запропоновані Шияном А.А. і з моєї точки зору дозволяють представити конкретну систему, як об'єкта захисту з урахуванням функціонування його у множині середовищ: інформаційному, технологічному, виробничому, навколишньому тощо.

Визначення. Функціональне середовище системи – це характерна для системи сукупність правил і параметрів (часто сформульованих у вигляді законів або алгоритмів), за якими здійснюється взаємодія (обмін, взаємовідносини) між елементами системи та функціонування (розвиток) системи в цілому.

Визначення. Елемент системи – це умовно неподільна частина системи, що самостійно функціонує. Підкреслимо, що виділення елементів (розбивка системи на елементи) – це операція, у певному сенсі, суб'єктивна. І хоча вона найчастіше повністю визначає успіх або невдачу всього дослідження, вона надзвичайно важко піддається регламентації. Як правило, таке розчленовування системи здійснюється відповідно до апріорних уявлень дослідника. І, звичайно, виділення елементів істотно залежить від постановки задачі та мети, яка стоїть перед дослідником.

Визначення. Структура системи – це сукупність «ключових» елементів, які перебувають між собою в «сильних» зв'язках, що забезпечують такий обмін інформацією між елементами системи, який є визначальним для функціонування системи в цілому та способів її взаємодії із зовнішнім середовищем. Такі «структурозадаючі» елементи є свого роду «унікальними», виділеними. Проте вони є виділеними не за своєю індивідуальною специфікою, але за їх місцем розташування та їх роллю у функціонуванні системи.

Визначення. Границя системи – це сукупність пов'язаних між собою елементів, які – взяті у своїй сукупності – дозволяють здійснювати поділ на «внутрішнє» (наприклад, функціональне середовище системи) і «зовнішнє» середовища для розглянутої системи. Через такі «прикордонні» елементи, а, точніше, «місця», які вони займають, і відбувається весь обмін інформацією між системою та її оточенням.

Важливим практичним наслідком наведених визначень є можливість ідентифікації системи з обов'язковим описом таких даних: а) універсальних складових одиниць – (функціональних) елементів системи; б) зв'язків, які існують між цими елементами; в) особливо виділити структуру системи (як сукупність «специфічних місць», потрапляючи в які елементи здобувають «особливу вагу і значення», а також систему зв'язків між такими «виділеними» місцями); г) сукупність «прикордонних» елементів (скоріше навіть тих «місць», тих положень елементів у системі, знаходження в яких і надає цим елементам здатність «відмежовувати» внутрішність системи від навколишнього середовища. І кожен із перерахованих вище 4-х пунктів потрібно описати як в статично так і в динаміці, які формально можна представити у вигляді ситуаційних моделей. Таким чином, приходимо остаточно до опису довільної системи у вигляді восьми введених вище класів даних.

Задачу оцінювання та забезпечення комплексної інформаційної безпеки СТС будемо розглядати як наскрізну, тобто процес забезпечення безпеки необхідно реалізовувати на всіх етапах її життєдіяльності.

Для подальшого формулювання аксіоматики наведемо узагальнені ситуаційні моделі, які формалізують ймовірну поведінку об'єктів – конкурентів взаємодії під час ведення спеціальних інформаційних операцій.

Розглянемо два часткових варіанти взаємодії об'єктів. На рис.1 представлена взаємодія об'єктів А і В, яка відбувається за сценарієм – об'єкт А – пасивний, а об'єкт В – активний. Тобто об'єкт А не виконує жодних дій (протидій) по відношенню до об'єкта В.

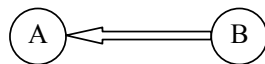


Рисунок 1 – Ситуаційна модель «активний-пасивний»

Реалізація такого сценарію може привести до змін структури і зв'язків об'єкта А, що у свою чергу дозволить об'єкту В отримати перемогу над об'єктом А.

На рис.2. представлена другий сценарій взаємодії об'єктів: об'єкт А активний і об'єкт В активний.

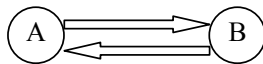


Рисунок 2 – Ситуаційна модель «активний-активний»

Реалізація другого сценарію може привести до змін структури і зв'язків, як об'єкта А так і об'єкта В. Переможець реалізації такого сценарію буде визначатися множиною чинників головними з яких є такі: наявність і кваліфікація працівників інформаційно-аналітичної служби, ефективність прогнозування розвитку подальших подій, якісне планування, спеціальне керування персоналом, координація і контроль прийнятих рішень.

Як відомо, є такі етапи життєдіяльності будь-якої системи, зокрема СТС:

- етап постановки задачі;
- етап проектування системи;
- етап створення системи;
- етап експлуатації системи;
- етап закінчення життєдіяльності системи.

Виходячи з вищевикладеного доцільно запропонувати такі аксіоми відносно кожного етапу життєдіяльності:

На етапі постановки задачі оцінювання та забезпечення безпеки формалізується такою аксіомою.

Аксіома 1. На етапі постановки задачі рівень безпеки визначається елементами системи, зв'язками між елементами, а також умовами експлуатації майбутньої системи і не може бути меншою допустимого рівня.

$$Z_{\tilde{n}\tilde{n}} \geq Z_{\tilde{a}\tilde{i}}$$

де $Z_{\text{сист}}$ - рівень безпеки системи, що проектується, $Z_{\text{доп}}$ - допустимий рівень безпеки.

На етапі проектування системи поставлена задача описується такою аксіомою.

Аксіома 2. На етапі проектування рівень безпеки забезпечується тривіальними методами, засобами та заходами з урахуванням умов експлуатації, які забезпечують виконання:

$$Z_{\tilde{n}\tilde{n}\tilde{o}} \geq Z_{\tilde{a}\tilde{i}}$$

Для етапу створення системи пропонується така система:

Аксіома 3. На етапі створення рівень безпеки забезпечується реалізацією методів, засобів та заходів, передбачених на етапі проектування.

Етап експлуатації системи характеризується декількома аксіомами:

Аксіома 4. На етапі експлуатації рівень комплексної безпеки СТС забезпечується в 2 етапи:

1. Оцінювання рівня комплексної безпеки;
2. Забезпечення необхідного рівня комплексної безпеки.

Аксіома 5. На етапі експлуатації, оцінювання та забезпечення необхідного рівня комплексної безпеки відбувається у конкурентному середовищі, яке включає інші антагоністичні системи.

Аксіома 6. На етапі експлуатації системи під інформаційним впливом конкуруючої системи може відбуватися зміна її структури або зв'язків між елементами, що може привести до зменшення необхідного рівня комплексної безпеки.

$$S = (E, ST, B, Q)$$

$$S1 = F(S) = (E1, ST1, B1, Q1),$$

де S - початкова система, яка має необхідний рівень безпеки, E - елементи цієї системи, ST - структура цієї системи, B - поведінка системи, Q – середовище, у якому система експлуатується, F - оператор перетворення. Відповідно для системи $S1$ маємо нові параметри $E1, ST1, B1, Q1$, які змінені внаслідок інформаційного впливу конкуруючої системи, і які вже не забезпечують необхідного рівня безпеки.

Аксіома 7. На етапі експлуатації системи може відбуватися її знищення антагоністичними системами або знищенням антагоністичних систем. У такому випадку можливі 2 варіанти:

$$1. \quad Z_{\tilde{n}\tilde{n}\tilde{o}} < Z_{\tilde{a}\tilde{i}}$$

якщо у результаті експлуатації відбулося знищення системи.

$$2. \quad Z_{\text{не}i} \geq Z_{\text{а}i}$$

якщо у результаті експлуатації відбулося знищення антагоністичної системи.

Етап закінчення життєдіяльності системи описується такою загальносистемною аксіомою.

Аксіома 8. Кожна система має закінчення життєвого циклу.

Сформульовані аксіоми в подальшому дозволять розвинути загальну теорію безпеки складних систем, зокрема систем, які є критичними і функціонують в умовах інформаційної війни і досягти головної мети - забезпечення необхідного рівня безпеки.

Висновки

Рішення задачі оцінювання та забезпечення комплексної безпеки сучасних соціотехнічних систем, яка у багатьох випадках функціонує в умовах проведення спеціальних інформаційних операцій, які можуть бути або типовими або навпаки невідомими у великій мірі є унікальним. Забезпечення комплексної безпеки СТС – це забезпечення їх здатності до динамічного ефективного розвитку в умовах постійної зміни, як внутрішнього так і зовнішнього середовищ.

Запропоновані аксіоми дозволяють продовжити розвиток загальної теорії комплексної безпеки СТС і досягти поставленої мети - забезпечити необхідний рівень комплексної безпеки.

Список літератури

1. В.С.Харченко. Безопасность критических инфраструктур: математические и инженерные методы анализа и обеспечения / В.С.Харченко.– Харьков: Изд-во Национальный аэрокосмический университет им. Н.Е.Жуковского. ("ХАИ"), 2011. – 641 с.

2. Шиян А.А. Теоретико-ігровий аналіз раціональної поведінки людини та прийняття рішень в управлінні соціально-економічними системами. Монографія./ Шиян А.А. – Вінниця: УНІВЕРСУМ-Вінниця, 2009. – 404 с.

Відомості про авторів

Дудатсьв Андрій Веніамінович – доцент кафедри захисту інформації, Вінницький національний технічний університет, (0432)598243, м. Вінниця, вул. Хмельницьке шосе 95.

УДК 681.3

А.Я. КУЛИК, С.Г. КРИВОГУБЧЕНКО, Ю.Ю. ІВАНОВ

Вінницький національний технічний університет, Вінниця

ПОРІВНЯЛЬНИЙ АНАЛІЗ СКЛАДНОСТІ РЕАЛІЗАЦІЇ МЕТОДІВ ДЕКОДУВАННЯ ТУРБО-КОДІВ

Анотація. Розглядаються питання турбо-кодування, перспективного для застосування у розподілених комп'ютерних системах різного функціонального призначення. Проаналізовано алгоритми SOVA, MAP, log-MAP та max-log-MAP. Проведено оцінку складності декодування турбо-коду. Виконано моделювання застосування методів декодування турбо-кодів для побудови апаратно-програмних засобів декодування сигналів у системах зв'язку з використанням обчислювальної потужності технічних засобів та дослідженням ймовірно-енергетичних характеристик.

Ключові слова: завадостійке кодування, турбо-код, узагальнена структура для роботи з інформацією, LLR, алгоритм Вітербі, SOVA, MAP, log-MAP, max-log-MAP, складність декодування турбо-кодів.

Аннотація. Рассматриваются вопросы турбо-кодирования, перспективного для применения в распределенных компьютерных системах различного функционального назначения. Проанализированы алгоритмы SOVA, MAP, log-MAP и max-log-MAP. Проведено оценку сложности декодирования турбо-кода. Проведено моделирование применения методов декодирования турбо-кодов для построения аппаратно-програмных средств декодирования сигналов в системах связи с использованием вычислительной мощности технических средств и исследованием вероятностно-энергетических характеристик.

Ключевые слова: помехоустойчивое кодирование, турбо-код, обобщенная структура для работы с информацией, LLR, алгоритм Витерби, SOVA, MAP, log-MAP, max-log-MAP, сложность декодирования турбо-кодов.

Abstract. Considered problems of turbo coding, promising for using in distributed computer systems of different functional purpose. Analysed SOVA, MAP, log-MAP and max-log-MAP algorithms. Completed the estimation complexity turbo-code decoding. Simulating the use of methods decoding turbo-codes for building hardware and software decoding signals in communication systems using the processing power of technical means and with research of probabilistic-energy characteristics.

Key words: protecting coding, turbo-code, a generalized structure for work with information, LLR, the Viterbi algorithm, SOVA, MAP, log-MAP, max-log-MAP, the complexity of decoding turbo-codes.

Вступ

Процес передавання даних в інформаційних системах чутливий до помилок, оскільки будь-яка помилка може суттєво порушити обчислення. Управління правильністю процесу передавання інформації забезпечується за допомогою завадостійкого кодування, яке забезпечує надійність і вірогідність переданої інформації в умовах завад, дозволяючи здійснити захист даних в пам'яті обчислювального пристрою та накопичувача інформації.

Найбільш помітним досягненням в теорії завадостійкого кодування за останні роки є турбо-коди. Вони описані в наукових працях С. Berrou, А. Glavieux, Р. Titimajshima та є досить потужним методом виправлення помилок. Володіють особливою здатністю забезпечувати характеристики завадостійкості передавання інформації, близькі до теоретично отриманих значень, використовуючи не дуже складну реалізацію пристроїв кодування [1].

Одними з найважливіших сфер використання турбо-кодування на практиці є цифрове телебачення, космічні та мобільні супутникові комунікації. Турбо-коди дозволяють збільшити швидкість передавання інформації, не вимагаючи збільшення потужності передавача, а також вони можуть бути використані для зменшення необхідної потужності при передаванні із заданою швидкістю. Отже, турбо-коди застосовують, щоб кодувати об'ємні інформаційні повідомлення для високошвидкісного передавання даних з великою завадостійкістю [2].

Використання турбо-кодування дозволяє підвищити завадостійкість приймання сигналів на 1-3 дБ і більше в залежності від типу сигналів, швидкості, виду кодування, якості каналу зв'язку. Порогові значення у турбо-кодів лежать на 2-3 дБ нижче, ніж у системах, що використовують декодування за Вітербі або Рідом-Соломоном.

Актуальність

Реалізація розроблених обчислювальних процедур для роботи з даним кодом в техніці, формування та приймання турбо-кода з використанням обчислювальної потужності технічних пристроїв, а також дослідження ймовірно-енергетичних характеристик процесу передавання інформації з використанням технічних засобів при роботі з турбо-кодами є актуальними проблемами [1, 3].

Мета досліджень

Метою статті є порівняння та оцінювання складності реалізації методів декодування SOVA, MAP, log-MAP та max-log-MAP, які застосовуються для турбо-кодів у розподілених комп'ютерних системах різного функціонального призначення.

Постановка задачі

Недостатність програмного забезпечення вільного розповсюдження та складність алгоритмів декодування турбо-кодів перешкоджають їх впровадженню [2]. Але чи є доцільним застосування розглянутого коду в сучасних розподілених комп'ютерних системах різного функціонального призначення? Щоб відповісти на це питання необхідно розв'язати проблему складності декодування та реалізації цієї процедури під час роботи з турбо-кодом.

Таким чином, необхідно визначити особливості алгоритмів декодування турбо-кодів та проаналізувати їх. Існує достатня кількість наукових праць, які надають ідеї для вирішення поданих проблем, наприклад, [1-7] або інші, пов'язані з даною темою.

Розв'язання поставленої задачі

Для розв'язання задачі використовується узагальнена структура для роботи з інформацією (рис. 1), в якій кодуєчий пристрій (кодер каналу) системи завадостійкого кодування отримує інформаційні символи від джерела і додає до них надлишкові символи таким чином, щоб могла бути виправлена велика частина помилок, що виникають в процесі модуляції сигналів, їх передавання каналом з шумом і демодуляції [4].

Для декодування турбо-кодів можна використовувати декілька методів: SOVA (soft-output Viterbi algorithm – алгоритм Вітербі з "м'яким" виходом); MAP (maximum a posteriori probability – алгоритм декодування за максимумом апостеріорної ймовірності) та його модифікації для зниження обчислювальної складності (log-MAP, max-log-MAP алгоритми).

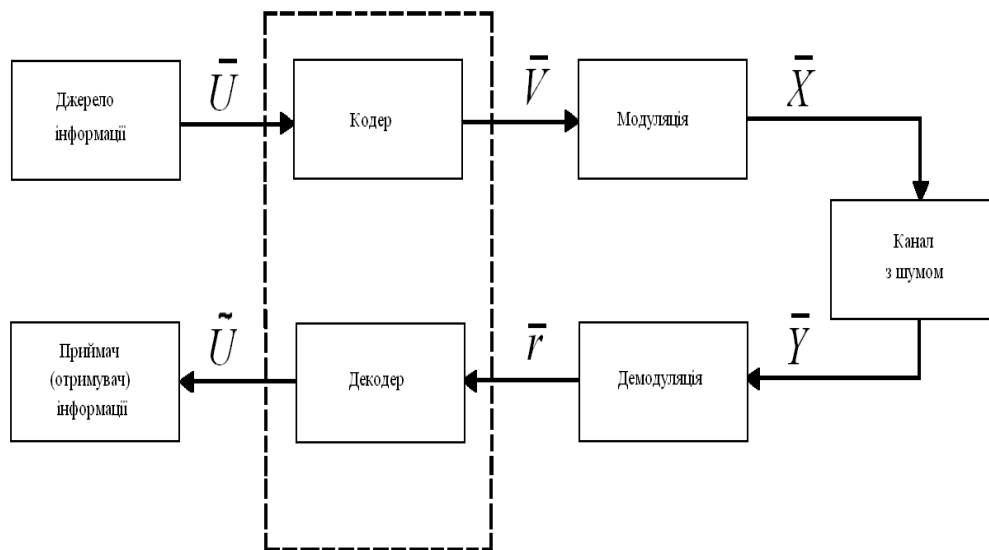


Рисунок 1 — Узагальнена структура для роботи з інформацією: \bar{U} — початкові дані, які надаються пристрою кодування інформації; \bar{V} — дані після процедури кодування; \bar{X} — дані після модуляції, які надходять до каналу із шумом; \bar{Y} — дані на виході каналу з шумом; \bar{r} — дані після демодуляції, які надходять до пристрою декодування інформації; \tilde{U} — інформація з джерела після проходження вузлів системи.

Алгоритм Вітербі з "м'яким" виходом (SOVA) [3] – це категорія м'яких алгоритмів, що базуються на визначенні максимальної правдоподібності, яка дозволяє мінімізувати ймовірність помилки на біт інформації за рахунок прогнозу апостеріорної ймовірності кожного окремого біту. Дана модифікація алгоритму Вітербі [5] дозволяє обчислити надійність інформаційних символів як логарифм відношення правдоподібності (log-likelihood ratio – LLR), що можна записати у формулі

$$LLR(\bar{U}_i) = \log \left(\frac{p(\bar{U}_i = 1 | \bar{r})}{p(\bar{U}_i = 0 | \bar{r})} \right), \quad (1)$$

де \overline{U}_i – інформаційний символ.

Робота декодера SOVA ділиться на дві частини. Перша частина процедури декодування здійснюється аналогічно звичайному алгоритму Вітербі, при цьому вибирається найбільш ймовірна кодова послідовність, яка відповідає шляху, прокладеному на trellis-діаграмі "вперед". Але потрібно зберегти всі метрики на кожному кроці і для кожного стану декодера. У другій частині SOVA алгоритм Вітербі виконується "назад", зберігаючи метрики для кожного стану решітки, при цьому формуються дві метрики для вузла решітки.

На стадії формування "м'якого" виходу алгоритму SOVA для i -ої секції визначається найбільш правдоподібне значення інформаційного символу $\overline{U}_i = a$, де $a \in \{0, 1\}$, і відповідна йому максимальна метрика, яка знайдена на прямому проході за Вітербі, встановлюється рівною $Y_i(\overline{U}_i)$. Метрика найкращого шляху для доповнення $Y_i(\overline{U}_i \oplus 1)$ може бути знайденою за формулою:

$$Y_i(\omega \oplus 1) = \min_{k_1, k_2} \{Y_f(S_{i-1}^{(k_1)}) + BY_i^{(b_1)}(\overline{U}_i \oplus 1) + Y_b(S_i^{(k_2)})\}, \quad (2)$$

де ω – найбільш правдоподібна послідовність, знайдена за алгоритмом Вітербі; $k_1, k_2 \in \{0, 1, 2, \dots, 2^n - 1\}$; $Y_f(S_{i-1}^{(k_1)})$ – метрика шляху, що "вижив" на прямому проході за алгоритмом Вітербі для $(i-1)$ -ої секції і стану $S_{i-1}^{(k_1)}$; $BY_i^{(b_1)}(\overline{U}_i \oplus 1)$ – метрика ребра для інвертованого інформаційного символу, асоційованого з переходом зі стану $S_{i-1}^{(k_1)}$ в $S_i^{(k_2)}$; $Y_b(S_i^{(k_2)})$ – метрика шляху, що "вижив" на зворотному проході за алгоритмом Вітербі для i -го моменту та стану $S_i^{(k_2)}$.

Якщо позначити $Y_i(1) \approx Y_{\max}$ та $Y_i(0)$ назвати максимальною метрикою шляху, асоційованого з доповненням інформаційного символу \overline{U}_i , то можна навести формулу для обчислення м'якого виходу [4, 5].

$$LLR(\overline{U}_i) = Y_i(1) - Y_i(0). \quad (3)$$

Одним із методів, що знижує обчислювальну складність та вирішує складну задачу декодування турбо-кодів, є алгоритм MAP [6], який представлений Bahl, Cocke, Jelinek і Raviv. Це метод оптимального алгоритму посимвольного MAP декодування лінійних блокових кодів, який мінімізує BER (bit error rate – ймовірність помилки на біт) та формує основу для MAP декодування турбо-кодів. Має альтернативну назву VBCJR на честь авторів-розробників.

Ідея MAP декодування [6] полягає в обчисленні апостеріорних ймовірностей інформаційних символів, використовуючи задану прийнятну послідовність \bar{r} та логарифм відношення правдоподібності, що можна записати у формулі

$$LLR(\overline{U}_i) = \log \left(\frac{p(\overline{U}_i = 1 | \bar{r})}{p(\overline{U}_i = 0 | \bar{r})} \right) = \log \left(\frac{\sum_m \sum_{m'} \alpha_{i-1}(m') \cdot \gamma_i^{(1)}(m', m) \cdot \beta_i(m)}{\sum_m \sum_{m'} \alpha_{i-1}(m') \cdot \gamma_i^{(0)}(m', m) \cdot \beta_i(m)} \right), \quad (4)$$

де $\alpha_i(m) = \sum_{m'} \alpha_{i-1}(m') \cdot \sum_{j=0}^1 \gamma_i^{(j)}(m', m)$ – метрика прямого шляху на решітці коду, $\gamma_i^{(j)}(m', m)$ – метрика

ребра, $\beta_i(m) = \sum_{m'} \beta_{i+1}(m') \cdot \sum_{j=0}^1 \gamma_i^{(j)}(m', m)$ – метрика зворотного шляху на решітці коду.

Отже, формула (4) є узагальненим виразом для обчислення LLR процедури MAP декодування. Реалізація MAP декодера виконується за допомогою обчислювальної процедури з формули, наведеної вище. "Жорсткий" розв'язок можна знайти як:

$$\hat{U}_i = \text{sgn}\left(\text{LLR}(\overline{U}_i)\right). \quad (5)$$

Для того, щоб уникнути обчислювальної нестабільності, при роботі з ймовірностями $\alpha_i(m)$ та $\beta_i(m)$ потрібно зберігати умову нормування ймовірностей:

$$p(\alpha_i(m)) + p(\beta_i(m)) = 1. \quad (6)$$

Необхідно зазначити, що метрики ребер залежать від щільності потужності шуму N_0 , тому потрібно знати її оцінку для збереження оптимальності декодера.

Для того, щоб знизити обчислювальну складність MAP алгоритму, можуть бути застосовані логарифми метрик. Алгоритм, який використовує такий перехід, називають log-MAP алгоритмом [7]. Якщо здійснити логарифмування метрик прямого та зворотного шляху на решітці коду та метрики ребра, при цьому ввівши позначення $\log a_i(m) = A_i(m)$; $\log \gamma_i^{(j)}(m', m) = Y_i^{(j)}(m', m)$; $\log \beta_i(m) = B_i(m)$, то у результаті перетворень та спрощень можна отримати алгоритм для процедури log-MAP декодування, який використовується в логарифмічній формі:

$$\begin{aligned} \text{LLR}(\overline{U}_i) = & \log\left(\sum_m \sum_{m'} \exp(A_{i-1}(m') + Y_i^{(1)}(m', m) + B_i(m))\right) - \\ & - \log\left(\sum_m \sum_{m'} \exp(A_{i-1}(m') + Y_i^{(0)}(m', m) + B_i(m))\right). \end{aligned} \quad (7)$$

Якщо застосувати формулу логарифму якобіана, то можна уникнути операції додавання для експоненціальних складових:

$$\begin{aligned} \log(\exp(\Omega) + \exp(\Psi)) = & \max(\Omega, \Psi) + \log(1 + \exp(-|\Omega - \Psi|)) = \\ = & \max(\Omega, \Psi) + f(|\Omega - \Psi|), \end{aligned} \quad (8)$$

де $f(|\Omega - \Psi|) = \log(1 + \exp(-|\Omega - \Psi|))$.

Оскільки функція $\exp(x)$ відносно повільна та коштовна в апаратному виконанні, то для економії часу, спрощення та швидкості обчислень для виразу $f(|\Omega - \Psi|)$ складають таблицю. Достатньо декілька значень, щоб майже досягти точності MAP алгоритму. Для алгоритмів MAP та log-MAP ймовірність помилки мінімальна й однакова.

Модифікований MAP алгоритм, відомий як max-log-MAP [7], використовується для зменшення обчислювальної складності декодування турбо-кодів, але при цьому втрачаються властивості оптимальності. Ідея цієї модифікації заснована на тому, що логарифмується MAP метрика та застосовується апроксимація у вигляді:

$$\log(\exp(\Omega) + \exp(\Psi)) \approx \max(\Omega, \Psi). \quad (9)$$

Використовуючи дану апроксимацію та властивість логарифму дробу, формула логарифма відношення правдоподібності для інформаційного символу \overline{U}_i матиме вид:

$$\begin{aligned} \text{LLR}(\overline{U}_i) \approx & \max_{m', m} \left\{ A_{i-1}(m') + Y_i^{(1)}(m', m) + B_i(m) \right\} - \\ & - \max_{m', m} \left\{ A_{i-1}(m') + Y_i^{(0)}(m', m) + B_i(m) \right\}. \end{aligned} \quad (10)$$

Складність алгоритму max-log-MAP приблизно вдвічі менша, ніж для log-MAP.

Отже, для порівняння та аналізу складності алгоритмів декодування турбо-кодів доцільно скласти таблицю, а також ввести позначення K , яке визначає конструктивну довжину складових кодів (табл.1).

Таблиця 1 – Складність алгоритмів декодування турбо-кодів

Можлива операція	Алгоритм Вітербі	SOVA	log-MAP	max-log-MAP
Порівняння	2^{K-1}	$6 \cdot K$		
Множення на ± 1		8	8	8
Додавання	$2 \cdot 2^{K-1} + 5$	$2 \cdot 2^{K-1} + 8$	$15 \cdot 2^{K-1} + 9$	$10 \cdot 2^{K-1} + 11$
Вибір максимуму		$2^{K-1} + 3 \cdot K$	$5 \cdot 2^{K-1} - 2$	$5 \cdot 2^{K-1} - 2$

На основі складності реалізації методів декодування складових кодів O_1 є можливість провести оцінку складності декодування всього турбо-коду за формулою:

$$O_N = O_1 \cdot k \cdot N, \quad (11)$$

де k – число ітерацій декодування, N – число складових кодів (два та більше).

Результат роботи алгоритмів декодування турбо-кодів (рис. 2) в каналі з AWGN (additive white Gaussian noise — адитивним білим гаусівським шумом) показує значення BER, що прямує до мінімуму, який визначений теоретично можливою межею для даного коду, коли існує безпомилкове передавання та декодування інформації. Для турбо-кодів ці значення максимально наближені до даної межі.

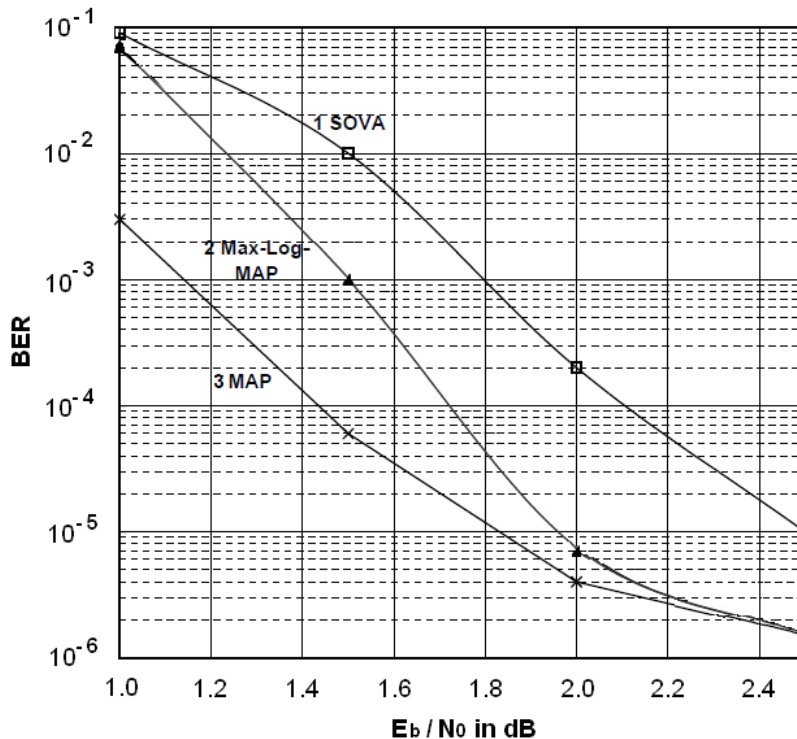


Рисунок 2 – Моделювання енергетичної ефективності та характеристик BER для алгоритмів декодування турбо-кодів

Висновки

Отже, розглянуто алгоритми, що ведуть до розуміння реалізації, складності та принципів декодування турбо-кодів, які стають стандартом для кодування сучасної техніки. Для надійної роботи алгоритм декодування турбо-коду не повинен бути обмеженим поданням на декодери “жорсткої” схеми рішень. Тому, для кращого використання інформації, яка отримана з кожного декодера, застосовують “м’яку” схему декодування.

Для декодування турбо-кодів можна використовувати SOVA – різновид алгоритму Вітербі. Оптимальним є алгоритм MAP декодування турбо-кодів. Доцільно розглядати також модифікації log-MAP та max-log-MAP, які знижують обчислювальну складність процесу декодування. Дослідження систем із застосуванням турбо-кодів актуальні. Представлені алгоритми можуть бути використані для побудови апаратно-програмних засобів декодування сигналів у системах зв’язку.

Проведений аналіз підтверджує перспективність використання турбо-кодів у розподілених комп’ютерних системах різного функціонального призначення.

Список літератури

1. Hanzo L. Turbo Coding, Turbo Equalisation and Space-Time Coding for Transmission over Wireless Channels / L. Hanzo, T.H. Liew, B.L. Yeap. – Southampton: Department of Electronics and Computer Science of UK, 2002. – 746 p. – ISBN 0-470-84726-3.
2. Single- and Multi-Carrier CDMA. Multi-User Detection, Space-Time Spreading, Synchronisation and Standards / L. Hanzo, L-L. Yang, E-L. Kuan, K. Yen – New York: John Wiley & Sons Inc, 2003. – 1077 p. – ISBN 0-470-86309-9.
3. Кулик А.Я. Турбо-код: властивості та застосування для завадостійкого кодування / А.Я. Кулик, С.Г. Кривогубченко, Ю.Ю. Іванов // Оптико-електронні інформаційно-енергетичні технології. – 2011. – № 2(22). – С. 25-33.
4. Морелос-Сарагоса Р. Искусство помехоустойчивого кодирования. Методы, алгоритмы, применение: пер. с англ.; под ред. В.Б. Афанасьева / Р. Морелос-Сарагоса. — М.: Техносфера, 2005. — 320 с. — ISBN: 5-94836-035-0.
5. Витерби А.Д. Принципы цифровой связи и кодирования: пер. с англ.; под ред. К.Ш. Зигангирова / А.Д. Витерби, Дж.К. Омура. — М.: Радио и связь, 1982. — 536 с.
6. Bahl, Cocke, Jelinek and Raviv (BCJR) Algorithm [Web resource] / BCJR Algorithm. – Access mode: http://www.comlab.hut.fi/opetus/311/Bahl_mod.pdf.
7. Valenti M. An Efficient Software Radio Implementation of the UMTS Turbo Codec [Web resource] / Matthew C. Valenti. – Access mode: <http://www.csee.wvu.edu/~mvalenti/documents/valenti2001c.pdf>.

Відомості про авторів

Кулик Анатолій Ярославович – д.т.н., професор, професор кафедри автоматики та інформаційно-вимірювальної техніки, Вінницький національний технічний університет, (0432)598-437, kulyk@inaeksu.vinnica.ua, Хмельницьке шосе, 95, м. Вінниця, Україна, 21021.

Кривогубченко Сергій Григорович – к.т.н., доцент, доцент кафедри автоматики та інформаційно-вимірювальної техніки Вінницького національного технічного університету, (0432) 598-437, kriv@inaeksu.vinnica.ua, Хмельницьке шосе, 95, м. Вінниця, Україна, 21021.

Іванов Юрій Юрійович – студент факультету автоматики та комп’ютерних систем управління ІнАЕКСУ Вінницького національного технічного університету, (0432) 546-443, YuraII@yandex.ru, Хмельницьке шосе, 95, м. Вінниця, Україна, 21021.

УДК 004.9:81'374.82

Е.В.ПОТАПОВА

Таврический национальный университет им.В.И Вернадского, Симферополь

ИНСТРУМЕНТАЛЬНАЯ СИСТЕМА ДЛЯ СОЗДАНИЯ МНОГОЯЗЫЧНОЙ ОНТОЛОГИИ ПРЕДМЕТНОЙ ОБЛАСТИ

Анотація. У статті подано концептуальну модель багатомовної онтології з Фізики магнітних явищ як лексикографічної системи особливого типу. На її підставі розроблено інструментальну систему для створення і управління багатомовною онтологією на основі засобів і методів СУБД. Програмний інтерфейс і алгоритм опису термінів не залежать від предметної галузі й можуть бути використані для інших предметних галузей.

Ключові слова: програмний інтерфейс, багатомовна онтологія, лексикографічна система, електронний словник.

Аннотация. В статье представлена концептуальная модель многоязычной онтологии по Физике магнитных явлений как лексикографической системы особого типа. На основе данной модели разработана инструментальная система для создания и управления многоязычной онтологией на основе средств и методов СУБД.

Ключевые слова: программный интерфейс, многоязычная онтология, лексикографическая система, электронный словарь.

Abstract. The article presents a conceptual model for multilingual domain ontology (Physics of the magnetic phenomena) as lexicographic system of special type. On the base of this model the instrumental system for multilingual ontology creating and managing was designed. The instrumental system based on the tools and methods of the data base management systems. The program interface and the algorithm of the terms description does not depend on domain and they can be used to create the multilingual ontology for other domains.

Keywords: programming interface, multilingual ontology lexicographical system, electronic dictionary.

Введение

Создание многоязычных онтологий имеет практическое значение в области согласования терминологии в различных предметных областях (ПрО), развития технологий информационного поиска (организации многоязычного поиска), машинного перевода в рамках заданной предметной области и т.д. Одной из важных проблем в данном направлении является моделирование взаимодействия терминологических систем для нескольких (более двух) языков. Применение теории лексикографических систем [1, 2] к концептуальному моделированию онтологии предметной области «Физика магнитных явлений» как лексикографической системы особого типа позволило создать концептуальную модель многоязычной онтологии ПрО [3]. Предложенная модель реализована в виде лингвистической базы знаний по Физике магнитных явлений с представлением информации на русском, украинском и английском языках. В статье представлен опыт разработки инструментальной системы для создания и управления трехязычной (русский-украинский-английский) лингвистической онтологией «Физика магнитных явлений» (ФМЯ). Концептуальная модель многоязычной онтологии, разработанная ранее [3, 4], здесь приводится кратко, для связности изложения.

Постановка цели и задачи научного исследования

Цель исследования: разработка методов описания предметных и лингвистических знаний на примере предметной области «Физика магнитных явлений» с помощью трехязычной лингвистической онтологии, а также методов и инструментальных средств управления лингвистической онтологией.

Задачи исследования:

- 1) Создание инструментальной среды, реализующей предложенную концептуальную модель.
- 2) Разработка интерфейса настройки на определенную предметную область.
- 3) Обеспечение визуализации онтологии в виде комплекса словарных статей и когнитивных карт терминов в электронном виде.

Концептуальная модель

Формальная модель онтологии — это упорядоченная тройка конечных множеств [5]:

$$O = \langle T, R, F \rangle, \quad (1)$$

где T — термины ПрО, R — отношения между терминами, F — функции интерпретации, заданные на терминах и/или отношениях онтологии ПрО.

В результате концептуального моделирования предметной области ФМЯ посредством лингвистической онтологии предложена оригинальная модель представления знаний ПрО с расширенным описанием ассоциативных – проблемно-специфических связей (подробнее в [6]). Для определения типов проблемно-специфических связей было введено понятие «лексико-онтологического класса» (ЛОК) представляющего собой некоторое нечеткое множество, исследованы особенности перехода терминов онтологии между классами и существование нечетких семантических состояний [4]. Для их описания в концептуальной онтологической модели ПрО ФМЯ использованы нечеткие множества Л. Заде и теория семантических состояний Колмогорова-Широкова [3]. Определена функция, описывающая нечеткое семантическое

состояние понятия онтологии. В формальную модель онтологии были введены параметры, отражающие данную ситуацию.

Таким образом, лексикографическая параметризация для любого термина $t \in \langle T \rangle$ из формулы (1) включает в себя следующие параметры:

$$t(t^{L1}, t^{L2}, t^{L3}, t^{OC} \langle t^G \rangle, t^D, \langle t^C \rangle, \langle t^I \rangle, \langle t^R \rangle), \quad (2)$$

где t^{L1} – орфографический стандарт (русский язык - $L1$), t^{L2} – переводной эквивалент на язык $L2$ (украинский язык), t^{L3} – переводной эквивалент на язык $L3$ (английский язык). t^{OC} - ЛОК термина, значение которого является элементом $\langle OntC \rangle$ - множества ЛОК установленных для данной ПрО. $\langle t^G \rangle$ - множество грамматических параметров, t^D – дефиниция текстовая, $\langle t^C \rangle$ – множество контекстных примеров, $\langle t^I \rangle$ - множество вспомогательной информации. $\langle t^R \rangle$ - множество связей с другими терминами - онтологическая окрестность термина.

Однако данная модель не разрешает проблему установления адекватного соответствия между переводными эквивалентами трех языков, так как в лексикографической параметризации термина онтологии параметры t^{L2} и t^{L3} могут содержать несколько значений. Для этого концептуальная модель онтологии была переработана в рамках теории лексикографических систем (Л-систем) [3].

В типологическом смысле Л-система – это информационный объект, сочетающий в себе черты модели данных, модели знаний и логико-лингвистического исчисления [1,2]. Основными системообразующими отношениями Л-системы являются: «субъект-объект» и «форма-содержание». Основным системообразующим инвариантом Л-системы является лексикографический эффект в информационных системах. Лингвистическая онтология предметной области отличается от других Л-систем тем, что помимо языкового уровня развития лексикографических эффектов, имеется надязыковой уровень – уровень метаописаний понятий и их отношений. Такое разделение системы на два уровня позволяет создать модель для лингвистической онтологии предметной области на двух и более языках с общим ядром понятий и отношений между ними.

Преобразование концептуальной модели 3-х язычной онтологии ПрО в лексикографическую систему (LS) позволило выделить в ней два уровня: $LS \rightarrow (LS^{L1}, LS^{L2}, LS^{L3})$ и определить взаимодействие уровней и компонентов полученной системы [3].

Уровень 1 в целом соответствует формальной метаонтологии понятий ПрО и содержит информацию о системе, не связанную с конкретным языком. Множество объектов $Ob(LS) = \{T, OntC, Pers, Br, R\}$, где T -множество терминологических понятий (метаописаний - индексов), R – множество типов связей между понятиями, $OntC$ – множество ЛОК. Множество персоналий - $Pers$ и множество разделов Физики магнитных явлений - Br соответствуют $\langle t^I \rangle$ - множеству дополнительной информации.

Множество отношений $RelOb(LS) = \{TT, TPers, TBr\}$, где TT – множество связей между понятиями, что в онтологической модели соответствует $\langle t^R \rangle$. $TPers$ - множество связей между множеством понятий и множеством персоналий, TBr - множество связей между множеством понятий и множеством разделов ПрО.

Уровень 2 – отображение формальной онтологии понятий в языковые системы. Множество объектов $Ob(LS^{Li}) = \{T^{Li}, TCont^{Li}, TDef^{Li}\}$, где T^{Li} -множество терминов языка Li (в онтологической модели это множества t^{L1}, t^{L2}, t^{L3}); для $\forall t \in \langle T^{Li} \rangle$ $t(t^{Li}, \langle t^G \rangle, tDef^{Li})$, где t^{Li} - орфографический стандарт, T^G - множество грамматических характеристик термина ($\langle t^G \rangle$ в онт.м.). Текстовая дефиниция термина - $tDef^{Li}$ (t^D в онт.м.) и $\langle TCont^{Li} \rangle$ -множество контекстов термина на языке Li ($\langle t^C \rangle$ в онт.м.).

Множество отношений $RelOb(LS^{Li}) = \{TTCont^{Li}\}$ - это множество связей между множеством терминов и множеством контекстных примеров (коллекцией текстов).

Ограничение целостности заключается в том, что :

- 1) никакой элемент из T^{Li} уровня 2 не может существовать без связи с соответствующим элементом T уровня 1;
- 2) одному элементу T уровня 1 может соответствовать несколько элементов из T^{Li} уровня 2 в каждом из языков;
- 3) связь между объектами (терминами) различных языков (т.е. перевод) осуществляется только через элемент T уровня 1.

Таким образом, уровень 1 концептуальной модели содержит формальные данные онтологии, в то время как уровень 2 содержит лингвистическую часть информации о предметной области. Такое разделение позволяет проектировать многоязычную систему, для 2-х и более языков. Взаимодействие между комплексами лингвистических данных различных языков осуществляется только через формальную онтологию уровня 1.

Описанная концептуальная модель многоязычной лингвистической онтологии справедлива только для терминологических систем предметных областей и не используется для моделирования общезыковой лингвистической онтологии для двух и более языков.

Формула словарной статьи

Процесс абстрагирования словарной (лексикографической) структуры представляет собой своеобразную расшифровку, реконструкцию того лексикографического эффекта (ЛЭ), который привел к образованию этой структуры. Вследствие того, что в лингвистической онтологии наблюдается два уровня ЛЭ, структура словарной (онтологической) статьи для некоторого понятия T выраженного термином T^{Li} в языке $L1$ с грамматическими параметрами G^{L1} имеет вид :

$$T^{L1}(T, G^{L1}) = C^O + C^L ;$$

$$C^O = OntC + TT + TPers + TBr ;$$

$$C^L = S^{L1} + T^{L2} + T^{L3} + TDef^{L1} + TCont^{L1} , \text{ где } T^{L2} \supseteq S^{L2} \text{ и } T^{L3} \supseteq S^{L3} ,$$

где C^O - комплекс онтологических параметров, C^L - комплекс лингвистических параметров.

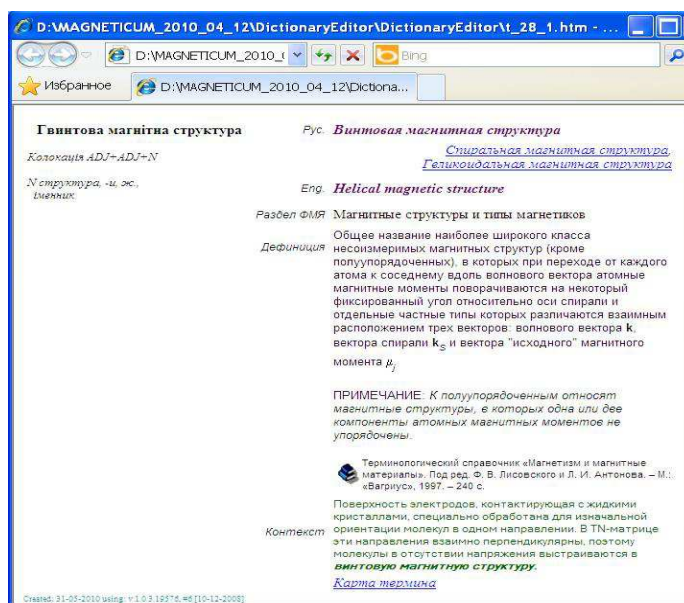


Рисунок 1 – Интерфейс пользователя. Словарная статья термина

Разделы переводных эквивалентов T^{L2} и T^{L3} на языки $L2$ и $L3$ соответственно – включают в себя весь синонимический ряд (классы условных эквивалентов S^{L1}, S^{L2}, S^{L3} соответственно). Остальные обозначения соответствуют вышеописанной концептуальной модели.

При отображении словарной статьи термина для пользователя в явном виде присутствуют лингвистические параметры. Из онтологических параметров отображаются только $TPers$ и TBr – персоналии и раздел ФМЯ связанные с понятием, выраженным данным термином. Информация об онтологических связях термина отображается в виде когнитивной карты термина – фрагмента онтологической схемы (подробнее в [7]). На рис.1- 2 - пример словарной статьи и карта термина.

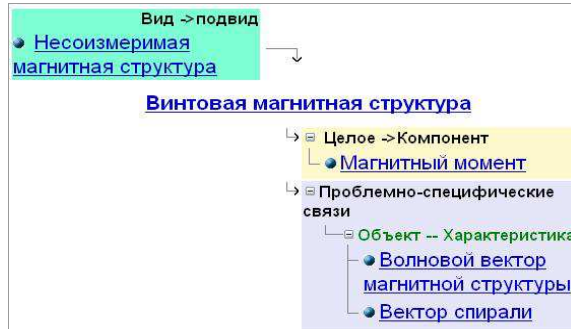


Рисунок 2 – Интерфейс пользователя. Пример когнитивной карты термина

Инструментальная система

Внутренняя модель. Исходя из двухуровневой концептуальной модели изложенной выше, в структуре онтологической лексикографической базы данных (ЛБД) данные, не соотносимые с конкретным языком (связи терминологических понятий, набор лексико-онтологических классов, персоналии, раздел ФМЯ) образуют отдельную группу таблиц. С другой стороны выделены три языковые группы таблиц: английская, русская и украинская (рис.3).

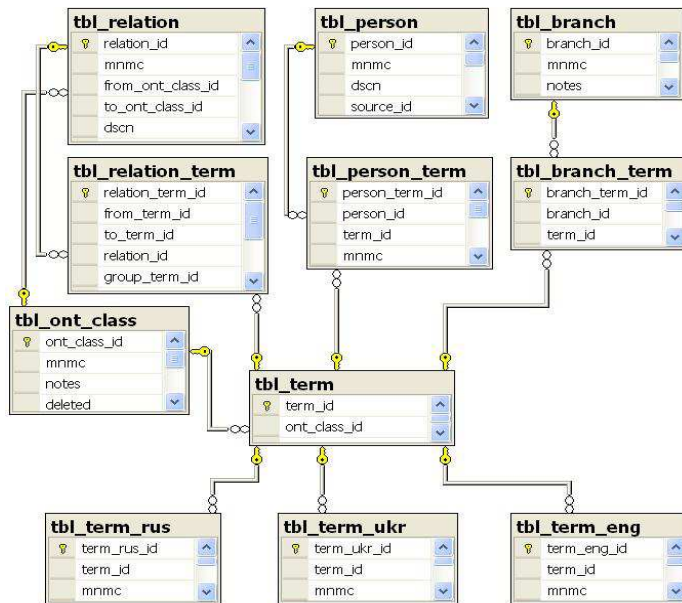


Рисунок 3 – Общая схема лексикографической базы данных

Главным реестром, связывающим все группы таблиц, является таблица `tbl_term`, которая содержит индексы терминологических понятий. Таблицы `tbl_term`, (`tbl_relation_term`, `tbl_relation`, `tbl_ont_class` – связи понятий и онтологические классы), (`tbl_person_term`, `tbl_person` – информация о персоналиях связанная с открытием или исследованием тех или иных физических явлений), (`tbl_branch_term`, `tbl_branch` – перечень разделов физики магнитных явлений) соответствуют 1 уровню концептуальной модели – формальной метаонтологии. Все множества объектов хранятся в отдельных таблицах. Связь между множествами объектов (связь типа «многие-ко-многим») обеспечивается через таблицы составных (сложных) ключей `tbl_relation_term`, `tbl_person_term`, `tbl_branch_term`.

Второй уровень образуют языковые группы таблиц, связанные друг с другом через таблицу `tbl_term`. Рассмотрим на примере группы таблиц русского языка: `tbl_term_rus` – главный реестр терминов, каждый термин может иметь одно текстовое определение (`tbl_term_rus_dscn`) и несколько входов в коллекцию научных текстов (`tbl_collection_rus`). В то же время каждый текст из `tbl_collection_rus` может быть связан с несколькими терминами, что обеспечено сложным ключом `tbl_term_collection_rus`. Список библиографических источников содержится в таблице `tbl_source`.

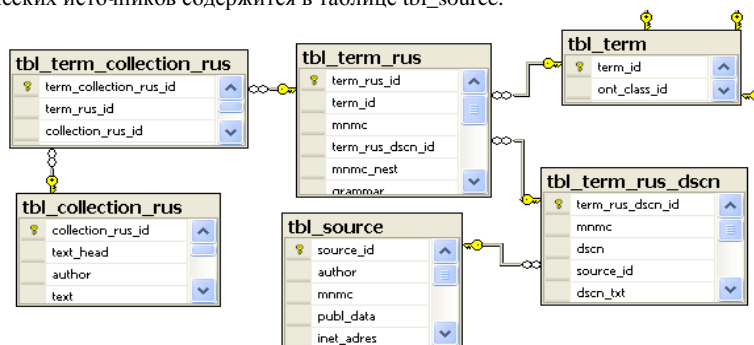


Рисунок 4 – Группа таблиц русского языка

Ввиду наличия формул и рисунков в текстовых описаниях терминов, в коллекции текстов и текстах представляющих персоналии, для хранения такой смешанной информации использован формат *rtf*.

Внешнее представление и инструментальная Л-система. Редактор онтологии реализует концептуальную модель и предоставляет возможность управления данными ЛБД представленными на Рис.3-4. Редактор реализован как локальное Windows-приложение (C#) с доступом к источнику данных (SQL Server).

Графический интерфейс редактора лексикографической базы разработан в соответствии с концептуальной моделью и структурой словарной статьи. Главное окно предоставляет пользователю выбор режима работы: 1)Онторедактор, 2)Редактор языковых реестров.

Функции **Онторедактора** :

- 1) ввод нового понятия (должен быть введен хотя бы один языковой эквивалент);
- 2) установку переводных соответствий (связь между языковыми реестрами);
- 3) установка связей понятия с другими понятиями (рис.5.);
- 4) установка связей с персоналиями, с разделом ФМЯ;
- 5) демонстрация онтологической схемы связей понятия (карты термина).

Функции **Редактора языковых реестров**:

- указание грамматических характеристик термина; модель коллокации; ввод условных языковых эквивалентов (синонимов);
- ввод текстовой дефиниции;
- установка связей с коллекцией текстов.

Кроме того здесь тоже есть вкладка «Связи» для поиска и установки связей термина в результате анализа текстовой дефиниции или текста из коллекции текстов. Поиск связи автоматизирован, тип связи определяет эксперт. Введенные данные отображаются в виде словарной статьи термина и его когнитивной карты (рис.1,2).

Для всех множеств объектов уровня 1 и уровня 2 концептуальной модели существуют отдельные окна для редактирования: «Редактор связей», «Онтологические классы», «Персоналии», «Разделы ФМЯ», «Коллекция текстов», «Библиографические источники».

Перечислим технические и процедурные особенности разработанной инструментальной системы и онтологической лексикографической базы данных:

1. Модульная структура: позволяет использовать отдельные библиотеки объектов и фрагменты (группы таблиц БД) как отдельные приложения или для модернизации существующих информационных систем. Т.е. разработанная инструментальная система позволяет разрабатывать информационные ресурсы следующих типов:

- 3-х язычная онтология ПрО с лингвистическим компонентом описания информационных единиц и когнитивными картами терминов, совокупность, которых представляет собой графическое отображение онтологии ПрО;
- Формальная онтология понятий (реестр понятий должен быть выражен на одном из языков) + карты терминов;
- 3-х язычный терминологический словарь;
- Словарь терминов + коллекция текстов, размеченная этими терминами (реализовано для русского языка).

2. Адаптация многоязычной лингвистической онтологии ПрО под определенную предметную область включает в себя, помимо создания словаря ПрО, создание классификации понятий и определение типов проблемно-специфических связей между понятиями на основе заданной классификации. Кроме того, инструментальная система позволяет настраивать таксономические связи.

3. Реализована функция определения онтологического окружения термина на основе теории нечетких семантических состояний и системы лексико-онтологических классов (подробнее в [4]). Результат может быть использован для уточнения поисковых запросов.

4. Доказательство и интроспекция БЗ обеспечивается структурой SQL запросов и другими возможностями СУБД.

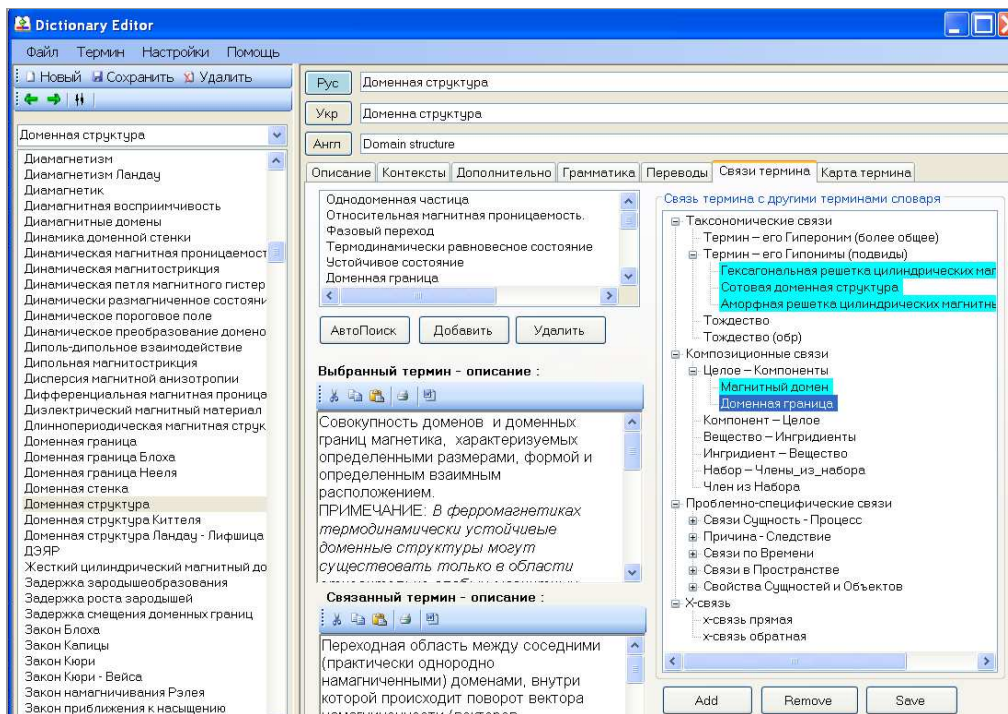


Рисунок 5 - Инструментальная система. Онторедактор. Вкладка «Связи термина»

Посредством разработанной инструментальной системы была проведена работа по установлению связей между понятиями в ряде разделов Физики магнитных явлений и получены соответствующие схемы связей – карты терминов, которые являются взаимосвязанными фрагментами онтологической схемы предметной области. На данном этапе существует 276 карт терминов. Карты терминов генерируются инструментальной системой в автоматическом режиме в формате HTML.

Пользовательская версия. На любом этапе редактирования ЛБД содержимое онтологической базы знаний может быть опубликовано в формате html для конечного пользователя в виде отдельного

электронного продукта «Словарь терминов по ФМЯ» на трех языках (русский-украинский-английский) с когнитивными картами терминов.

Заключение

Расширение онтологической модели ПрО «Физика магнитных явлений» средствами теории лексикографических систем позволило создать модель многоязычной лингвистической онтологии ПрО как лексикографической системы особого типа.

Реализация предложенной концептуальной модели многоязычной лингвистической онтологии ПрО в виде инструментальной среды интегрированной с СУБД позволяет успешно решать ряд задач управления онтологией (логический вывод, нахождение противоречий) и пополнения онтологии как вручную, так и путем экстракции терминов и их связей из текстов (реализовано для русского языка).

Список литературы

1. Широков В.А. Элементы лексикографії.– К.: Довіра, 2005. – 304с.
2. Широков В.А. Інформаційна теорія лексикографічних систем. – К.:Довіра, 1998. – 331с.
3. Широков В.А., Потапова Е.В. Онтология предметной области как лексикографическая система особого типа. //«Казанская наука», Казань. №12 2012, С.209-213.
4. Потапова Е.В. Модель лингвистической онтологии предметной области с нечеткими семантическими состояниями терминов. // Научно-технический журнал «Бионика интеллекта», Харьков, ХНУРЭ, № 2(79), 2012, с.95-102.
5. Добров Б.В., Соловьев В.Д., Лукашевич Н.В., Иванов В.В. Онтологии и тезаурусы. Модели, инструменты, приложения. Бинوم, 2009. - 173 с.
6. Бержанский В.Н., Потапова Е.В. Классификация связей между понятиями в онтологии по физике магнитных явлений. / MegaLing 2010: сб.научн. трудов – К.: Довіра, 2010, С.12-21.
7. Бержанский В.Н., Дикарева С.С., Потапова Е.В. Принципы конструирования когнитивных карт терминов в многоязычной лексикографической системе по физике магнитных явлений./ Актуальні проблеми прикладної лінгвістики. Матеріали міжнародної Інтернет-конференції (23-24 червня 2011). – Умань, 2011, С. 23-26.

Сведения об авторе

Потапова Елена Владимировна - м.н.с. Межведомственного центра когнитивной и прикладной лингвистики Научно-исследовательской части Таврического национального университета им.В.И. Вернадского, e-mail: helen1pota@rambler.ru , Украина, АР Крым, 95050 Симферополь, ул.Киевская, д.147, кв. 51.

УДК 004.896

ШУШУРА А.Н., ТАРАСОВА І.А.

Донецкий национальный технический университет, Донецк

СПОСОБ ЗАДАНИЯ МНОГОМЕРНЫХ ФУНКЦИЙ ПРИНАДЛЕЖНОСТИ ТЕРМОВ
ЛИНГВИСТИЧЕСКИХ ПЕРЕМЕННЫХ

Анотация. У статті наведено спосіб побудови багатомірних функцій належності термів лінгвістичних змінних на основі методів нечіткої кластеризації. Проведено чисельне дослідження способу в задачі автоматизації процесу діагностування і введення медикаментів при лікуванні вагітних жінок із преєклампсією різного ступеня тяжкості. Застосування розробленого способу дозволяє підвищити ефективність нечіткого управління об'єктами, що погано формалізуються, зі складною структурою взаємозв'язків вхідних і вихідних змінних, знизивши вплив суб'єктивності експерта за рахунок автоматизації процесу побудови функцій належності термів лінгвістичних змінних з використанням кластерного аналізу.

Ключові слова: нечітке управління, багатомірні функція належності, нечітка кластеризація, терм лінгвістичної змінної, автоматизація процесу.

Аннотация. В статье приведен способ построения многомерных функций принадлежности термов лингвистических переменных на основе методов нечеткой кластеризации. Проведено численное исследование способа в задаче автоматизации процесса диагностики и введения медикаментов при лечении беременных женщин с преэклампсией различной степени тяжести. Применение разработанного способа позволяет повысить эффективность нечеткого управления плохо формализуемыми объектами со сложной структурой взаимосвязей входных и выходных переменных, снизив влияние субъективности эксперта за счет автоматизации процесса построения функций принадлежности термов лингвистических переменных с использованием кластерного анализа.

Ключевые слова: нечеткое управление, многомерная функция принадлежности, нечеткая кластеризация, терм лингвистической переменной, автоматизация процесса.

The Abstract. The article is devoted to method of constructing multidimensional membership functions of terms of linguistic variables on the basis of fuzzy clustering has been done. A numerical study of the method in the problem of automation of the process of diagnosis and medicines introduction for the treatment of pregnant women with preeclampsia of different severity has been given. Application of the developed method improves the efficiency of fuzzy control of badly formalized objects with the complicated structure of intercommunications of entrance and output variables, reducing the influence of the subjectivity of the expert by automation of the construction process of the membership functions of linguistic variables terms with the usage of cluster analysis.

Key words: fuzzy control, multidimensional membership function, fuzzy clustering, term of the linguistic variable, automation of the process.

Введение

Вопрос о построении функций принадлежности является одним из самых важных вопросов в теории нечетких множеств, которому посвящены работы С.Д. Штовбы, А.П. Ротштейна, А.Е. Алтунина, В.В. Борисова [1-4]. Существующие методы нечеткого управления в основном используют нечеткие переменные с функциями принадлежности одного аргумента. Это позволяет использовать простое и наглядное представление функций принадлежности (например, треугольное или трапециевидное), а также применять несложные вычислительные процедуры при проведении всех этапов нечеткого вывода. Однако использование нечетких переменных с функциями принадлежности одного аргумента часто приводит к потере зависимости между управляющими переменными, которая обусловлена наличием ограничений на управление в некоторых задачах. Кроме того, лингвистические переменные могут иметь достаточно сложную физическую природу, требующую при определении их значений использовать несколько связанных величин [5].

Способы построения функций принадлежности существенно зависят от экспертного мнения. Методы задания и определения вида многомерных функций принадлежности в настоящее время достаточно слабо развиты. Рассмотренные в работах [1-4] представления многомерных функций принадлежности позволяют построить функции принадлежности заранее известного вида и не обеспечивают задания произвольной формы областей, в которых определены лингвистические термы переменных. Недостатком изложенного метода в работе [6] можно считать то, что использование нейронных сетей привело к отсутствию возможности выделить как саму функцию принадлежности, так и базу правил, а также функции, описывающие заключения каждого конкретного правила.

При формировании многомерных функций принадлежности с целью снижения влияния субъективности эксперта могут быть использованы методы нечеткой кластеризации.

Целью данной работы является повышение эффективности управления плохо формализованными объектами за счет автоматизации процесса построения функций принадлежности термов лингвистических переменных на основе нечеткой кластеризации.

Для решения поставленной цели в работе решаются следующие задачи:

- обзор методов нечеткой кластеризации и выбор одного из них для формирования многомерных функций принадлежности термов;
- формализация способа задания многомерных функций принадлежности термов лингвистических переменных с использованием кластерного анализа;

– численное исследование разработанного способа задания многомерных функций принадлежности в задаче автоматизации процесса диагностирования и введения медикаментов при лечении беременных женщин с преэклампсией различной степени тяжести.

Обоснование выбора метода нечеткой кластеризации для задания многомерных функций принадлежности

Концептуальная взаимосвязь между кластерным анализом и теорией нечетких множеств основана на том, что при решении задач структуризации сложных систем большинство формируемых классов объектов размыты по своей природе. Эта размытость состоит в том, что переход от принадлежности к непринадлежности элементов к данным классам постепенен. Требование нахождения однозначной кластеризации элементов исследуемой проблемной области является достаточно грубым и жестким, особенно при решении плохо или слабо структурированных задач. Методы нечеткой кластеризации ослабляют это требование за счет введения в рассмотрение нечетких кластеров и соответствующих им функций принадлежности, что позволяет применять их при задании функций принадлежности термов лингвистических переменных.

Различные аспекты нечеткой кластеризации исследовались в работах [7-8]. Нечеткий подход к решению задачи кластеризации в ряде случаев позволяет разделить кластеры сложной формы и открывает новые возможности интерпретации результатов кластеризации. В отличие от вероятностного пространства, структура нечетких множеств представлена не булевой решеткой, а векторной, так что с точки зрения формальной логики для нечетких классов не выполняются такие законы, как закон противоречия и закон исключения третьего. Одним из наиболее распространенных методов нечеткой кластеризации является метод нечетких s -средних. Этот метод в силу высокой степени обобщенности стал основой для ряда модификаций, отличающихся формой выделенных кластеров. Однако следует учитывать, что модификации метода стремятся создавать кластеры определенных форм, даже если в исследуемом наборе данных таких кластеров на самом деле нет. Учитывая высокую распространенность и общность метода нечетких s -средних при необходимой простоте численной реализации, предложено его использовать в задаче формирования многомерных функций принадлежности.

Формализация способа задания многомерных функций принадлежности

Способ задания многомерных функций принадлежности термов лингвистических переменных включает следующие этапы:

- выделение лингвистических переменных и сопоставление им характеристик объекта моделирования;
- сбор статистических данных о деятельности объекта моделирования;
- нечеткая кластеризация статистических данных по каждой лингвистической переменной;
- формирование термов лингвистической переменной путем присваивания полученным кластерам названий;
- формализация функций принадлежности термов.

Исходной информацией являются статистические данные о деятельности объекта моделирования, которые сгруппированы относительно выделенных лингвистических переменных и представлены в виде матриц, число которых соответствует количеству лингвистических переменных:

$$X_l = \begin{bmatrix} x_{11}^l & x_{12}^l & \dots & x_{1n}^l \\ x_{21}^l & x_{22}^l & \dots & x_{2n}^l \\ & & \dots & \\ x_{q1}^l & x_{q2}^l & \dots & x_{qn}^l \end{bmatrix}, \quad (1)$$

где X_l – матрица экспериментальных данных, относящихся к l -ой лингвистической переменной; q – число экспериментов; n – число измеренных характеристик, сопоставленных с лингвистической переменной.

Так как сопоставленные лингвистическим переменным характеристики могут быть измерены в разных единицах, необходимо их нормирование путем деления центрированной величины на среднее квадратическое отклонение. Таким образом переходим от матрицы X_l к нормированной матрице X_l^* с элементами:

$$x_{ij}^{l*} = \frac{x_{ij}^l - \bar{x}_j}{s_j} \quad (2)$$

где x_{ij}^l – значение j -ой характеристики i -го эксперимента, относящегося к l -ой лингвистической переменной; \bar{x}_j – среднее значение j -ой характеристики; s_j – среднее квадратическое отклонение j -ой характеристики:

$$s_j = \sqrt{\frac{1}{q} \sum_{i=1}^q (x_{ij}^l - \bar{x}_j)^2} \quad (3)$$

Задача кластеризации состоит в разбиении множества данных об экспериментах на группы «похожих» между собой. В n -мерном метрическом пространстве признаков мерой «сходства» двух экспериментов будем считать расстояние между ними. Применение нечёткой кластеризации позволяет каждому эксперименту принадлежать с различной степенью нескольким кластерам. Число кластеров k определяется на основании экспертных знаний ведущих специалистов в моделируемой предметной области. Кластерная структура задаётся матрицей принадлежности M_l :

$$M_l = \begin{bmatrix} \mu_{11}^l & \mu_{12}^l & \dots & \mu_{1q}^l \\ \mu_{21}^l & \mu_{22}^l & \dots & \mu_{2q}^l \\ & & \dots & \\ \mu_{k1}^l & \mu_{k2}^l & \dots & \mu_{kq}^l \end{bmatrix}, \quad (4)$$

где μ_{ij}^l – степень принадлежности j -го набора данных i -му терму l -той лингвистической переменной.

Матрица принадлежности должна удовлетворять следующим условиям [8]:

- а) $\mu_{ij}^l \in [0, 1]$, $i = \overline{1, k}$, $j = \overline{1, q}$,
- б) $\sum_{i=1}^k \mu_{ij}^l = 1$, $j = \overline{1, q}$,
- в) $0 < \sum_{j=1}^q \mu_{ij}^l < q$, $i = \overline{1, k}$.

Для оценки качества разбиения будем использовать критерий разброса, показывающий сумму расстояний от экспериментов до центров кластеров с соответствующими степенями принадлежности:

$$J_l = \sum_{i=1}^k \sum_{j=1}^q (\mu_{ij}^l)^w d(v_i^l, x_j^{l*}), \quad (5)$$

где $d(v_i^l, x_j^{l*})$ – Евклидово расстояние между экспериментом и центром кластера

$$x_j^{l*} = (x_{j1}^{l*}, x_{j2}^{l*}, \dots, x_{jn}^{l*}) \quad (6)$$

$$v_i^l = (v_{i1}^l, v_{i2}^l, \dots, v_{in}^l). \quad (7)$$

$w \in (1, \infty)$ – экспоненциальный вес, определяющий нечёткость, размытость кластеров.

Матрица V_l – матрица координат центров кластеров:

$$V_l = \begin{bmatrix} v_{11}^l & v_{12}^l & \dots & v_{1n}^l \\ v_{21}^l & v_{22}^l & \dots & v_{2n}^l \\ & & \dots & \\ v_{k1}^l & v_{k2}^l & \dots & v_{kn}^l \end{bmatrix}. \quad (8)$$

Элементы матрицы V_l вычисляются по формуле:

$$v_{im}^l = \frac{\sum_{j=1}^q (\mu_{ij}^l)^w x_{jm}^{l*}}{\sum_{j=1}^q (\mu_{ij}^l)^w}, \quad m = \overline{1, n}. \quad (9)$$

Задачей является нахождение матрицы M_l , минимизирующей критерий J_l . Для этого используется алгоритм нечётких с-средних, в основе которого лежит метод множителей Лагранжа.

На первом шаге матрица принадлежности M_l , удовлетворяющая описанным выше условиям, генерируется случайным образом. Далее запускается итерационный процесс вычисления центров кластеров и пересчёта элементов матрицы степеней принадлежности:

$$\mu_{ij}^l = \frac{1}{\left(\left(d_{ij}^2 \right) \sum_{m=1}^k \frac{1}{d_{mj}^2} \right)^{\frac{1}{w-1}}} \quad \text{при } d_{ij} > 0, \quad (10)$$

$$\mu_{mj}^l = \begin{cases} 1, & m = i \\ 0, & m \neq i \end{cases} \quad \text{при } d_{ij} = 0, \quad (11)$$

где $d_{ij} = d(v_i^l, x_j^{l*})$ для $i = \overline{1, k}$, $j = \overline{1, q}$.

Вычисления продолжают до тех пор, пока изменение матрицы M , характеризующееся величиной $\|M - M^*\|^2$, где M^* – матрица на предыдущей итерации, не станет меньше заранее заданного параметра остановки ε .

Термы лингвистических переменных образуются путем присвоения полученным кластерам названий. Для каждого термина задается таблично многомерная функция принадлежности в виде матрицы, включающей матрицу значений аргументов X_l и столбец значений функции, в качестве которых выступает соответствующая терму строка матрицы M_l . Расчет значений функции принадлежности может быть выполнен на основе методов многомерной аппроксимации.

Для разработанного способа формирования многомерных функций принадлежности проведено численное исследование.

Формирование функций принадлежности термов в задаче управления лечением пациентов с преэклампсией различной степени тяжести

Одной из задач, характеризующихся рядом нелинейных ограничений на управление, является задача автоматизации процесса диагностирования и введения медикаментов при лечении беременных женщин с преэклампсией различной степени тяжести. За период 2009-2012 гг. при обследовании больных, находившихся на лечении в отделениях интенсивной терапии, были собраны статистические данные о медикаментозном влиянии на комплекс регистрируемых показателей. Выделены лингвистические переменные с соответствующей группировкой измерений по относящимся к ним величинам, определяющим состояние больного:

- параметр $\alpha 1$ детрентного флуктуационного анализа (DFA);
- уровень омега-потенциала;
- тип variability сердечного ритма, включающий в себя вагосимпатический индекс как отношение LF/HF, отношение амплитуд компонентов VLF и HF спектрального вида variability сердечного ритма, общую мощность спектра variability сердечного ритма, TINN.

В качестве примера рассмотрим лингвистическую переменную «тип variability сердечного ритма». Для определения элементов терм-множеств и их функций принадлежности проведена нечеткая кластеризация на основе собранных статистических данных, которые представляют собой матрицу данных размерности 126×4 . В этом случае матрица данных соответствует 126 обследованиям, для каждого из которых выполнены измерения по четырем показателям. В результате применения нечеткой кластеризации для рассматриваемой переменной было выделено 6 кластеров, образовавшие терм-множество $T = \{\text{«оптимальный»}, \text{«бародисфункциональный»}, \text{«ваготонический»}, \text{«депрессивный вариант 1»}, \text{«депрессивный вариант 2»}, \text{«гиперадаптивный»}\}$. В качестве примера часть полученных значений функции принадлежности для термина «оптимальный» приведена в табл. 1.

Полученные термы и их функции принадлежности были положены в основу алгоритма нечеткого управления процессом введения медикаментов больным, что позволило значительно улучшить показатели терапии и повысить качество обслуживания пациентов.

Таблица 1 – Значения функции принадлежности для термина «оптимальный» лингвистической переменной «тип variability сердечного ритма»

LF/HF	VLF/HF	TP	TINN	μ
0,86	0,092	976	81,4	0,00538519
0,875	0,0925	990	82,25	0,019585547
0,89	0,093	1004	83,1	0,064793687
0,905	0,0935	1018	83,95	0,150721829
0,92	0,094	1032	84,8	0,281761316
0,935	0,0945	1046	85,65	0,450265078
0,95	0,095	1060	86,5	0,632541337
0,965	0,0955	1074	87,35	0,795718225
0,98	0,096	1088	88,2	0,913903471
0,995	0,0965	1102	89,05	0,979234486

Выводы

Представленный способ задания многомерных функций принадлежности термов лингвистических переменных на основе нечеткой кластеризации позволяет сохранить в нечеткой модели управления нелинейные взаимосвязи между переменными объекта, а также снизить влияние субъективности эксперта. Результаты численного исследования показали эффективность метода в задаче автоматизации процесса диагностирования и введения медикаментов при лечении пациентов с преэклампсией различной степени тяжести.

Литература

1. Штовба С.Д. Проектирование нечетких систем средствами MATLAB. / Штовба С.Д. – М.: Горячая линия – Телеком, 2007. – 288 с.
2. Ротштейн А.П. Интеллектуальные технологии идентификации: нечеткая логика, генетические алгоритмы, нейронные сети. / Ротштейн А.П. – Винница: УНІВЕРСУМ – Вінниця, 1999. – 320 с.
3. Алтуни А.Е. Модели и алгоритмы принятия решений в нечетких условиях: Монография. / Алтуни А.Е., Семухин М.В. – Тюмень: Издательство Тюменского государственного университета, 2000. – 352с.

4. Борисов В.В. Нечеткие модели и сети. / Борисов В.В., Круглов В.В., Федулов А.С. – М.: Горячая линия – Телеком, 2007. – 284 с.
5. И.А. Тарасова Нечеткое управление на основе переменных с многомерными функциями принадлежности в диагностике и лечении гипертензивных осложнений беременности // Радиоэлектронные и компьютерные системы. – 2012. – №4. – С. 169-173.
6. Рутковская Д. Нейронные сети, генетические алгоритмы и нечеткие системы: Пер. с польск. И.Д. Рудинского./ Рутковская Д., Пилиньский М., Рутковский Л. – М.: Горячая линия – Телеком, 2006. – 452 с.
7. Вятчинин Д.А. Нечеткие методы автоматической классификации: монография / Д.А. Вятчинин – Минск: Технопринт, 2004. – 219 с.
8. А.А. Барсемян Методы и модели анализа данных: OLAP и Data Mining / Барсемян А.А., Куприянов М.С., Степаненко В.В., Холод И.И. – СПб.: БХВ-Петербург, 2004. – 336с.

Сведения об авторах

Шушура Алексей Николаевич – к.т.н., доцент, доцент кафедры системного анализа и моделирования, Донецкий национальный технический университет, leshu@i.ua.

Тарасова Ирина Александровна – аспирант кафедры системного анализа и моделирования, Донецкий национальный технический университет, irina_tarasova@i.ua.

КОМП'ЮТЕРНІ СИСТЕМИ ТА КОМПОНЕНТИ

УДК 621.382.3

О. Д. АЗАРОВ, М. Ю. ТЕПЛИЦЬКИЙ, В. Є. ЯЦИК

Вінницький національний технічний університет, Вінниця

СПЕЦІАЛІЗОВАНІ ВІДБИВАЧІ СТРУМУ З ПАРАФАЗНИМИ ВХОДАМИ ДЛЯ ДВОТАКТНИХ ПІДСИЛЮВАЛЬНИХ СХЕМ

Анотація. У статті розглядаються методи покращення статичних і динамічних характеристик спеціалізованих двотактних відбивачів струму з парафазними входами, а також двотактних підсилювальних схем, до складу яких вони входять. Оцінюються параметри статичних і динамічних характеристик, зокрема, рівні вхідних і вихідних опорів, а також смуга пропускання для малосигнальних коефіцієнтів передачі по струму. Розглянуто приклади побудови двотактних підсилювальних схем із застосуванням системних спеціалізованих відбивачів струму. Здійснено порівняльний аналіз їх статичних і динамічних характеристик.

Ключові слова: двотактний відбивач струму, ДППС, вхідний опір, вихідний опір.

Аннотация. В статье рассматриваются методы улучшения статических и динамических характеристик специализированных отражателей тока с парафазными входами, а также двухтактных усилителей, в состав которых они входят. Оцениваются параметры статических и динамических характеристик, в частности, уровни входных и выходных сопротивлений, а также полоса пропускания для малосигнальных коэффициентов передачи по току. Рассмотрены примеры построения двухтактных усилителей с применением системных специализированных отражателей тока. Совершен сравнительный анализ их статических и динамических характеристик.

Ключевые слова: двухтактный отражатель тока, ДУПТ, входное сопротивление, выходное сопротивление.

Abstract. This article describes methods of improving the static and dynamic characteristics of specialized current mirrors with paraphase inputs, and push-pull amplifier, in which they are used. Parameters of static and dynamic characteristics, in particular, levels of the input and output resistances, as well as bandwidth for small-signal current transfer ratio are estimated. Examples of the construction of push-pull amplifier system using specialized current mirrors are considered. Authors had committed a comparative analysis of their static and dynamic characteristics.

Key words: push-pull current mirror, push-pull amplifier DC, input resistance, output resistance.

Вступ

В останні два десятиліття набувають поширення двотактні підсилювальні схеми, зокрема, двотактні підсилювачі постійного струму (ДППС). Вони характеризуються високими статичними і динамічними характеристиками [1-4]. Водночас їх можна додатково покращити за умови застосування в сигнальних колах ДППС спеціалізованих двотактних відбивачів струму (СпВС). Відмінною особливістю останніх можуть слугувати низький вхідний і високий вихідний опори, а також достатньо широка смуга пропускання, яка практично наближається до частоти одиничного підсилення по струму біполярного транзистора.

Актуальність

Існують різноманітні схеми відбивачів струму (ВС). Найпростішими серед них є так звані параметричні [2, 5]. Проте їх недоліками є невисокий вихідний опір і, як наслідок, зменшення коефіцієнту передачі струму в діапазоні прикладеної напруги. Для покращення статичних характеристик ВС, зокрема, збільшення вихідного опору застосовується низка схемотехнічних прийомів [2, 6, 7]. Окремо розглядаються методи зменшення вхідного опору [6, 7]. Водночас при побудові двотактних підсилювальних схем у сигнальних колах застосовані відбивачі струму використовуються переважно у вигляді окремих компонент, не зв'язаних між собою системно. Це не дозволяє в повній мірі отримати переваги двотактного принципу підсилення і перетворення сигналів. Автори пропонують під час проектування двотактних підсилювальних пристроїв використовувати спеціалізовані двотактні ВС із покращеними статичними і динамічними характеристиками, побудованими у вигляді системних вузлів, елементи яких зв'язані між собою як єдине ціле. Проте матеріали, присвячені такому підходу, в науково-технічній літературі зустрічаються відносно рідко і до того ж мають розпорошений характер. У зв'язку з цим, тема статті, присвячена принципам побудови згаданих СпВС, є актуальною.

Мета досліджень

Покращення статичних і динамічних характеристик спеціалізованих двотактних ВС із парафазними входами, а також двотактних підсилювальних схем, до складу яких вони входять, шляхом структурно-схемотехнічної організації балансових СпВС у вигляді функціонального вузла, елементи якого зв'язані між собою на системному рівні.

Задачі

1. Проаналізувати запропоновані підходи щодо структурно-схемотехнічної організації спеціалізованих двотактних ВС із парафазними входами, побудованих у вигляді системного функціонального вузла.

2. Оцінити параметри статичних і динамічних характеристик двотактних СпВС із парафазними входами, зокрема, рівні вхідних і вихідних опорів, а також смугу пропускання для малосигнальних коефіцієнтів передачі по струму.

3. Розглянути приклади побудови двотактних підсилювальних схем із застосуванням системного СпВС із парафазними входами. Порівняти параметри статичних і динамічних характеристик, зокрема, двотактних підсилювачів струму з вибіркоким зворотним зв'язком на базі згаданих СпВС і відомих ВС.

Розв'язання задач

Спеціалізований двотактний відбивач струму з парафазними входами по суті являє собою двотактний підсилювальний пристрій із глибокими зворотними зв'язками. Причому, залежно від структурно-функціональної організації такі СпВС можна поділити на прості (двоступеневі) і на модифіковані (триступеневі). Безумовно такий поділ є умовним, але він дозволяє виділити деякі принципи відмінності в особливостях роботи цих пристроїв.

Структурно-функціональну схему двоступеневого двотактного СпВС наведено на рисунку 1а. Вона містить першу вхідну ступінь, реалізовану на транзисторах T' і T'' , другу ступінь у вигляді двонаправленого відбивача струму (ДВС) та параметричні відбивачі струму BC' і BC'' . Балансовий вихід ДВС слугує виходом схеми та підключений до опору навантаження R_H . Задання робочих точок пристрою здійснюється за допомогою генератора робочого струму I_p . Для аналізу АЧХ і ФЧХ у схему уведено керований генератор вхідного струму КГС, який інвертує фазу $I_{вх}$. Двотактний ДВС може бути реалізований у двох варіантах. Перший – це застосування параметричних ВС, як показано на рисунку 1б. Другий – це застосування комплементарних схем Уілсона, як зображено на рисунку 1в.

Пристрій працює таким чином. Прирости вхідного струму $I_{вх}$ відповідно у фазі і протифазі подаються на бази T' і T'' . Якщо значення $I_{вх}$ збільшується, то T' привідкривається, а T'' - прикривається. При цьому $I'_{вих}$ збільшується, а $I''_{вих}$ - зменшується. За таких умов $I_{вих}$ починає витікати з виходу схеми в навантаження і напруга на R_H стає додатною. Якщо $I_{вх}$ зменшується, то вихідний струм $I_{вих}$ через навантаження втікає у вихідну шину і напруга на R_H стає від'ємною.

Бази транзисторів T' і T'' слугують парафазними входами схеми, відповідно, Vx' і Vx'' . Малосигнальні вхідні опори СпВС можна оцінити за допомогою аналітичних виразів:

$$r'_{вх} = \frac{r_{вх} T'}{1 + k'_i}, \quad (1)$$

$$r''_{вх} = \frac{r_{вх} T''}{1 + k''_i}, \quad (2)$$

де $r_{вх} T'$ і $r_{вх} T''$ – малосигнальні опори транзисторів, відповідно, T' і T'' , включених за схемою із загальним емітером; k'_i і k''_i – коефіцієнти передачі по струму відповідно верхнього і нижнього трактів підсилення при розірваній петлі зворотного зв'язку: $k'_i = \beta' \cdot k_{двс} k'_{вс}$, $k''_i = \beta'' \cdot k_{двс} k''_{вс}$, де β' , і β'' – коефіцієнти передачі по струму T' і T'' , і $k'_{вс}$, і $k''_{вс}$ – коефіцієнти передачі по струму BC' і BC'' .

Значення коефіцієнта передачі струму k_i двотактного підсилювача залежить від типу ДВС [4]. Для схеми на рисунку 1б:

$$k_i \approx \frac{2\beta' \cdot \beta''}{\beta' + \beta''},$$

де β' і β'' – малосигнальні коефіцієнти підсилення струму, відповідно р – n – p і n – p – n транзисторів.

У випадку застосування у ДВС схем Уілсона (рис. 1б) маємо:

$$k_i \approx \frac{\beta' \cdot \beta''}{\beta' + \beta''}$$

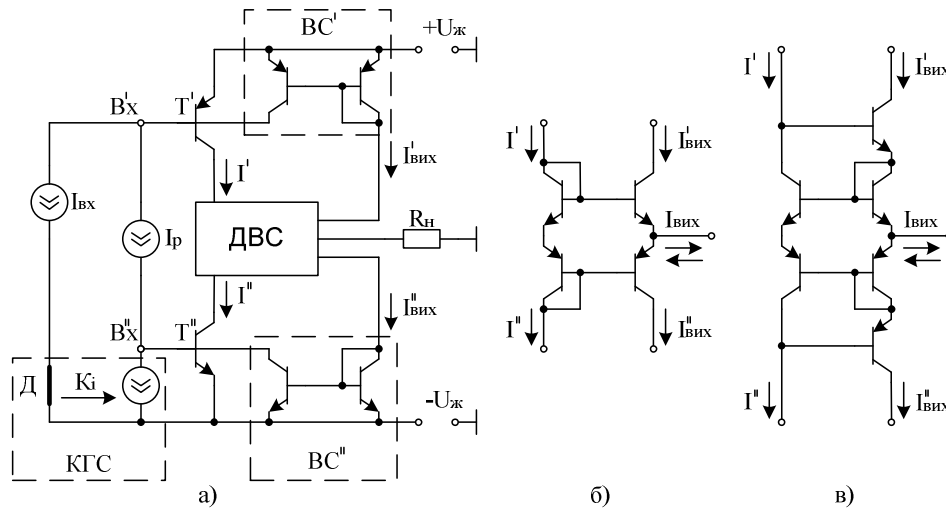


Рисунок 1 – Двотактний двоступеневий спеціалізований відбивач струму з парафазними входами: а) структурно-функціональна схема; двонаправлені ВС: б) параметричного типу; в) на основі схеми Уїлсона

Значення $r_{вх T'}$ і $r_{вх T''}$ визначають у вигляді [8]:

$$r_{вх T'} = r'_0 + r_e(1 + \beta''), \quad r_{вх T''} = r''_0 + r_e(1 + \beta'),$$

де $r_e = \frac{\Phi_T}{I_p}$ – малосигнальний опір емітерного переходу; Φ_T – термопотенціал; r_0 – опір бази.

Враховуючи вищенаведене для випадку ДВС, зображеного на рисунку 1б, маємо:

$$r'_{вх} \approx r''_{вх} \approx \frac{r_e(\beta' + \beta'')}{2\beta' \cdot \beta''},$$

на рисунку 1в:

$$r'_{вх} \approx r''_{вх} \approx \frac{r_e(\beta' + \beta'')}{\beta' \cdot \beta''}$$

За умови задання $I_{вх} = 1\text{mA}$ і застосування інтегральних транзисторів фірми Intersil серії HFA 3046/3096 [9] маємо для першої схеми $r'_{вх} \approx r''_{вх} \approx 1,7 \text{ Ом}$, для другої $r'_{вх} \approx r''_{вх} \approx 1,6 \text{ Ом}$.

Загальні коефіцієнти передачі по струму верхнього і нижнього трактів для визначеної смуги пропускання:

$$K'_{\Pi i} = \frac{K'_i}{1 + K'_i} \approx 1,0; \quad K''_{\Pi i} = \frac{K''_i}{1 + K''_i} \approx 1,0.$$

Слід відзначити, що ці схеми за принципом дії є балансними. Це означає, що як режими робочих точок, так і коефіцієнти передачі струмів зі входів на вихід у верхній і нижній частинах схеми є взаємно залежними. При цьому змінення струмів і напруг в одному фрагменті компенсується інверсними зміненнями значень величин в іншому. Для оцінювання вихідного опору $r_{вих}$ вказаних схем їх доцільно розбити на дві симетричні частини: верхню і нижню. Розглянемо це на прикладі двотактного СпВС, зображеного на рисунку 1а, із застосуванням ДВС, показаного на рисунку 1б. При цьому фрагмент нижньої частини схеми і його малосигнальну електричну схему заміщення наведено, відповідно, на рисунках 2а і б.

Для схем заміщення, наведених на цих рисунках, маємо такі позначення елементів: $r_{вх\ p-n-p}$ – малосигнальний вхідний опір p-n-p транзистора, включеного за гібридною схемою загальний емітер; $r_{k\ p-n-p}^*$ і $r_{k\ n-p-n}^*$ – опори колекторних переходів p-n-p і n-p-n транзисторів; β_{p-n-p} і β_{n-p-n} – коефіцієнти передачі струму відповідно p-n-p і n-p-n транзисторів, що реалізується керованими генераторами струму; R_H – опір навантаження; $I_{зм}$ – струм зміщення.

Застосовуючи I і II закони Кірхгофа, знайдемо вихідний опір у вигляді $r_{вих} = \frac{U_{вих}}{I_{вих}}$, а саме:

$$r_{вих} = \frac{r_{k\ n-p-n}^* \cdot \beta_{n-p-n} \cdot r_{k\ p-n-p}^* \cdot \beta_{p-n-p} - R_H}{r_{k\ n-p-n}^* \cdot \beta_{n-p-n} - r_{k\ p-n-p}^* \cdot \beta_{p-n-p}} \quad (3)$$

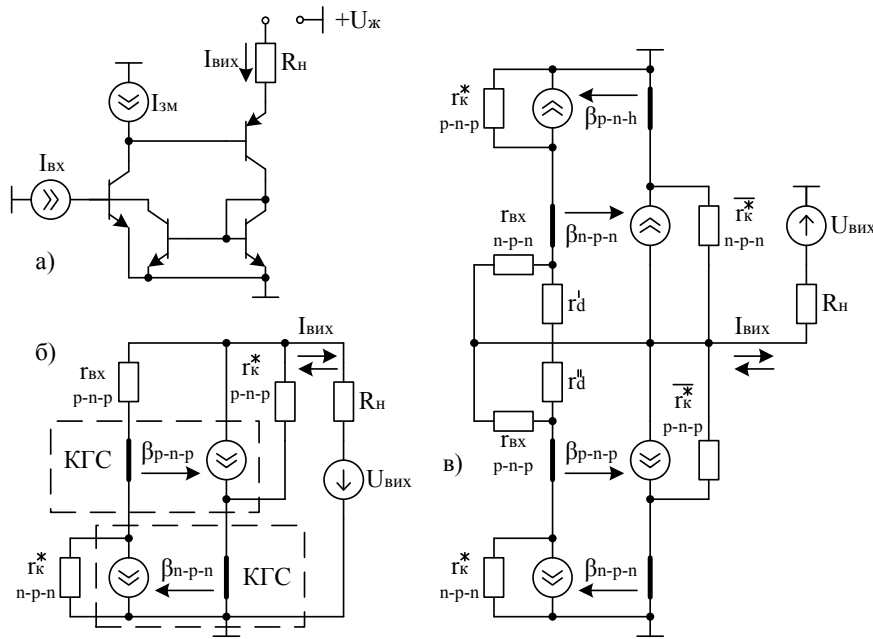


Рисунок 2 – Схеми заміщення двоступеневого СпВС: а) фрагмент принципової схеми нижньої частини; б) еквівалентна електрична схема нижньої частини; в) загальна еквівалентна схема

У випадку використання вищезазначених інтегральних транзисторів [9], $I_p = 1\text{мА}$, $U_{ж} = 15\text{В}$, маємо $\Gamma_{к\text{-}п\text{-}п}^* = 72\text{кОм}$, $\Gamma_{к\text{-}р\text{-}п}^* = 39\text{кОм}$, $\beta_{п\text{-}р\text{-}п} = 100$, $\beta_{р\text{-}п\text{-}р} = 64$, $r_{вх\text{-}п\text{-}р\text{-}п} = 3\text{кОм}$, $r_{вх\text{-}р\text{-}п\text{-}р} = 1,8\text{кОм}$, опори р-п переходів (транзисторів у діодному вмиканні) $r'_d = r''_d \approx 30\text{Ом}$. Підставляючи ці дані у формулу (3), для верхнього фрагменту отримаємо:

$$\Gamma'_{вих} = 7\text{МОм},$$

для нижнього:

$$\Gamma''_{вих} = 4,2\text{МОм}.$$

Замінивши в (3) добутки $\Gamma_{к\text{-}п\text{-}п}^* \cdot \beta_{п\text{-}р\text{-}п}$ і $\Gamma_{к\text{-}р\text{-}п}^* \cdot \beta_{р\text{-}п\text{-}р}$ на опори колекторних переходів транзисторів, включених за схемою із загальною базою, відповідно, $\Gamma_{к\text{-}п\text{-}п}$ і $\Gamma_{к\text{-}р\text{-}п}$, отримаємо вихідний опір у вигляді:

$$\Gamma_{вих} = \frac{\Gamma_{к\text{-}п\text{-}п} \cdot \Gamma_{к\text{-}р\text{-}п}}{\Gamma_{к\text{-}п\text{-}п} - \Gamma_{к\text{-}р\text{-}п}} - R_n. \quad (4)$$

При цьому слід вказати, що під час визначення за допомогою (4) вихідних опорів $\Gamma'_{вих}$ і $\Gamma''_{вих}$ значення $\Gamma_{к\text{-}п\text{-}п}$ і $\Gamma_{к\text{-}р\text{-}п}$ будуть для першого і другого випадків дещо відрізнятися. Це залежить від місця розташування п-р-п і р-п-р транзисторів у верхньому і нижньому фрагментах.

Неважко помітити, що опори $\Gamma_{вих}$, знайдені за допомогою (3) або (4) можуть бути як додатними, так і від'ємними. Це залежить від співвідношення $\Gamma_{к\text{-}р\text{-}п}^*$ і $\Gamma_{к\text{-}п\text{-}п}^*$, а також $\beta_{р\text{-}п\text{-}р}$ і $\beta_{п\text{-}р\text{-}п}$. У випадку, якщо $\Gamma_{к\text{-}п\text{-}п}^* \cdot \beta_{п\text{-}р\text{-}п} = \Gamma_{к\text{-}р\text{-}п}^* \cdot \beta_{р\text{-}п\text{-}р}$, то вихідний опір гіпотетично наближається до нескінченності.

Для загальної схеми заміщення, наведеної на рисунку 2в, вихідний опір також може бути визначений за допомогою узагальненого закону Кірхгофа, проте це призведе до досить громіздких виразів. Замість цього доцільно використовуючи принцип суперпозиції для малосигнальних схем, перейти від виразу (4) до:

$$\Gamma_{вих} = \frac{\Gamma'_{вих} \cdot \Gamma''_{вих}}{\Gamma'_{вих} - \Gamma''_{вих}} - R_n. \quad (5)$$

Тут $\Gamma'_{вих}$ і $\Gamma''_{вих}$ – вихідні опори, відповідно для верхнього і нижнього фрагментів. Підставляючи $\Gamma'_{вих}$ і $\Gamma''_{вих}$ у (5), маємо $\Gamma_{вих} \approx 10,5\text{МОм}$. Таким чином загальний вихідний опір у балансних СпВС може бути більшим за опори фрагментів. Фізичний сенс цього явища можна трактувати так, що паралельне з'єднання «додатного» і «від'ємного» резисторів дасть підсумковий резистор, модуль значення опору якого, буде більше опорів окремих резисторів.

Структурно-функціональну схему модифікованого (триступеневого) двотактного СпВС зображено на рисунку 3а. Вона містить вхідні транзистори T' і T'' , модифікований двонаправлений відбивач струму (МДВС), коригуючі конденсатори C'_k і C''_k і комбіновані відбивачі струму $КБВС'$ і $КБВС''$. Ці вузли призначено як для організації локальних і загальних зворотних зв'язків у верхньому і нижньому трактах підсилення, так і для забезпечення двотактного режиму роботи пристрою по виходу.

Схему модифікованого двонаправленого відбивача струму зображено на рисунку 3б. Відмінність полягає в тому, що в модифікованому варіанті замість діодів-давачів використовують п-р-п і р-п-р тран-

зистори, включені як каскади із загальною базою. Це дає можливість отримати по виходах не одну, а дві пари струмів, відповідно, $I'_{\text{двс}}$, $I''_{\text{двс}}$, а також їх інверсії $\bar{I}'_{\text{двс}}$ і $\bar{I}''_{\text{двс}}$.

Такий підхід дозволяє використовуючи відносно прості схеми КБВС' і КБВС'', покращити в цілому як статичні, так і динамічні модифікованого двотактного СпВС. Вхідні опори такого СпВС у загальному випадку можна оцінити за допомогою виразів (1) і (2). Для вибраної елементної бази при заданому $I_{\text{вх}} = 1\text{mA}$ маємо: $r'_{\text{вх}} \approx r''_{\text{вх}} \approx 0,55\text{ Ом}$. Вихідний опір схеми можна оцінити за допомогою комп'ютерного моделювання [10], він дорівнює $r_{\text{вих}} \approx 34\text{ МОм}$.

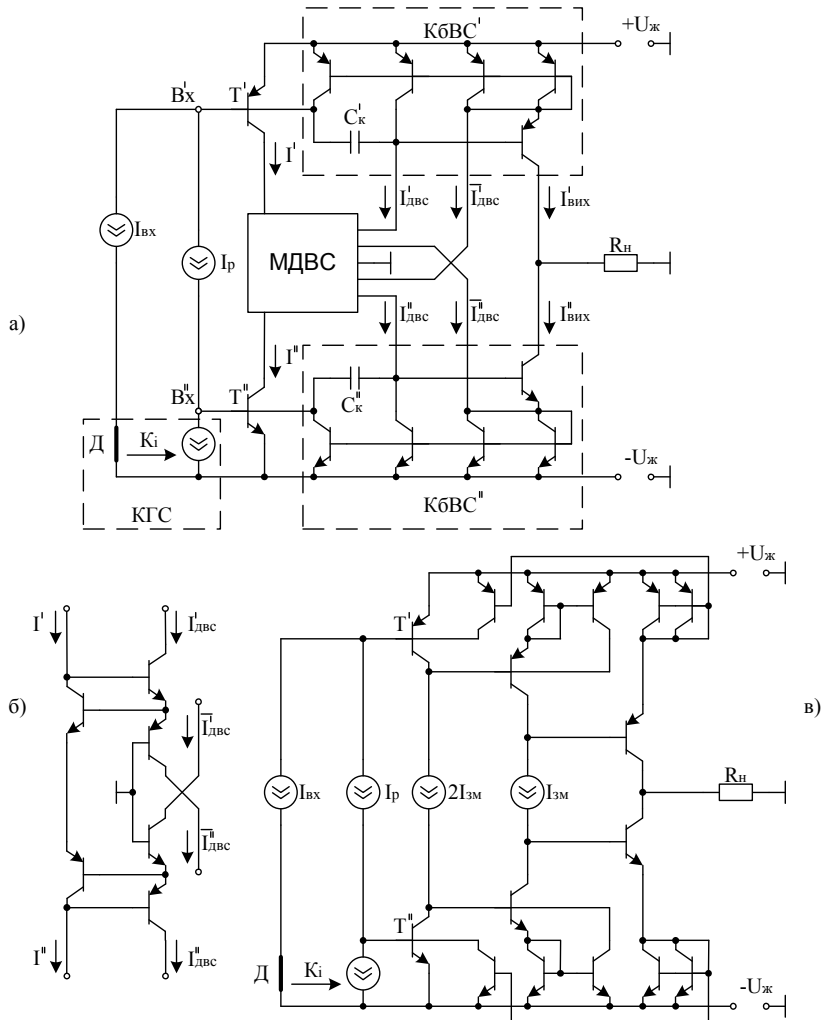


Рисунок 3 – Триступеневі двотактні СпВС: а) структурно-функціональна схема СпВС із модифікованим ДВС; б) принципова схема модифікованого ДВС; в) ДВС на «трійках» каскадів «загальний емітер»

Динамічні характеристики двотактних СпВС із парафазними входами доцільно проаналізувати в рамках комп'ютерного схемотехнічного аналізу, зокрема, за допомогою інтегрованого пакету прикладних програм MicroCAP 9. На рисунку 4 наведено АЧХ і ФЧХ цих пристроїв при $R_{\text{н}} = 1\text{ кОм}$. Тут криві 1 і 2 характеризують двоступеневий двотактний СпВС, зокрема, у випадку застосування схем ДВС, наве-

дених, відповідно на рисунках 1б і в, крива 3 – модифікований триступеневий двотактний СпВС (рисунок 3а), а крива 4 – СпВС на «трійках» каскадів із загальним емітером (рисунок 3в).

Аналіз отриманих результатів показує, що модифікований СпВС має найширшу смугу пропускання на рівні (-3db) і найменший фазовий зсув. Для простого СпВС перевагу має варіант, в якому використано ДВС на базі схеми Уілсона.

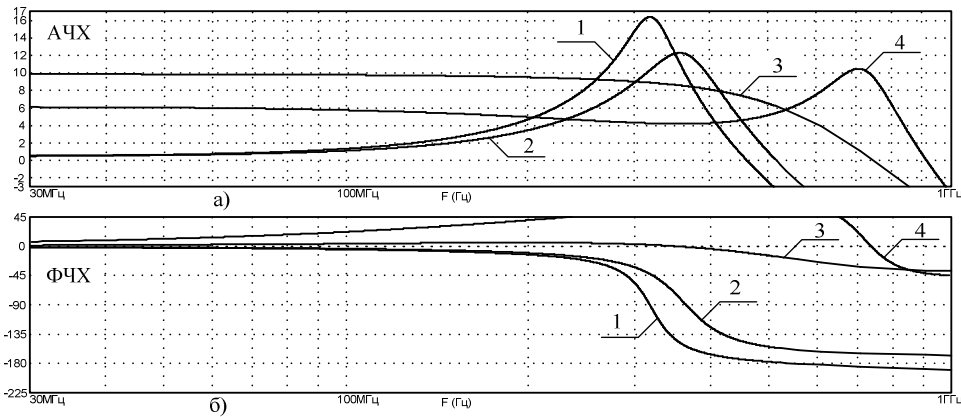


Рисунок 4 – Динамічні характеристики двотактних СпВС: а) амплітудно-частотна; б) фазочастотна

Розглянемо приклади застосування СпВС у двотактному підсилювачі струму з вибірковою зворотним зв'язком. Узагальнену структурно-функціональну схему цього пристрою [11] наведено на рисунку 5а. Вона містить: двотактний вхідний каскад (Вх. к.), побудований, на комплементарній парі n-p-n і p-p-r транзисторів, включених по схемі із загальною базою; двонаправлений відбивач струму (ДВС) з параметром $m=2$; параметричні відбивачі струму (BC' і BC''); генератори струму зміщення I_{3M} , а також каскади проміжного підсилення на транзисторах T' і T'' . Бази транзисторів вхідного каскаду підключено до генераторів напруг зміщення $+U_{3M}$ і $-U_{3M}$, які задають його робочу точку. Емітери T' і T'' проміжних каскадів під'єднано до парафазних входів двотактного СпВС.

Динамічні, а також статичні характеристики цього підсилювача струму залежать від типу СпВС, що застосовується. Критеріями вибору СпВС можуть слугувати значення малосигнального коефіцієнта передачі струму $K_i = \frac{\Delta I_{\text{вих}}}{\Delta I_{\text{вх}}}$, ширина смуги пропускання на рівні одиничного коефіцієнта підсилення

$f_c(0 \text{ db})$, а також рівень нелінійності статичної передатної характеристики. Останній параметр можна оцінити через похибку лінійності у вигляді:

$$\Delta I_{\text{л}} = I'_{\text{вих max}} - I''_{\text{вих max}}$$

де $I'_{\text{вих max}}$ і $I''_{\text{вих max}}$ – значення вихідних струмів у протилежних квадрантах залежності $I_{\text{вих}} = f(I_{\text{вх}})$, що протікають через вихідну шину підсилювача при однакових максимальних вхідних струмах протилежного напрямку, відповідно $+I'_{\text{вх max}}$ і $-I''_{\text{вх max}}$.

Абсолютна $\Delta I_{\text{л}}$ або відносна похибка лінійності $\delta I_{\text{л}} = \frac{\Delta I_{\text{л}}}{I'_{\text{вих max}}}$, де $I'_{\text{вих max}}$ – відповідає мен-

шому значенню величини $I'_{\text{вих max}}$ або $I''_{\text{вих max}}$, повинні визначитися по краях лінійного діапазону передатної характеристики. Слід відзначити, що такий підхід до оцінювання цих похибок відповідає найгіршому випадку, коли підсилювач включено без зворотного зв'язку.

На рисунку 5б наведено АЧХ ДППС за умови застосування згаданих інтегральних біполярних транзисторів, і заданні струму робочої точки $I_p = 1 \text{ MA}$ при $R_n = 1 \text{ kOM}$. Тут криві на графіках відповіда-

ють випадкам: 1 – коли замість СпВС використано відбивачі струму Уїлсона; 2, 3 – використано двоступеневі двотактні СпВС з різними типами ДВС; 4 – триступеневий СпВС (рис. 3а); 5 – СпВС на «трійках» транзисторів рисунк 3в.

Для оцінювання $\Delta I_{\text{л}}$ і $\delta I_{\text{л}}$ будемо задавати на вході ДППС такі $I'_{\text{вх max}}$ і $I''_{\text{вх max}}$, при яких $I'_{\text{вх}}$ або $I''_{\text{вх}}$ приблизно дорівнює 1 мА.

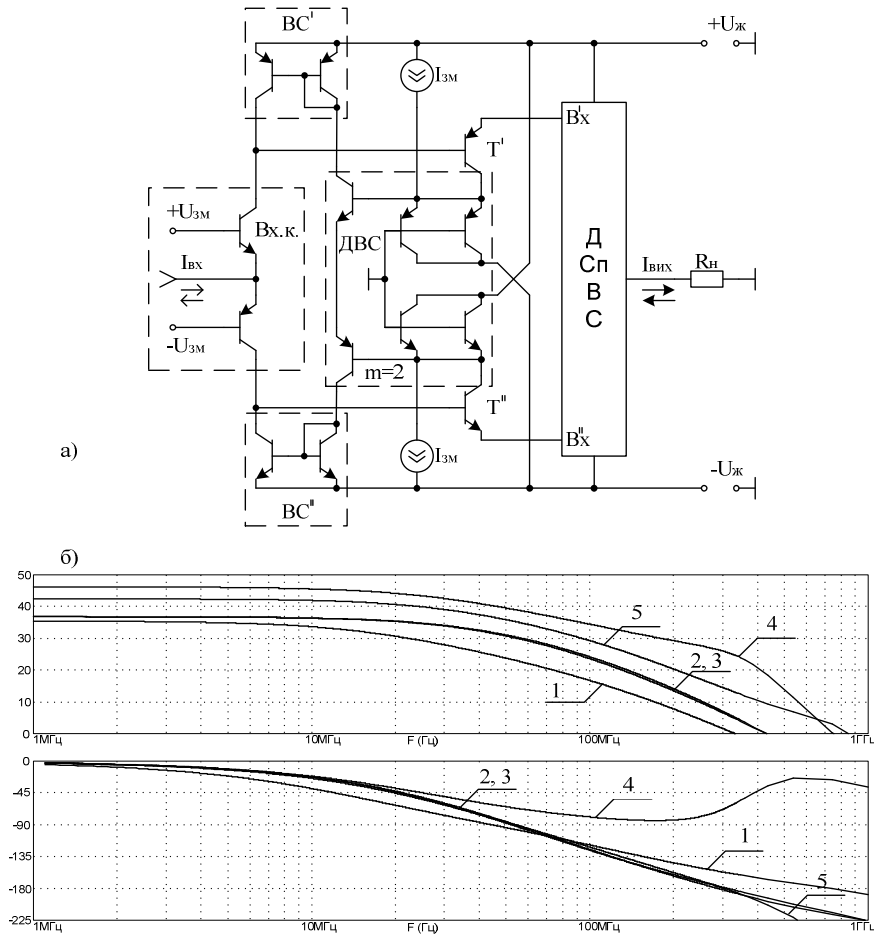


Рисунок 5 – Двотактний підсилювач струму з організацією вихідного вузла на базі ДСпВС: а) структурно-функціональна схема; б) АЧХ і ФЧХ

Параметри ДППС на базі СпВС, отримані в ході досліджень, наведено в таблиці 1.

Таблиця 1 – параметри ДППС і СпВС

Параметри ДППС				Параметри СпВС		
Тип СпВС	K_i	f_c (0 db) МГц	$\delta I_{\text{л}}$ (%)	$r_{\text{вих}}$ (МОм)	$r'_{\text{вх}}$ (Ом)	$r''_{\text{вх}}$ (Ом)
ВС Уїлсона	60	330	5,1	0,93	59	58
Двоступеневий СпВС	I	70	430	10,3	1,67	1,73
	II	70	430	11,8	1,49	1,78
Триступеневий СпВС	I	130	750	34,2	0,49	0,61
	II	200	870	1,8	0,3	0,52

Тут для двоступеневих СпВС варіанти I і II характеризують тип використаного ДВС (рисунки 1б, в відповідно). Для триступеневих: варіант I – відповідає модифікованому ДВС, варіант II – схемі на «трийках» транзисторів.

Аналіз отриманих результатів показує, що домінуючим параметром, що впливає на динамічні і статичні характеристики підсилювача є вхідний опір СпВС. Так у випадку застосування ВС Уілсона, ДППС має найменший K_i і найвужчу смугу пропускання. Водночас похибка лінійності при цьому є максимальною. Найкращі характеристики має ДППС, в якому застосовується триступеневий СпВС із модифікованим ДВС.

Висновки

1. Аналіз запропонованих підходів структурно-схемотехнічної організації СпВС із парафазними входами у вигляді системного функціонального вузла дозволив отримати рішення, що дозволяють досягти кращих статичних і динамічних характеристик.

2. Отримано оцінки статичних і динамічних характеристик двотактних СпВС із парафазними входами. Доведено, що найширшу смугу пропускання, найбільший малосигнальний коефіцієнт передачі і найвищий вихідний опір має триступеневий СпВС із модифікованим ДВС.

3. Розглянуті приклади побудови двотактних підсилювачів постійного струму з вибіркоким зворотним зв'язком дозволили визначити структуру з найкращими статичними і динамічними характеристиками, а саме ДППС на базі СпВС із модифікованим ДВС.

Список літератури

1. Grebene A. Bipolar and MOS Analog Integrated Circuit Design / Alan B. Grebene – , 2003. – 879 p.
2. Титце У. Полупроводниковая схемотехника / У. Титце, К. Шенк [пер. с нем. ДМК Пресс]. – М. : – ДМК Пресс, 2008 Т. 1: 2008. – 832 с. – ISBN 5-94074-148-7.
3. Yodprasit U. High-precision CMOS current conveyor / U. Yodprasit // – Electron. Lett. – 2000. – vol.36, №27(23). – 609-610 p. – ISSN 0013-5194.
4. Азаров О.Д. Двотактні підсилювачі постійного струму для багаторозрядних перетворювачів форми інформації, що самокалібруються : монографія / О. Д. Азаров, В. А. Гарнага. – Вінниця : ВНТУ, 2011. – 156 с. – ISBN 978-966-641-435-2
5. Гребен А. Б. Проектирование аналоговых интегральных схем / А. Б. Гребен; [пер. с англ.]. – М. : «Энергия», 1976. – 256 с., ил. – ISBN 978-5-458-38911-2.
6. Азаров О. Д. Методи покращення статичних характеристик відбивачів струму. / О. Д. Азаров, В. Є. Яцик. // Інформаційні технології та комп'ютерна інженерія. – 2012. – №1 (23). – С. 31-39. – ISSN 1999-9941.
7. Азаров О. Д. Відбивачі струму для аналогових пристроїв із покращеними статичними і динамічними характеристиками. / О. Д. Азаров, В. А. Гарнага, В. Є. Яцик. // Інформаційні технології та комп'ютерна інженерія. – 2012. – №2 (24). – С. 48-55. – ISSN 1999-9941.
8. Степаненко И. П. Основы микроэлектроники : учебное пособие для вузов. [2-е изд. перераб. и доп.] / И. П. Степаненко – М. : Лаборатория Базовых Знаний, 2003. – 488 с., ил. – ISBN 5-932208-045-0.
9. Технічний опис HFA3046/3096/3127/3128 Transistor Array SPICE Models: Application Note, Intersil, 1994. – 5 p.
10. Амелина М. А., Програма схемотехнічного моделювання Micro-Cap. Версии 9, 10. / М. А. Амелина, С. А. Амелин – Смоленск, Смоленский филиал НИУ МЭИ, 2012. – 617 с., ил.
11. Азаров О. Д. Двотактні підсилювачі постійного струму на базі двонаправлений відбивачів струму / О. Д. Азаров, М. Ю. Теплицький, В. А. Гарнага // Проблеми інформатизації та управління. – Київ, 2011. – №2 (34). – с. 15-22.

Відомості про авторів

Азаров Олексій Дмитрійович – д.т.н., проф., завідувач кафедри обчислювальної техніки, Вінницький національний технічний університет, azarov2@vntu.edu.ua, (0432)598-437, Хмельницьке шосе, 95, м. Вінниця, Україна, 21021.

Теплицький Михайло Юхимович – аспірант кафедри ОТ, Вінницький національний технічний університет, michaeltepl@gmail.com, (0432)598-437, Хмельницьке шосе, 95, м. Вінниця, Україна, 21021.

Яцик Володимир Євгенійович – магістрант кафедри ОТ, Вінницький національний технічний університет, yatsyk.ve@gmail.com, (0432)598-437, Хмельницьке шосе, 95, м. Вінниця, Україна, 21021.

УДК 681.586

БАЗІЛО К.В., СОТУЛА Ж.В., ТКАЧЕНКО О.С.

Черкаський державний технологічний університет, м. Черкаси

ДОСЛІДЖЕННЯ ВПЛИВУ ДОДАТКОВИХ ЕЛЕМЕНТІВ НА АМПЛІТУДНО-ЧАСТОТНІ ХАРАКТЕРИСТИКИ П'ЄЗОЕЛЕКТРИЧНИХ ПЕРЕТВОРЮВАЧІВ

Анотація. Досліджено вплив додаткових індуктивності та ємності на амплітудно-частотні характеристики п'єзоелектричних перетворювачів. Визначено, що при різних варіантах підключення ємності разом з індуктивністю на вхід п'єзоелектричного перетворювача його амплітудно-частотна характеристика змінюється. Одержана експериментальна залежність амплітудно-частотної характеристики від зміни додаткової ємності.

Ключові слова: додаткові елементи, амплітудно-частотна характеристика, п'єзоелектричний перетворювач.

Аннотация. Исследовано влияние дополнительных индуктивности и емкости на амплитудно-частотные характеристики пьезоэлектрических преобразователей. Определено, что при различных вариантах подключения емкости вместе с индуктивностью на вход пьезоэлектрического преобразователя его амплитудно-частотная характеристика изменяется. Получена экспериментальная зависимость амплитудно-частотной характеристики от изменения дополнительной емкости.

Ключевые слова: дополнительные элементы, амплитудно-частотная характеристика, пьезоэлектрический преобразователь.

Abstract. The research of the influence of additional inductance and capacitance on the amplitude-frequency responses of piezoelectric transducers. It was determined that in different ways to connect capacitance with the inductance of the input of the piezoelectric transducer amplitude-frequency response varies. The experimental dependence of the amplitude-frequency responses from changes of additional capacity was obtained.

Key words: additional elements, amplitude-frequency response, piezoelectric transducer.

Вступ

П'єзоелектричні перетворювачі і датчики широко застосовуються в комп'ютерних системах, гідроакустиці, електроакустиці, в ультразвуковій, медичній, вимірювальній техніці, в скануючих зондових наномікроскопах, п'єзодвигунах і в інших галузях науки і техніки [1-3].

Як відомо, перетворювач – це пристрій, який перетворює одну фізичну величину або енергію в іншу фізичну величину або енергію, наприклад, теплову енергію – в електричну, силу – в переміщення, тиск – в електричну напругу або струм, електричну напругу одного рівня – в електричну напругу іншого рівня і т.д. [3].

Актуальність

В роботі [4] описані методи синтезу (створення, проектування) п'єзоелектричних перетворювачів.

Ці методи дозволяють не тільки розширити можливості технології проектування, але і створювати перетворювачі з необхідними характеристиками.

Серед описаних технологій особливий інтерес представляє технологія додаткових елементів, так як в цьому випадку не потрібно будь-яка зміна в структурі п'єзоелектрика, а зміна характеристик перетворювача здійснюється за рахунок зовнішніх для п'єзоелемента ланцюгів [4].

Суть описаної технології полягає в тому, що до п'єзоелемента приєднують додаткові елементи, які змінюють характеристики перетворювача. До п'єзоелемента механічно приєднується другий п'єзоелемент, металева пластина або ультразвуковий концентратор, або електричними ланцюгами приєднується ємність, індуктивність, опір, ще один п'єзоелемент або частина п'єзоелемента.

Мета

Метою даної роботи є вивчення впливу індуктивності та ємності, одночасно підключених до п'єзоелектричного перетворювача, на його характеристики.

Задачі

1. Розробити схеми підключення п'єзоелектричних перетворювачів з додатковими індуктивностями й ємностями.
2. Дослідити та проаналізувати амплітудно-частотні характеристики розроблених п'єзоелектричних перетворювачів.

Розв'язання задач

Як відомо [1, 3], п'єзоелемент (рис. 1) являє собою електромеханічну коливальну систему, еквівалентною електричною схемою якого являється послідовно-паралельний коливальний контур.

Якщо на вхід п'єзоелемента можна підключити індуктивність $L_{доп}$ таким чином, щоб ця індуктивність і міжелектродна ємність C_n утворили послідовний коливальний контур, характеристики п'єзоелемента будуть змінені [6-8].

Експериментальну перевірку виконано в п'єзотрансформаторному режимі, для чого один з електродів розділено на три частини – два півкільця і диск на електроакустичному перетворювачі ЗП-19 виробництва ВАТ «Аврора» (Волгоград). Перетворювач складається із пластини сталі 40Х діаметром 32 і тов-

щиною 0,15 мм. До пластини приклеєний епоксидним компаундом п'єзоелемент Ø23 і товщиною 0,2 мм. Біморфний елемент закріплений в корпусі з ударостійкого полістиролу.

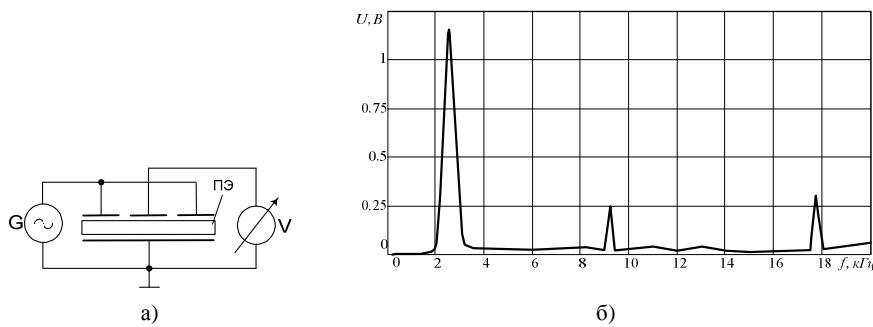


Рисунок 1 – Вимірювальна схема (а) та АЧХ п'єзоперетворювача ЗП-19 (б)

Міжелектродні ємності п'єзоелемента вимірювалися за допомогою цифрового мультиметра Voice kraft M890G. На вхід п'єзоелемента подавався синусоїдальний сигнал амплітудою 1В. Для вимірювань амплітудно-частотної характеристики (АЧХ) використовувався генератор ГЗ-109 і мілівольтметр ВЗ-41.

На рис. 2 показана схема підключення п'єзоелектричного перетворювача з додатковою індуктивністю.

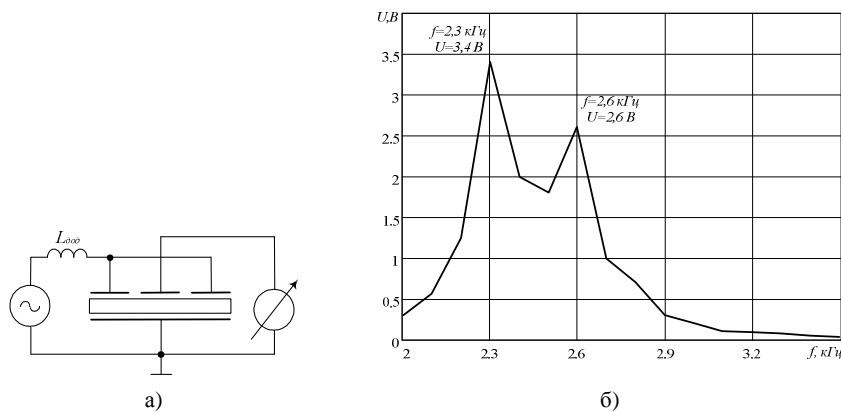


Рисунок 2 – Вимірювальна схема (а) та АЧХ п'єзоперетворювача ЗП-19 (б) при $L_{доп} = 325$ мГн

Розглянемо вхідний ланцюг перетворювача, що містить додаткову індуктивність – $L_{доп}$ і міжелектродну ємність – C_{ex} . Якщо вибрати $L_{доп}$ зі співвідношення:

$$L_{доп} = \frac{1}{4\pi^2 f_p^2 C_{ex}} \quad (1)$$

де f_p – резонансна частота п'єзоелемента, тоді резонансна частота f_k послідовного коливального контуру $L_{доп} C_{ex}$ буде дорівнює резонансній частоті п'єзоелемента, тобто $f_p = f_k$.

Як відомо [9], на резонансній частоті коефіцієнт передачі коливального контуру дорівнює добротності контура Q .

У цьому випадку напруга на вході п'єзоелемента (тобто ємності C_{ex}) буде рівна:

$$U_C = U_{\text{вх}} \cdot Q, \quad (2)$$

де Q – добротність контуру $L_{доп} C_{ex}$; $U_{ген}$ – напруга, що створюється генератором.

Отже, можна очікувати збільшення вихідної потужності (звукового тиску) перетворювача або коефіцієнта трансформації п'єзотрансформатора.

Дослідимо вплив додаткової ємності, яка під'єднується до п'єзоелектричного перетворювача одночасно з додатковою індуктивністю, в цьому разі можливо як мінімум три варіанта підключення ємності:

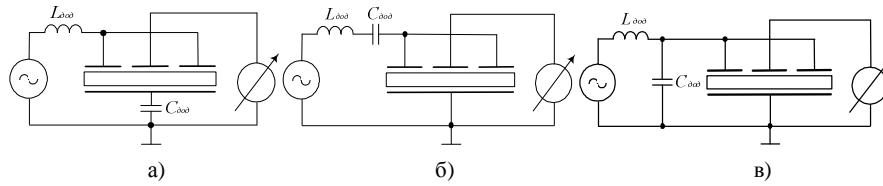
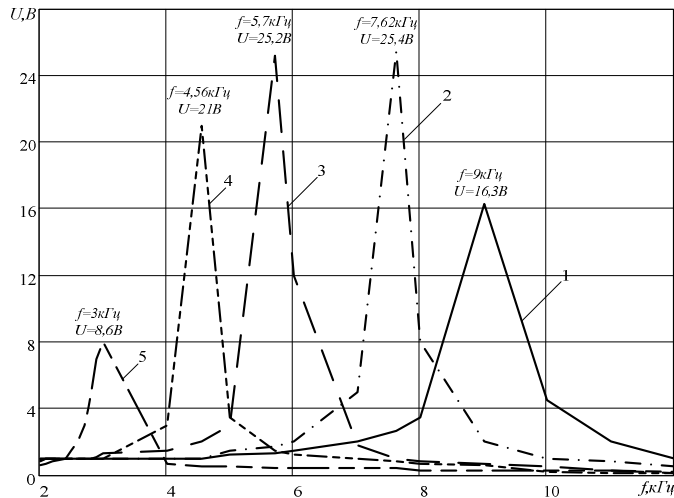


Рисунок 3 – Перетворювач ЗП-19 з додатковими індуктивністю і ємністю

В першому випадку (рис. 3, а), додаткова ємність C_{000} під'єднується між загальним проводом схеми і загальним проводом п'єзоелемента.

Як відомо, при послідовному з'єднанні конденсаторів на обкладинках окремих конденсаторів електричні заряди по величині рівні: $Q_1=Q_2=Q$.

У другому випадку (рис. 3, б) додаткова ємність C_{000} під'єднується послідовно L_{000} і вхідній ємності п'єзоелектричного перетворювача C_{ex} , та утворюється послідовний коливальний контур.

Рисунок 4 – АЧХ п'єзоелектричного перетворювача ЗП-19 показаного на рис. 3, а при $L_{000} = const$, 1 – $C = 0,5 \text{ нФ}$, 2 – $C = 1 \text{ нФ}$, 3 – $C = 2 \text{ нФ}$, 4 – $C = 4 \text{ нФ}$, 5 – $C = 18 \text{ нФ}$

Розглянемо випадок (рис. 4), коли додаткова ємність C_{000} підключається між загальним проводом схеми і загальним проводом п'єзоелектричного перетворювача, в наслідок чого сумарна ємність коливального контуру – $C_{заг} = Q/U = Q/(U_1 + U_2)$, тобто при послідовному з'єднанні ємностей величина, обернено-пропорційна загальній ємності, дорівнює сумі обернено-пропорційних величин окремих ємностей.

Напруга між обкладками окремих ємностей при їх послідовному з'єднанні буде рівна: $U_1 = Q/C_{000}$, $U_2 = Q/C_{ex}$, а загальна напруга рівна $U_{заг} = U_1 + U_2$, що пояснює зростання амплітуди на виході п'єзоелектричного перетворювача, а оскільки загальна ємність коливального контуру зменшилася, керуючись формулою (1), то відбувся зсув резонансу в область високих частот.

Створення додаткового коливального контуру на вході п'єзоелектричного перетворювача призвело до появи другого резонансного піку (рис. 5), а також до зростання амплітуди резонансу самого п'єзоелектричного перетворювача. Як і в першому випадку, ємності C_{000} та C_{ex} з'єднані послідовно, що пояснює зсув обох резонансних піків при зменшенні додаткової ємності C_{000} в область високих частот, керуючись формулою (1).

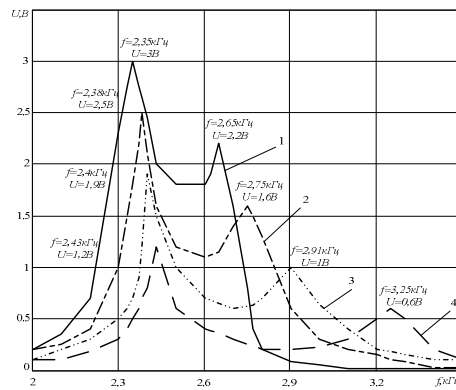


Рисунок 5 – АЧХ п'єзоелектричного перетворювача ЗП-19 показаного на рис. 3, б при: $L_{доод} = const$,
 1 – $C = 60$ нФ, 2 – $C = 33$ нФ, 3 – $C = 22$ нФ, 4 – $C = 14$ нФ

Нарешті, випадок коли ємність приєднана паралельно входу п'єзоелектричного перетворювача (рис. 3, в), додаткова ємність $C_{доод}$ включається паралельно вхідній ємності п'єзоелектричного перетворювача C_{ex} та утворюється додатковий коливальний контур $L_{доод}(C_{доод} + C_{ex})$.

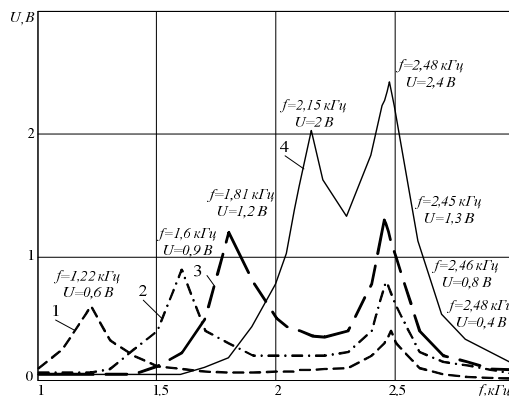


Рисунок 6 – АЧХ п'єзоелектричного перетворювача ЗП-19 показаного на рис. 3, в при: $L_{доод} = const$,
 1 – $C = 33$ нФ, 2 – $C = 14$ нФ, 3 – $C = 7$ нФ, 4 – $C = 1$ нФ

Створення додаткового коливального контуру $L_{доод}(C_{доод} + C_{ex})$ на вході п'єзоелектричного перетворювача призвело до появи другого резонансного піку (рис. 6). При паралельному з'єднанні конденсатора та ємності п'єзоелектричного перетворювача напруги, що підводяться до них, однакові – $U_1 = U_2 = U$. Заряди, які виникають при цьому, дорівнюють $Q_1 = C_{доод}U$, $Q_2 = C_{ex}U$, а загальний заряд, отриманий від джерела $Q_{заг} = Q_1 + Q_2$. Загальна ємність $C_{заг}$ коливального контуру $L_{доод}(C_{доод} + C_{ex})$ дорівнює сумі їхніх ємностей $C_{заг} = Q/U = (Q_1 + Q_2)/U = C_1 + C_2$, що і пояснює зменшення амплітуди на виході п'єзоелектричного перетворювача, а оскільки загальна ємність коливального контуру збільшилась, то керуючись формулою (1), відбувся зсув резонансів в область низьких частот.

Подальші дослідження можуть бути спрямовані на комп'ютерне моделювання еквівалентних електричних схем заміщення п'єзоелектричних перетворювачів та їхніх характеристик.

За матеріалами дослідження отримано патент на корисну модель [10].

Висновки

1. При підключенні ємності між загальним проводом п'єзоелектричного перетворювача та загальним проводом схеми, досягнуто зростання вихідної напруги в 25 раз порівняно з п'єзоперетворювачем без додаткових елементів та в 7,5 рази порівняно з п'єзоперетворювачем з додатковою індуктивністю.

2. Експериментально отримано залежність амплітудно-частотної характеристики від зміни додаткової ємності для різних конструкцій п'єзоелектричного перетворювача, що може бути використано при проектуванні п'єзоелектричних перетворювачів.

Список літератури

1. Шарапов В.М. Пьезоэлектрические датчики / В.М. Шарапов, М.П. Мусиенко, Е.В. Шарапова. – М.: Техносфера, 2006. – 632 с.
2. Sharapov V. Piezoceramic sensors / V. Sharapov. – Springer, 2011. – 498 p.
3. Шарапов В.М. Пьезокерамические трансформаторы и датчики / В.М. Шарапов, И.Г. Минаев, Ж.В. Сотула, К.В. Базило, Л.Г. Куницкая. – Черкасы: Вертикаль, 2010. – 278 с.
4. Шарапов В.М. Технологии синтеза пьезокерамических датчиков / В.М. Шарапов // Вісник ЧДТУ. – 2010. – №3.
5. Физика и техника мощного ультразвука. Физические основы ультразвуковой технологии / Под ред. Л.Д. Розенберга. – М.: Наука, 1970. – 688 с.
6. Сотула Ж.В. Совершенствование пьезокерамических трансформаторов для датчиков механических величин: диссертация на соискание ученой степени кандидата технических наук / Ж.В. Сотула. Черкасы, 2009. – 174 с.
7. Пат. 56930 Україна, МПК H04R 17/00. Спосіб створення акустичних коливачів за допомогою п'єзоелемента. / Шарапов В.М. – № 201012462 ; заявл. 22.10.10 ; опубл. 25.01.11, Бюл. № 2.
8. Пат. 56932 Україна, МПК H04R 17/00. Електроакустичний перетворювач / Шарапов В.М., Сотула Ж.В., Чорноший І.В., Куницкая Л.Г., Базило К.В. – № 201012464 ; заявл. 22.10.10 ; опубл. 25.01.11, Бюл. № 2.
9. Котельников В.А. Основы радиотехники / В.А. Котельников. – М.: Гостехиздат, 1950. – 372 с.
10. Пат. 62844 Україна, МПК H01F 30/00. П'єзоелектричний трансформатор / Шарапов В.М., Ткаченко О.С., Базило К.В., Сотула Ж.В. – № 201107543 ; заявл. 15.06.11 ; опубл. 12.09.11, Бюл. № 17.

Відомості про авторів

Базіло Костянтин Вікторович – к.т.н., старший викладач кафедри комп'ютеризованих та інформаційних технологій у приладобудуванні, тел.: 0472 730260, e-mail: b_constantine@mail.ru, бул. Шевченка, 460, 18006, м. Черкаси, Україна.

Сотула Жанна Василівна – к.т.н., старший викладач кафедри комп'ютеризованих та інформаційних технологій у приладобудуванні, бул. Шевченка, 460, 18006, м. Черкаси, Україна, тел.: 093 5641505, e-mail: janna-sotula@yandex.ru.

Ткаченко Олександр Сергійович – аспірант кафедри комп'ютеризованих та інформаційних технологій у приладобудуванні, тел.: 093 3592016, e-mail: tkachenko.a.s@yandex.ru, бул. Шевченка, 460, 18006, м. Черкаси, Україна.

УДК 004.222.2 + 004.312.2

О.Я. АНОПРІЄНКО, С.В. ІВАНИЦЯ, С.В. КУЛБАБА

Донецький національний технічний університет, Донецьк

ПРИНЦИП РОБОТИ, СТРУКТУРА І МОДЕЛЮВАННЯ БЛОКУ ПЕРЕТВОРЮВАЧА ФОРМАТІВ У СКЛАДІ ПОСТБІНАРНОГО СПІВПРОЦЕСОРА

Анотація. Показана необхідність створення постбінарного перетворювача форматів як ключового блоку постбінарного співпроцесора. Розглянуто принцип роботи та структуру цього блоку, а також особливості його апаратної реалізації. Наведено результати синтезу сигналів і моделювання розробленої мовою VHDL спрощеної моделі блоку.

Ключові слова: постбінарний формат даних, тетракод, VHDL, модель, співпроцесор.

Аннотация. Показана необходимость создания постбинарного преобразователя форматов как ключевого блока постбинарного сопроцессора. Рассмотрены принцип работы и структура этого блока, а также особенности его аппаратной реализации. Приведены результаты синтеза сигналов и моделирования разработанной на языке VHDL упрощенной модели блока постбинарного преобразователя форматов.

Ключевые слова: постбинарный формат данных, тетракод, VHDL, модель, сопроцессор.

Abstract. Necessity of creating postbinary formats converter of the key block postbinary coprocessor is shown. Principles of operation and structure of this unit, and features of hardware-based implementation are observed. Results of simulation of signals and modelling of simplified model of converter unit postbinary formats, which is developed in VHDL, are pointed presented.

Key word: postbinary data format, tetracode, VHDL, model, coprocessor.

Вступ

Вся множина сучасного апаратного забезпечення обчислювальних пристроїв оперує числами, представленими двома форматами: цілочисельними та дійсними. При цьому цілочисельна арифметика реалізована без похибок, на відміну від дійсної, яка не в змозі оперувати нескінченною множиною дійсних значень при наявності кінцевих числових наборів в дійсних форматах. Дійсні числа в комп'ютерних системах представляються в форматах чисел з плаваючою комою. Стандарт чисел з плаваючою комою IEEE 754-2008 [1] передбачає в ряді випадків наближене представлення початкового числа в плаваючих форматах. Похибка при цьому носить накопичувальний характер та залежить від подальших операцій з представленим значенням. Накопичення похибки в деяких випадках може призвести до повної втрати достовірності результату.

Така ситуація призводить до неприйнятних на сучасному етапі розвитку обчислювальної техніки наслідків. Для більшості прикладних задач сучасної точності чисел з плаваючою комою цілком достатньо, але, наприклад, для мультимасштабного моделювання макро- та мікросвіту, такі похибки досягають неприпустимо великих величин [2]. А в тих випадках, коли мова йде про безпеку і життя людини, достовірність результату стає одним з найважливіших критеріїв. Зокрема, директор Інституту математики та її застосувань в Міннеаполісі, США, Дуглас Н. Арнольд стверджує, що цілий ряд найбільших аварій з людськими жертвами та мільярдними збитками пов'язаний із сучасними технологіями комп'ютерних обчислень, в яких дані представлені згідно стандарту IEEE 754-2008 [3].

Перехід від бінарних до **постбінарних форматів** представлення дійсних чисел зумовлений двома важливими моментами:

1. Можливість постбінарного кодування даних (з використанням значень неоднозначності, яка є особливістю тетракоду, як одного з можливих наборів гіперкоду) [4].

2. Зміна структури самого формату, в якому крім стандартних полів знака, порядку та мантиси присутнє поле ідентифікатора (ID), що складається з **коду постбінарного формату CF** (визначає «ширину» формату — кількість розрядів, точність) і **модифікатора постбінарного формату MF** (визначає тип даних, «упакованих» у полі формату) [5, с. 198].

У світлі вищевикладеного представляється актуальною задача проектування вбудованої процесорної системи, що підтримує плаваючі формати чисел і забезпечує коректність їх обчислень, які не призводять до втрати достовірності. Рішення такої задачі можна розбити на два етапи: 1-й етап — модифікація форматів представлення дійсних чисел (корекція представлення даних); 2-й етап — введення додаткових методів контролю точності для стандартних алгоритмів обчислення (корекція обчислень).

Метою даної роботи є розробка блоку перетворювача постбінарних форматів (ППФ, рFC — postbinary formats converter), що виконує наступні функції:

- виділення формату завантаженого операнда за значенням поля коду формату CF;
- визначення типу даних в форматі за значенням поля модифікатора формату MF;
- реалізація постбінарного кодування (перетворення: двійковий код → тетракод);
- підготовка даних до завантаження в **постбінарний арифметико-логічний пристрій** (ПАЛП, або рALU — postbinary arithmetic logic unit) з розділенням розрядів операнду на поля мантиси, порядку та знака.

Важливість блоку ППФ у складі постбінарної процесорної системи (можливо, постбінарного співпроцесора) очевидна, оскільки на нього покладено функції розпізнавання типу значень, виділення та обробки полів формату і підготовки їх до завантаження у відповідні регістри ПАПП.

В рамках даної роботи розглянута апаратна реалізація блоку ППФ для підготовки операндів до завантаження в регістри 120-розрядного ПАПП. Наведено результати синтезу вихідних сигналів, функціональна схема ППФ та результати моделювання.

Принцип роботи та структура перетворювача постбінарних форматів

Розробку блоку ППФ можна позиціонувати як один з перших етапів розробки **постбінарного співпроцесора**, який представляє собою аналог математичного співпроцесора (або модуля операцій з плаваючою комою), здатного обробляти числа в постбінарних форматах. На рис. 1 наведено спрощену функціональну схему постбінарного співпроцесора (pFPU — postbinary floating point unit).

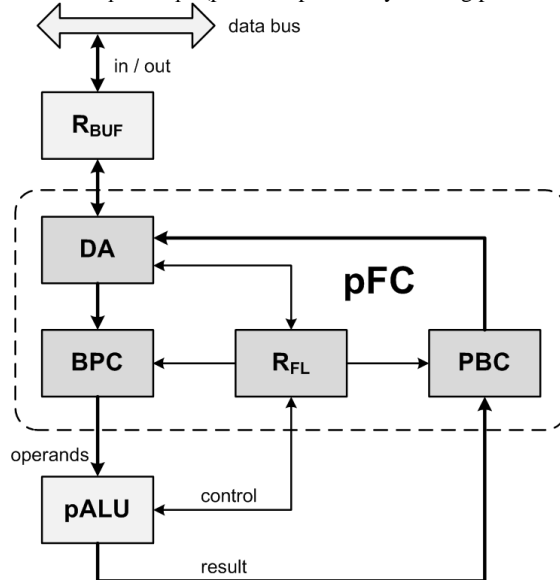


Рисунок 1 – Спрощена функціональна схема постбінарного співпроцесора pFPU з виділеною структурою постбінарного перетворювача форматів pFC (data bus – шина даних; in/out – вхідна/вихідна шина; operands – шина операндів; control – шина керування; result – шина результату)

Дані (бітове поле формату), що надійшли на вхід співпроцесора, а також результат обчислень, що надійшов з постбінарного АЛП (pALU) зберігаються в **буферному регістрі R_BUF** (buffer register). **Аналізатор даних DA** (data analyzer) розпізнає формат і тип розміщених у ньому даних, а також виконує виділення з вмісту R_BUF інформаційної та службової частин. Інформаційна частина (мантиса, порядок і знак) передається на **блок постбінарного перетворювача BPC** (binary-postbinary converter), а службова частина (код і модифікатор формату) обробляється та записується у **регістр прапорів R_FL** (flags register). У блоці BPC (при необхідності, на що вказують відповідні поля регістру R_FL) відбувається постбінарне кодування даних, тобто заміна бінарного коду тетракодом (фактично приведення бітів до тетритів [5, с. 65]), після чого постбінарні дані надходять до регістрів операндів pALU. Результат з постбінарного АЛП проходить через **блок бінарного перетворювача PBC** (postbinary-binary converter), в якому дані (при необхідності, на що вказують відповідні поля регістру R_FL) перетворюються з тетракоду в двійковий код (фактично представлення кожного тетриту парою біт), і знову надходять на блок DA. У блоці DA дані (що містять тільки інформаційну частину) збираються в первісний вигляд постбінарного формату шляхом формування і додавання службової частини згідно з необхідними полями в R_FL. Зібраний формат даних завантажується в регістр R_BUF.

Таким чином, робота зв'язки блоків BPC→DA←PBC дозволить отримати на виході pFPU результат в тому форматі, в якому було введено вихідні значення. Це означає, що з боку апаратури постбінарний співпроцесор pFPU ні чим не відрізнятиметься від бінарного співпроцесора FPU, а постбінарне представ-

лення числових значень буде використовуватися виключно для внутрішньої роботи блоків співпроцесора.

У таблиці істинності (табл. 1) блоку ППФ використовуються позначення постбінарних форматів **pb Ω , pb Ω / Ψ γ** , де:

- pb — скорочений запис постбінарного формату (від англ. **postbinary**);
- Ω — розрядність формату, що представляє значення $\Omega/32$ -арної точності (фактично, необхідна для зберігання формату розрядність);
- Ψ — розрядність значення, що входить до Ω -розрядного формату;
- γ — визначник групи Ψ -розрядного значення, що представляє множину {f, i, p, fp, ip}:
 - $\gamma = \emptyset$ – звичайне число Ω -розрядне двійкове число;
 - $\gamma = f$ – дріб виду a/b (Ψ -розрядні двійкові чисельник a та знаменник b звичайного дробу в одному полі даних та зі спільним знаковим розрядом, причому $\Psi = \Omega/2$);
 - $\gamma = i$ – інтервал виду $[x_1, x_2]$ (Ψ -розрядні двійкові значення меж інтервалу x_1 і x_2 в одному полі даних, причому $\Psi = \Omega/2$);
 - $\gamma = p$ – постбінарне число (Ψ -розрядний тетракод, виражений Ω -розрядним двійковим кодом, тому $\Psi = \Omega/2$);
 - $\gamma = fp$ – постбінарний дріб виду a'/b' (Ψ -розрядні постбінарні чисельник a' та знаменник b' звичайного дробу в одному полі двійкових даних та зі спільним знаковим постбінарним розрядом, тому $\Psi = \Omega/4$);
 - $\gamma = ip$ – постбінарний інтервал виду $[x'_1, x'_2]$ (Ψ -розрядні постбінарні значення меж інтервалу x_1 та x_2 в одному полі двійкових даних, тому $\Psi = \Omega/4$).

Таблиця 1 – Таблиця істинності сигналів-ознак формату

Постбінарний формат	Модифікатор формату MF	Код формату CF	Молодший байт формату X[7:0]	Вихідні сигнали*		
				CFm	MFm	
					Data_type	Data_form
pb32	0	0	xxxxxx00	00	0	00
pb32/16p	1		xxxxxx10		1	00
pb64	00	01	xxxx0001	01	0	00
pb64/32f	01		xxxx0101		0	01
pb64/32i	10		xxxx1001		0	10
pb64/32p	11		xxxx1101		1	00
pb128	000		011		00000011	10
pb128/64f	001	00001011		0	01	
pb128/64i	010	00010011		0	10	
pb128/64p	011	00011011		1	00	
pb128/32fp	100	00100011		1	01	
pb128/32ip	101	00101011		1	10	
pb256	000	0111		00000111	11	
pb256/128f	001		00010111	0		01
pb256/128i	010		00100111	0		10
pb256/128p	011		00110111	1		00
pb256/64fp	100		01000111	1		01
pb256/64ip	101		01010111	1		10

* – CFm[1:0] – фіксоване (2 біта) поле коду формату CF (вказує розрядність формату) для зберігання в регістрі прапорів R_{FL}; MFm[2:0] – фіксоване (3 біта) поле модифікатора формату MF (вказує тип даних) для зберігання в регістрі прапорів R_{FL}; Data_type[0] – кодовий базис (0 – двійкове кодування, 1 – постбінарне кодування); Data_form[1:0] – тип даних (00 – число, 01 – дріб, 10 – інтервал, 11 – резерв).

Поле CFm на відміну від поля коду формату CF має фіксовану ширину, що дозволяє економити реєстрову пам'ять і визначити розрядність вхідного числа, як $32 \cdot (CFm + 1)$ біта. Поле MFm представляє собою об'єднання сигналів $Data_type$ і $Data_form$. Така організація службової частини формату спрощує її передачу та зберігання, та водночас спрощує апаратну реалізацію (виключення зі схеми шифраторів і дешифраторів для об'єднання і розбиття цих сигналів в блоках ППФ).

В результаті синтезу було отримано наступні залежності:

$$CFm[1] = (\bar{x}_7 \bar{x}_6 \bar{x}_3 x_2 x_1 x_0) \vee (\bar{x}_7 \bar{x}_5 \bar{x}_3 x_1 x_0) \vee (\bar{x}_4 \bar{x}_2 x_1 x_0) \vee (\bar{x}_5 \bar{x}_2 x_1 x_0);$$

$$CFm[0] = (\bar{x}_7 \bar{x}_5 \bar{x}_3 x_2 x_0) \vee (\bar{x}_7 \bar{x}_6 \bar{x}_3 x_2 x_0) \vee (\bar{x}_1 x_0);$$

$$Data_type = (\bar{x}_7 \bar{x}_6 x_5 x_4 \bar{x}_3 x_2 x_1) \vee (\bar{x}_7 x_6 \bar{x}_5 \bar{x}_3 x_2 x_1) \vee (\bar{x}_5 x_4 x_3 \bar{x}_2 x_1) \vee (x_5 \bar{x}_4 \bar{x}_2 x_1) \vee (x_3 x_2 \bar{x}_1 x_0) \vee (x_1 \bar{x}_0);$$

$$Data_form[1] = (\bar{x}_7 \bar{x}_6 x_5 \bar{x}_4 \bar{x}_3 x_2 x_1 x_0) \vee (\bar{x}_7 x_6 \bar{x}_5 x_4 \bar{x}_3 x_1 x_0) \vee (\bar{x}_5 x_4 \bar{x}_3 \bar{x}_2 x_1 x_0) \vee (x_5 \bar{x}_4 x_3 \bar{x}_2 x_0) \vee (x_3 \bar{x}_2 \bar{x}_1 x_0);$$

$$Data_form[0] = (\bar{x}_7 x_6 \bar{x}_5 \bar{x}_4 \bar{x}_3 x_2 x_0) \vee (\bar{x}_7 \bar{x}_6 \bar{x}_5 x_4 \bar{x}_3 x_2 x_0) \vee (x_5 \bar{x}_4 \bar{x}_3 \bar{x}_2 x_1 x_0) \vee (\bar{x}_5 \bar{x}_4 x_3 \bar{x}_2 x_1 x_0) \vee (\bar{x}_3 x_2 \bar{x}_1 x_0).$$

Функціональність ППФ (зокрема, блок DA), крім перерахованих вище операцій, повинна виконувати ряд процедур, а саме:

1. Забезпечити послідовне завантаження операндів у реєстри $pALU$.
2. При роботі з «складовими» форматами (для дробів та інтервалів) забезпечити парціальне завантаження даних в $pALU$ у вибраному порядку (наприклад, спочатку обробляються знаменники операндів-дробів, потім їх чисельники, або спочатку ліві межі операндів-інтервалів, потім — праві).
3. Забезпечити виконання операції приведення типів. З двох операндів різних типів операції приведення підлягає операнд з меншою розрядністю. При неможливості приведення типів генерується відповідна помилка (помилка приведення типів постбінарних форматів).

Детальний опис перерахованих процедур, реалізація алгоритмів їхньої роботи і обробка виключень виходять за рамки теми даної роботи.

Апаратна реалізація блоку ППФ

Далі розглянуто апаратну реалізацію блоку ППФ для забезпечення узгодження постбінарних форматів зі 120-розрядним ПАЛП ($pALU120$), що оперує даними, представленими у тетракодах (рис. 2). Формати, що містять такі дані, надходять на вхід ПАЛП без змін, а дані, представлені двійковим кодом, приводяться до тетракоду в блоці бінарно-постбінарного кодування BPC . Для інших даних, що не підходять під вказані на рис. 2 формати, генерується сигнал помилки err :

$$err = (x_3 x_4 \bar{x}_2 x_1 x_0) \vee (x_6 x_5 x_2 x_1 x_0) \vee (x_7 x_2 x_1 x_0) \vee (x_3 x_2 x_1 x_0).$$

На відміну від рис. 1, на рис. 2 робота блоку BPC виглядає очевидніше і полягає в перетворенні 60-бітної двійкової шини в 120-розрядний тетракод. Це відбувається таким чином: кожен біт представляється парою двійкових розрядів ($0 \rightarrow 01$ та $1 \rightarrow 10$), тим самим стаючи вже значеннями тетритів 0 та 1.

Якщо вхідне число має «тетракодове» представлення, то воно передається в постбінарний АЛП без перетворень.

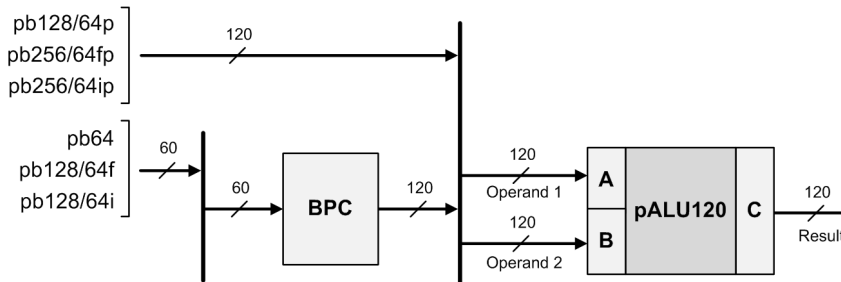


Рисунок 2 – Схема взаємодії постбінарного перетворювача (BPC) та 120-розрядного постбінарного АЛП ($pALU120$) із зазначенням допустимих постбінарних форматів

На рис. 3 представлена спрощена модель блоку ППФ, що реалізована мовою опису апаратури інтегральних схем $VHDL$ [6]. У спрощеній моделі виконана повна реалізація блоку BPC та часткова — блоку DA (визначення розрядності формату і типу даних в ньому, формування та занесення до R_{FL} сигналів-ознак, підготовка інформаційної частини формату до обробки в $pALU$).

На вхід моделі блоку ППФ по 256-розрядній шині даних подаються дані в передбачуваному постбінарному форматі. На виході — шини знаку, порядку та мантиси постбінарного числа, отриманого (у разі виправдання припущення) з ідентифікованого і обробленого формату. При цьому також формуються описані раніше сигнали-ознаки, у тому числі й сигнал помилки розпізнавання формату.

Модель блоку ППФ складається з тактованих функціональних модулів U1-U5:

- U1 — 256-розрядний буферний регістр RB (аналог блоку R_{BUF} на рис. 1) для зберігання поля формату вхідного операнду;
- U2 — модуль data_modifier, виконує визначення коду CF і модифікатора MF вхідного формату, і формування їхніх еквівалентів (CF_m, MF_m) на шинях bus_cfm, bus_data_form та bus_data_type, або помилку поля формату (якщо у вхідному регістрі дані не rb-форматів) на шині bus_DM_err.
- U3 — постбінарний перетворювач BPC (повна відповідність блоку BPC на рис. 1).
- U4 — мультиплексор data_MX, що вибирає на вихідну шину busOut значення з буферного регістра або з перетворювача BPC в залежності від рівня сигналу вибірки на шині bus_convert.
- U5 — блок convert_analizator (реалізація блоку DA (рис. 1) з обмеженою функціональністю), що визначає, чи потрібно виконувати перетворення вхідного числа в постбінарне представлення: в результаті аналізу сигналів CF_m і MF_m формує відповідний керуючий сигнал вибірки на шині bus_convert, або сигнал помилки представлення формату на шині bus_CA_err.

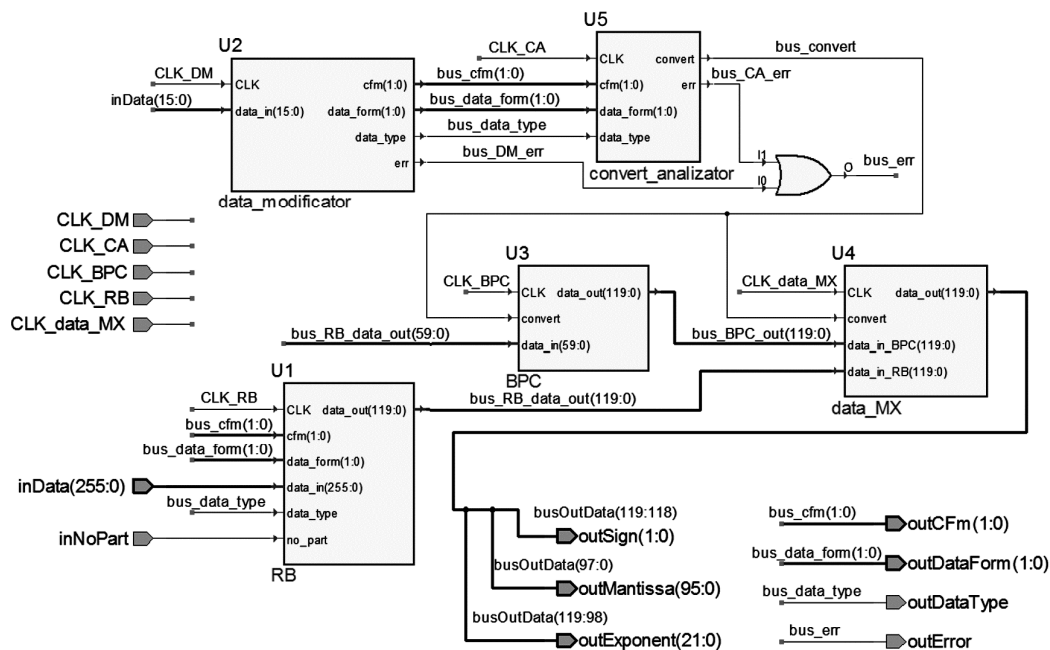


Рисунок 3 – Функціональна схема спрощеної моделі блоку ППФ

Робота моделі блоку ППФ організована за принципом роботи операційного автомата, але в даному випадку на кожен модуль подається власний сигнал синхронізації (CLK). Дана концепція заснована на можливості реалізації покрокового процесу моделювання для поточної моделі і пізніше буде переглянута для загальної моделі постбінарного співпроцесора, що складається з операційної та керуючої частин.

У табл. 2 приведені назви і призначення вхідних і вихідних даних моделі ППФ.

На рис. 4 зображено тимчасову діаграму перетворення формату числа з поля даних, що надійшли на вхід inData перетворювача. Розглянемо роботу моделі по кроках:

Крок 1 — формування значень на виході модуля U2 (data_modifier):

- outError: = 0h — дані розпізнані й справді містять поле постбінарного формату з плаваючою комою;
- outCFm: = 1h — довжина формату 64 біта;
- outDataForm: = 0h — формат представлений бінарним кодом, що означає необхідність постбінарного перетворення;

– outDataType = 0h — формат містить в собі одне число, отже це формат rb64.

Таблиця 2 – Вхідні та вихідні дані моделі ППФ

Назва	Розрядність	Опис
<i>Вхідні дані</i>		
inData	256	Вхідні дані для завантаження в буферний регістр та подальшого аналізу
inNoPart	1	Вказівник номеру частини числа, що підлягає вибірці в форматах дробових та інтервальних чисел (для дробів: 0 – чисельник, 1 – знаменник; для інтервалу: 0 – ліва межа, 1 – права межа). Для форматів одного числа – завжди дорівнює 0. У загальній моделі rFC (рис. 1) зберігається в регістрі R _{FL} .
<i>Вихідні дані</i>		
outCFm	2	Фіксоване поле коду формату для завантаження в регістр прапорів. Несе інформацію про розрядності перетвореного формату (сигнал CFm, табл. 1)
outDataForm	2	Сигнал Data_form (табл. 1) для завантаження в регістр прапорів
outDataType	1	Сигнал Data_type (табл. 1) для завантаження в регістр прапорів
outError	1	Помилка роботи моделі – сигнал err – завантаження формату числа, що не підтримується (0 – формат розпізнано, дані отримані; 1 – аварійне завершення роботи блоку, очікування наступної вибірки)
outSign	2	Постбінарне значення знаку отриманого з формату числа для завантаження в rALU
outExponent	22	Постбінарне значення порядку отриманого з формату числа для завантаження в rALU
outMantissa	96	Постбінарне значення мантиси отримане з формату числа для завантаження в rALU

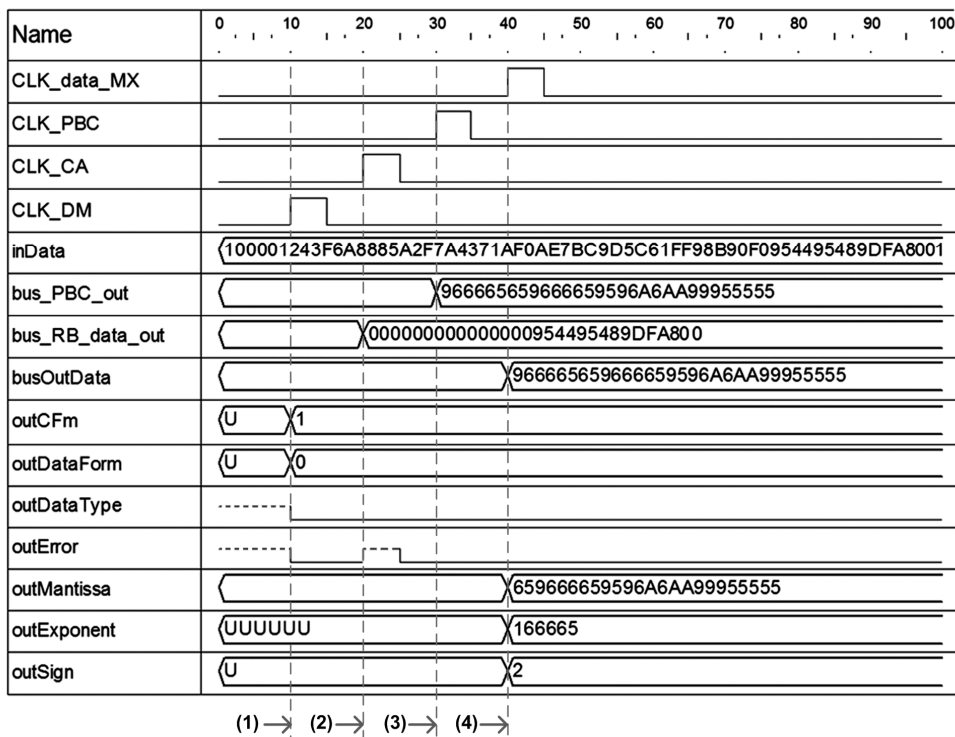


Рисунок 4 – Часова діаграма роботи спрощеної моделі ППФ

Крок 2 — формування сигналу convert := 1h модуля U5 (convert_analizator) — команда модулю PBC про виконання постбінарного перетворення і модулю data_MX для вибірки "постбінарної" шини. На

шину bus_RB_data_out подається інформаційна частина розпізаного формату: bus_RB_data_out := 954495489DFA800h.

Крок 3 — перетворення бінарного коду з шини bus_RB_data_out в постбінарний, та його передача на шину bus_BPC_data_out, оскільки сигнал convert модуля U5 (convert_analizator) дорівнює 1h — команда модулю BPC про виконання постбінарного перетворення.

Крок 4 — Робота модуля U4 (data_MX) яка полягає у формуванні даних на вихідній шині busOutData по керуючому сигналу convert:

$$\text{convert} = \begin{cases} 0 \rightarrow \text{busOutData} := \text{bus_PB_data_out}, \\ 1 \rightarrow \text{busOutData} := \text{bus_BPC_data_out}. \end{cases}$$

Оскільки convert := 1h, то на вихідну шину надходять дані з модуля BPC.

В результаті моделювання, з 256-розрядного бінарного поля даних було розпізано та вибрано число, представлене постбінарним форматом pb64:

954495489DFA8001h \approx -3,2055650328029E-206.

На відповідних виходах моделі сформовано знак, мантису та порядок цього числа у вигляді тетракоду.

Висновки

Таким чином, показана необхідність створення блоку постбінарного перетворювача форматів rFC як ключового блоку співпроцесора rFPU та описана його реалізація.

Представлена в роботі модель має обмежену функціональність, оскільки виконує частину функцій, покладених на блок перетворювача форматів (рис. 1) постбінарного математичного співпроцесора. Подальша реалізація моделі блоку rFC передбачає:

1) підтримку дробових та інтервальних чисел, з організацією черговості обробки частин цих чисел (чисельника і знаменника дробу, меж інтервалу);

2) реалізацію блоку постбінарно-бінарного конвертора BPC.

Модель розроблено мовою опису апаратури VHDL, що дозволяє виконати апаратну реалізацію інтегральної схеми будь-якої складності на ПЛІС [7]. Така реалізація обрана як альтернатива створення дорогих ВІС та НВІС, а також на користь спрощення модифікації і налагодження окремих частин пристрою.

Дану роботу можна вважати першим кроком створення прототипу rFPU — модуля для виконання широкого спектру математичних операцій над числами, представленими в постбінарних форматах, структурі та принцип роботи яких докладно розглянуті в монографії [5].

Перелік джерел

1. IEEE 754: Standard for Binary Floating-Point Arithmetic. [Електронний ресурс]. — Режим доступу: <http://grouper.ieee.org/groups/754/>.
2. Introduction to Programming in Java. Floating point // Computer Science Department of Princeton University <http://introcs.cs.princeton.edu/java/91float>
3. Юровицкий В.М. IEEE754-тика угрожает человечеству. МФТИ, РГСУ, Москва. [Електронний ресурс]. — Режим доступу: <http://www.yur.ru/science/computer/IEEE754.htm>.
4. Аноприенко А.Я. Тетралогика и тетракоды. / В кн. «Сборник трудов факультета вычислительной техники и информатики». Вып. 1. Донецк, ДонГТУ, 1996.
5. Аноприенко А.Я., Иванница С.В. Постбиннарный компьютеринг и интервальные вычисления в контексте кодо-логической эволюции. — Донецк: ДонНТУ, УНИТЕХ, 2011. — 248 с., ил.
6. VHDL coding tips and tricks. Get interesting tips and tricks in VHDL programming. [Електронний ресурс]. — Режим доступу: <http://vhdlguru.blogspot.com/>.
7. ПЛИС — Программируемые Логические Интегральные Схемы. [Електронний ресурс]. — Режим доступу: <http://www.fpga-cpld.ru/>.

Відомості про авторів

Анопрієнко Олександр Якович – к.т.н., проф., декан факультету КНТ ДонНТУ, Україна, 83001, м. Донецьк, вул. Артема, 58, тел. 050-67-77-270, e-mail: anoprien@gmail.com.

Іваниця Сергій Васильович – аспірант кафедри КІ ДоНТУ, Україна, 83001, м. Донецьк, вул. Артема, 58, тел. 050-258-71-18, e-mail: ivanitsa-serg@rambler.ru.

Кулібаба Сергій Володимирович – магістрант кафедри КІ ДоНТУ, Україна, 83001, м. Донецьк, вул. Артема, 58, тел. 095-387-80-55, e-mail: kulibaba@kulibaba.net.

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ТА ОБЧИСЛЮВАЛЬНІ МЕТОДИ

УДК 004.925

М. Д. ОБІДНИК, О. Н. РОМАНЮК

Вінницький національний технічний університет, Вінниця

ПРИСКОРЕНА НОРМАЛІЗАЦІЯ ВЕКТОРІВ ДЛЯ ФОРМУВАННЯ ЗОБРАЖЕНЬ
ВИСОКОПОЛІГОНАЛЬНИХ СЦЕН

Анотація. Запропоновано нові формули для прискореної нормалізації векторів нормалей для формування зображень високодеталізованих поверхонь. Доведено можливість розрахунку з достатньою точністю одиничних векторів без нормалізації. Оцінено похибку використання запропонованих формул для визначення інтенсивності дифузної складової кольору.

Ключові слова: комп'ютерна графіка, нормалізація, апроксимація, дифузна складова кольору.

Аннотация. Предложены новые формулы для ускоренной нормализации векторов нормалей для формирования изображений высокодетализованных поверхностей. Доказана возможность расчета с достаточной точностью единичных векторов без нормализации. Оценено погрешность использования предложенных формул для определения интенсивности диффузной составляющей цвета.

Ключевые слова: компьютерная графика, нормализация, аппроксимация, диффузная составляющая цвета. **Abstract.** The new formula for rapid normalization of vectors of normal imaging highly detailed surfaces. Proved possible to calculate with sufficient accuracy unit vectors without normalization. Error by using the proposed formulas for determining the intensity of the diffuse component of the color.

Keywords: computer graphics, normalization, approximation, diffuse color component.

Вступ

Особливістю високодеталізованої поверхні є висока щільність полігональної мережі, якою вона задається. Це означає, що полігони мають відносно невеликі розміри по відношенню до самого об'єкта візуалізації.

Для розрахунку дифузної та спекулярної складових освітленості кожного полігона розраховуються вектори (нормалі), що задають його орієнтацію в тривимірному просторі. При використанні складних методів тонування, наприклад Фонга, нормалі інтерполюються вздовж ребер, а у подальшому вздовж рядків rasterизації. Кожний вектор, отриманий у результаті інтерполяції, нормалізується. Таке тонування дозволяє створити візуальний ефект кривизни полігона. Проте, зі зменшенням розмірів полігонів, кут між нормалями, що знаходяться у вершинах, а відповідно і різниця між нормалізованим та ненормалізованим векторами зменшується. Це дає можливість спростити обчислювальну складність процесу тонування за рахунок використання простіших з обчислювальної точки зору формул для нормалізації.

Аналіз методів і постановка задачі

Загальна формула нормалізації має вигляд [1]:

$$\vec{N}_n = \frac{\vec{N}}{|\vec{N}|} = \frac{\vec{N}}{\sqrt{X^2 + Y^2 + Z^2}}, \quad (1)$$

де \vec{N} – вихідний вектор, \vec{N}_n – нормалізований вектор, а X , Y , Z – координати вихідного вектора \vec{N} .

Нормалізація векторів – це одна з найбільш поширених процедур, яка виконується у процесі візуалізації тривимірної сцени. У роботах [1, 2] був проведений аналіз методів прискореної нормалізації.

Барера, Хаст і Бенгтссон у своїй роботі отримали формулу, яка використовує один крок ітерації Ньютона-Рафсона [3]:

$$\vec{N}_n = \frac{\vec{N}}{3} (3 - \vec{N} \cdot \vec{N}).$$

Розрахунок згідно з наведеною формулою вимагає виконання 1 операції ділення, 9 операцій множення та 3 операцій типу додавання. Великі похибки визначення ортогональних складових вектора обмежує використання формули для зафарбовування, оскільки в цьому випадку мають місце суттєві візуальні відмінності порівняно з еталоном [2].

Згідно з підходом Р.Ліона [4] вираз $1/\sqrt{\vec{N} \cdot \vec{N}}$, який використовується для нормалізації вектора нормалі \vec{N} , розкладають у ряд Тейлора та обмежуються першими трьома членами:

$$\vec{N}_n = \vec{N} \left(1 - \frac{1}{2} (\vec{N} \cdot \vec{N} - 1) + \frac{3}{8} (\vec{N} \cdot \vec{N} - 1)^2 \right).$$

Формула хоча і придатна для апаратної реалізації, але не забезпечує прийнятної точності, що обмежує її використання [2].

Інтерполювати одиничні вектори нормалей між початковим \vec{N}_1 і кінцевим \vec{N}_m векторами рядку растеризації, які мають одиничну довжину, можна за формулою [5]:

$$\vec{N}_n(w) = \vec{N}_1 \frac{\sin((1-w)\psi)}{\sin\psi} + \vec{N}_m \frac{\sin(w\psi)}{\sin\psi},$$

де $w \in [0,1]$, а ψ – кут між векторами нормалей \vec{N}_1 і \vec{N}_m .

Суттєвим недоліком запропонованої формули є необхідність розрахунку тригонометричних функцій \sin і \arccos для розрахунку невідомих параметрів w і ψ . Крім того, розрахунок $\vec{N}_n(w)$ передбачає виконання операції ділення.

Розглянуті методи спрощення нормалізації векторів не враховують рівень деталізації тривимірної сцени, зображення якої формується. При підвищенні деталізації тривимірної сцени зменшуються як розміри полігонів, так і кути між нормаллями в вершинах полігонів. Це дає можливість використати простіші та точніші методи прискореної нормалізації.

Мета досліджень

Спрощення процедури нормалізації векторів при обчисленні інтенсивності дифузної складової кольору відповідно до рівня деталізації сцени.

Розв'язання задачі

Трикутні полігони, з яких складаються тривимірні моделі, задаються вершинами (вертексами), а також нормаллями у них. Якщо вектори нормалей у вершинах полігону не колінеарні, то трикутник, що зафарбовується, виглядає візуально опуклим. При цьому, чим більші кути між нормаллями у вершинах полігону, тим більша його візуальна опуклість. Кількісно опуклість можна оцінити косинусом найбільшого з кутів між нормаллями полігону ($\cos\varphi$). Задамо порогове значення опуклості полігону Θ . За умови $\cos\varphi \geq \Theta$, полігон відповідає заданому рівню опуклості. При збільшенні деталізації тривимірних об'єктів кути між нормаллями у вершинах полігонів зменшуються. Косинус між двома нормалізованими векторами можна знайти за формулою:

$$\cos\varphi = \vec{N}_1 \cdot \vec{N}_2 = X1 \cdot X2 + Y1 \cdot Y2 + Z1 \cdot Z2,$$

де $\vec{N}_1(X1, Y1, Z1)$ і $\vec{N}_2(X2, Y2, Z2)$ – нормалізовані вектори, а $\cos\varphi$ – косинус кута між ними.

Використовуючи порогове значення Θ , можна визначити відповідність тривимірної сцени заданому рівню опуклості. Якщо вона не відповідає цьому заданому рівню, то її можна привести до нього за рахунок виконання додаткової триангуляції [6].

Спрощення зафарбовування полігонів можна досягти за рахунок вилучення операції нормалізації з процесу растеризації або апроксимації загальної формули простішою. Доцільність такого спрощення можна оцінити за значеннями відносних похибок такої нормалізації.

Знайдемо максимальну різницю між нормалізованим і ненормалізованим векторами при інтерполяції нормалей вздовж рядка растеризації. Початкові точки векторів \vec{N}_A , і \vec{N}_B зведемо в одну точку, як показано на рис. 1. Оскільки всі вектори у результаті інтерполяції мають бути одиничної довжини, то їх кінці розміщуються на дузі, яка є частиною кола з одиничним радіусом. Найбільше значення похибки буде спостерігатись у середині ребра полігона, оскільки хорда найбільш віддалена від кола у середній точці. На рисунку 1 вектор, що знаходиться у середній точці ребра полігона, отриманий у результаті лінійної інтерполяції, позначено $\vec{N}_{1/2}$; нормалізований вектор у цій же точці – \vec{N} ; а різницю між нормалізованим і ненормалізованим векторами – $\Delta\vec{N}$.

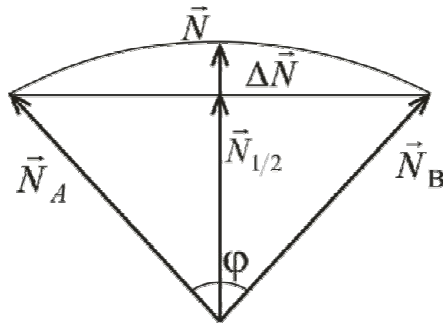


Рисунок 1 – Максимальна різниця нормалізації

Для зменшення обчислювальних затрат на процедуру нормалізації, можна апроксимувати загальну формулу поліномом Чебишева, враховуючи порогове значення Θ . При цьому доцільно обмежитись двома або трьома членами поліному залежно від необхідного рівня точності апроксимації. Формулу (1) можна записати у вигляді:

$$\vec{N}_n = \vec{N} \cdot \frac{1}{\sqrt{d}},$$

де $\sqrt{d} = |\vec{N}|$, $d = X^2 + Y^2 + Z^2 = |\vec{N}|^2$. d використаємо у якості змінної при апроксимації поліномом Чебишева.

Лінійна апроксимація формули (1) має вигляд:

$$\vec{N}_n = \vec{N} \cdot (k \cdot d + p), \quad (2)$$

де k і p – скалярні коефіцієнти апроксимації.

Порогове значення опуклості Θ полігону порівнюється з величиною косинуса найбільшого з кутів між нормаліями у його вершинах, а значення цього кута становить $\arccos \Theta$. Якщо у поточному полігоні значення опуклості не перевищує порогове, то виконується нерівність $\cos \varphi \geq \Theta$. Косинус найбільшого із кутів між нормаліями у вершинах полігону розраховується за формулою $\cos \varphi = \vec{N}_A \cdot \vec{N}_B$. Максимальний діапазон, у якому змінюється довжина векторів, отриманих у результаті лінійної інтерполяції, вздовж ребер полігону становить $(|\vec{N}_{1/2}|, |\vec{N}|)$. Із рис. 1 видно, що $|\vec{N}_{1/2}| = \cos(\varphi/2)$.

Знайдемо діапазон, у якому змінюється значення d . Оскільки $d = X^2 + Y^2 + Z^2 = |\vec{N}|^2$, то $d_{1/2} = X_{1/2}^2 + Y_{1/2}^2 + Z_{1/2}^2 = |\vec{N}_{1/2}|^2$, тому діапазон зміни d становить:

$$\left(|\vec{N}_{1/2}|^2, |\vec{N}|^2 \right).$$

У табл. 1 наведено формули лінійної апроксимації для різних порогових значень Θ . Порогові значення вибрані з міркувань простоти обчислень.

Для підвищення точності апроксимації можна застосувати формулу Чебишева з трьома членами. Хоча при цьому збільшується кількість обчислень на нормалізацію, але отримана формула простіша за (1). Квадратична апроксимація формули (1) має такий вигляд:

$$\vec{N}_n = \vec{N} \cdot (a \cdot d^2 + b \cdot d + c), \quad (3)$$

де a, b, c – скалярні коефіцієнти квадратичної апроксимації.

Таблиця 1 – Формули лінійної апроксимації формули (1)

Порогове значення, Θ	Макс. кут між нормаллями, $\arccos(\Theta)$	Діапазон зміни d , $(d_{1,2}, 1)$	Формула спрощеної нормалізації відповідно до (2)	Максимальна відносна похибка, δ_l
0	90°	(0.5, 1)	$-0.798 \cdot d + 1.778$	2.475 %
1/2=0.5	60°	(0.75, 1)	$-0.615 \cdot d + 1.611$	0.406 %
3/4=0.75	41.4°	(0.874, 1)	$-0.552 \cdot d + 1.551$	0.087 %
7/8=0.875	28.96°	(0.937, 1)	$-0.525 \cdot d + 1.525$	0.02 %
15/16=0.9375	20.4°	(0.968, 1)	$-0.512 \cdot d + 1.512$	0.005 %
31/32=0.96875	14.4°	(0.984, 1)	$-0.506 \cdot d + 1.506$	0.0012 %

У табл. 2 наведено формули для квадратичної апроксимації для різних порогових значень Θ . Як видно з даних, наведених у таблиці, точність квадратичної апроксимації значно вища, ніж лінійної.

Таблиця 2 – Формули квадратичної апроксимації формули (1)

Порогове значення, Θ	Макс. кут між нормаллями, $\arccos(\Theta)$	Діапазон зміни d , $(d_{1/2}, 1)$	Формула спрощеної нормалізації відповідно до (3)	Максимальна відносна похибка, δ_k
0	90°	(0.5, 1)	$0.838 \cdot d^2 - 2.07 \cdot d + 2.235$	0.323 %
1/2=0.5	60°	(0.75, 1)	$0.532 \cdot d^2 - 1.547 \cdot d + 2.016$	0.023 %
3/4=0.75	41.4°	(0.874, 1)	$0.443 \cdot d^2 - 1.382 \cdot d + 1.939$	0.0024 %
7/8=0.875	28.96°	(0.937, 1)	$0.407 \cdot d^2 - 1.312 \cdot d + 1.906$	0.00027 %
15/16=0.9375	20.4°	(0.968, 1)	$0.391 \cdot d^2 - 1.281 \cdot d + 1.89$	0.000034 %
31/32=0.96875	14.4°	(0.984, 1)	$0.383 \cdot d^2 - 1.265 \cdot d + 1.883$	0.000004 %

Оцінимо максимальну відносну похибку лінійної інтерполяції векторів, виконаної без нормалізації, за формулою:

$$\delta = \frac{|\vec{N}| - |\vec{N}_{1/2}|}{|\vec{N}|} \cdot 100\%.$$

Відповідно до заданих порогових значень Θ , розраховано значення відносних похибок лінійної інтерполяції, які зведено у таблицю 3. Також у таблиці наведено значення похибок лінійної та квадратичної апроксимації формули (1). Це дає змогу порівняти точність, з якою виконуються спрощенні розраху-

нки, а також доцільність застосування тої чи іншої формули залежно від порогового значення Θ .

Таблиця 3 – Порівняння похибок лінійної інтерполяції нормалей та отриманих апроксимаційних формул

Порогове значення, Θ	Макс. кут між векторами у ряд. раст. чи на ребрі полігону, φ	Максимальна відносна похибка розрахунків, виконаних без нормалізації, δ_i	Максимальна відносна похибка прискореної нормалізації, виконаної за формулою (2), δ_i	Максимальна відносна похибка прискореної нормалізації, виконаної за формулою (3), δ_k
0	90°	29.29 %	2.475 %	0.323 %
1/2=0.5	60°	13.4 %	0.406 %	0.023 %
3/4=0.75	41.4°	6.5 %	0.087 %	0.0024 %
7/8=0.875	28.96°	3.2 %	0.02 %	0.00027 %
15/16=0.9375	20.4°	1.6 %	0.005 %	0.000034 %
31/32=0.96875	14.4°	0.78 %	0.0012 %	0.000004 %

Наведені у таблицях 1 – 3 дані описують прискорену нормалізацію для високодеталізованих поверхонь з різними пороговими значеннями Θ . Оцінимо вплив похибок, які мають місце при використанні методів прискореної нормалізації, на результуюче зображення. Для цього потрібно розрахувати похибки визначення інтенсивності кольору, які мають місце у процесі зафарбовування.

Розрахуємо максимальну абсолютну похибку визначення інтенсивності кольору за умови, що визначення дифузної складової світла проводиться за формулою:

$$I_i = I_0 \cdot k_m \cdot (\vec{L} \cdot \vec{N}),$$

де I_i – інтенсивність дифузної складової світла у поточній точці, I_0 – інтенсивність випромінювання джерела світла, k_m – коефіцієнт дифузного відбиття, \vec{L} – вектор, який задає напрямок до джерела світла.

Максимальну абсолютну похибку визначення інтенсивності дифузної складової кольору ΔI_d можна визначити за формулою:

$$\Delta I_d = I_0 \cdot k_m \cdot (\vec{L} \cdot \vec{N}) - I_0 \cdot k_m \cdot (\vec{L} \cdot \vec{N}_{1/2}).$$

Враховуючи властивості скалярного добутку, можна записати:

$$\begin{aligned} \Delta I_d &= I_0 \cdot k_m \cdot (\vec{L} \cdot \vec{N}) - I_0 \cdot k_m \cdot (\vec{L} \cdot (\vec{N} - \Delta \vec{N})) = \\ &= I_0 \cdot k_m \cdot (\vec{L} \cdot (\vec{N} - (\vec{N} - \Delta \vec{N}))) = I_0 \cdot k_m \cdot (\vec{L} \cdot \Delta \vec{N}). \end{aligned}$$

Розрахуємо максимальну відносну похибку інтенсивності дифузної складової кольору δI_d

$$\delta I_d = \frac{\Delta I}{I_i} = \frac{I_0 \cdot k_m \cdot (\vec{L} \cdot \Delta \vec{N})}{I_0 \cdot k_m \cdot (\vec{L} \cdot \vec{N})} = \frac{\Delta \vec{N}}{\vec{N}}$$

Одиничний вектор нормалі в середній точці ребра трикутника розраховується за формулою [6]:

$$\vec{N} = \frac{\vec{N}_A + \vec{N}_B}{\sqrt{2(1 + \vec{N}_A \cdot \vec{N}_B)}}.$$

Вектор $\vec{N}_{1/2}$ розраховується за формулою

$$\vec{N}_{1/2} = \frac{\vec{N}_A + \vec{N}_B}{2}.$$

Знайдемо різницю $\Delta\vec{N}$:

$$\begin{aligned} \Delta\vec{N} &= \vec{N} - \vec{N}_{1/2} = \frac{\vec{N}_A + \vec{N}_B}{\sqrt{2 \cdot (1 + \vec{N}_A \cdot \vec{N}_B)}} - \frac{\vec{N}_A + \vec{N}_B}{2} = \\ &= \frac{\sqrt{2}(\vec{N}_A + \vec{N}_B) - (\sqrt{1 + \vec{N}_A \cdot \vec{N}_B}) \cdot (\vec{N}_A + \vec{N}_B)}{2 \cdot \sqrt{(1 + \vec{N}_A \cdot \vec{N}_B)}} = \\ &= \frac{(\vec{N}_A + \vec{N}_B) \cdot (\sqrt{2} - \sqrt{1 + \vec{N}_A \cdot \vec{N}_B})}{2 \cdot \sqrt{(1 + \vec{N}_A \cdot \vec{N}_B)}} \end{aligned}$$

Отже, при розрахунках векторів, виконаних без нормалізації, максимальна відносна похибка інтенсивності дифузної складової кольору δI_{di} визначається за формулою:

$$\begin{aligned} \delta I_{di} &= \frac{\Delta I}{I_i} = \frac{I_0 \cdot k_m \cdot (\vec{L} \cdot \Delta\vec{N})}{I_0 \cdot k_m \cdot (\vec{L} \cdot \vec{N})} = \frac{I_0 \cdot k_m \cdot \left(\vec{L} \cdot \frac{(\vec{N}_A + \vec{N}_B) \cdot (\sqrt{2} - \sqrt{1 + \vec{N}_A \cdot \vec{N}_B})}{2 \cdot \sqrt{(1 + \vec{N}_A \cdot \vec{N}_B)}} \right)}{I_0 \cdot k_m \cdot \left(\vec{L} \cdot \frac{\vec{N}_A + \vec{N}_B}{\sqrt{2 \cdot (1 + \vec{N}_A \cdot \vec{N}_B)}} \right)} = \\ &= \frac{I_0 \cdot k_m \cdot \left(\vec{L} \cdot \frac{(\vec{N}_A + \vec{N}_B) \cdot (\sqrt{2} - \sqrt{1 + \vec{N}_A \cdot \vec{N}_B})}{2 \cdot \sqrt{(1 + \vec{N}_A \cdot \vec{N}_B)}} \right)}{I_0 \cdot k_m \cdot \left(\vec{L} \cdot \frac{\vec{N}_A + \vec{N}_B}{\sqrt{2 \cdot (1 + \vec{N}_A \cdot \vec{N}_B)}} \right)} = \frac{\sqrt{2} - \sqrt{1 + \vec{N}_A \cdot \vec{N}_B}}{\sqrt{2}} = \\ &= 1 - \sqrt{\frac{1 + \vec{N}_A \cdot \vec{N}_B}{2}}. \end{aligned}$$

Визначимо максимальну відносну похибку інтенсивності дифузної складової кольору δI_{dl} , отриману в результаті застосування лінійної апроксимації формули (1):

$$\begin{aligned}\delta I_{dl} &= \frac{\Delta \bar{N}}{\bar{N}} = \frac{\bar{N}_{1/2} - \bar{N}_{1/2} \cdot (k \cdot d_{1/2} + p)}{\bar{N}_{1/2}} = \frac{1 - \frac{|\bar{N}_{1/2}| \cdot (k \cdot d_{1/2} + p)}{|\bar{N}_{1/2}|}}{\frac{1}{|\bar{N}_{1/2}|}} = \\ &= \frac{1 - \frac{|\bar{N}_{1/2}| \cdot (k \cdot d_{1/2} + p)}{|\bar{N}_{1/2}|}}{1} = 1 - |\bar{N}_{1/2}| \cdot (k \cdot d_{1/2} + p).\end{aligned}$$

Оскільки $d_{1/2} = X_{1/2}^2 + Y_{1/2}^2 + Z_{1/2}^2 = |\bar{N}_{1/2}|^2$, то

$$\delta I_{dl} = 1 - \sqrt{d_{1/2}} \cdot (k \cdot d_{1/2} + p).$$

Аналогічно, максимальна відносна похибка визначення інтенсивності дифузної складової кольору δI_{dk} , отримана в результаті застосування квадратичної апроксимації формули (1), можна розрахувати за формулою:

$$\delta I_{dk} = 1 - \sqrt{d_{1/2}} \cdot (a \cdot d_{1/2}^2 + b \cdot d_{1/2} + c).$$

Розраховані значення максимальних відносних похибок визначення інтенсивності дифузної складової кольору, для обчислень виконаних без нормалізації, а також із застосуванням лінійної та квадратичної апроксимації формули (1), наведено у табл. 4.

Таблиця 4 – Значення максимальних відносних похибок визначення інтенсивності дифузної складової кольору

Θ	φ	δI_{di}	δI_{dl}	δI_{dk}
0	90°	29.29 %	2.49 %	0.35 %
1/2=0.5	60°	13.4 %	0.429 %	0.026 %
3/4=0.75	41.4°	6.5 %	0.103 %	0.012 %
7/8=0.875	28.96°	3.2 %	0.04 %	0.09 %
15/16=0.9375	20.4°	1.6 %	0.0094 %	0.0026 %
31/32=0.96875	14.4°	0.78 %	0.0024 %	0.098 %

З наведених формул видно, що при використанні прискореної нормалізації досягається відносно висока точність визначення дифузної складової кольору.

Висновки

У результаті проведених досліджень встановлено, що спрощення розрахунку інтенсивності дифузної складової кольору при тонуванні високодеталізованих полігональних поверхонь можна досягти як за рахунок спрощення трудомісткої процедури нормалізації. Із даних, наведених у табл. 1–4, можна вибрати формули прискореної нормалізації залежно від заданого порогового значення Θ . Відомо, що візуальна ідентичність зображень досягається при умові, що відносна похибка визначення інтенсивності кольору не перевищує 4 %.

Результати досліджень можна використати для побудови високопродуктивних програмних та апаратних засобів комп'ютерної графіки.

Список використаної літератури

1. Tomas Akenine-Moller. Real-Time Rendering, Third Edition / Tomas Akenine-Moller, Eric Haines, Naty Hoffman // A K Peters, Ltd. Wellesley, Massachusetts. – 2008. – 1027 pp. – ISBN-13: 978-1-56881-424-7.
2. Романюк О. В. Аналіз методів нормалізації векторів нормалей для задач формування тривимірних зображень [Електронний ресурс] / Романюк О. В., Войтко В. В. // Наукові праці ВНТУ. – 2009. – № 1. – Режим доступу: <http://praci.vntu.edu.ua/article/view/1163/374>.
3. Barrera T. Fast Near Phong-Quality Software / T. Barrera, A. Hast, E. Bengtsson // Shading SCCG'06 full paper, 2006, pp 44-50.
4. Lyon R.F. Phong Shading Reformulation for Hardware Renderer Simplification / R.F.Lyon // Apple Technical Report № 43. – 1993.
5. Hast A. Improved Algorithms for Fast Shading and Lighting / A. Hast // Acta Universitatis Upsaliensis. Uppsala, 2004. – 116 p.
6. Романюк О. Н. Методи додаткової триангуляції / Романюк О. Н., Обідник М. Д. // Вестник Херсонського національного технічного університету. – 2011. – № 3(42). – С. 366–371. – ISSN 2078 – 4481.

Відомості про авторів

Обідник Микола Дем'янович – аспірант кафедри програмного забезпечення, Вінницький національний технічний університет, (0432)598243, м. Вінниця, вул. Хмельницьке шосе 95.

Романюк Олександр Никифорович – д. т. н., проф., перший проректор Вінницького національного технічного університету, (0432)598243, м. Вінниця, вул. Хмельницьке шосе 95.

УДК 681.518.3

С. Ш. КАЦІВ

Вінницький національний технічний університет, Вінниця

ШТУЧНА НЕЙРОПОДІБНА МЕРЕЖА ДЛЯ ДІАГНОСТУВАННЯ ДЕФЕКТІВ ГІДРОАГРЕГАТУ

Анотація. Гідроагрегат є дуже складною динамічною гідроелектромеханічною системою і адекватний математичний опис залежності віброакустичного сигналу від всіх чинників, які викликають вібрацію, практично неможливий. Тому в цій роботі запропоновано для автоматизованого діагностування існуючих дефектів гідроагрегату використовувати тришарову штучну нейроподібну мережу, яка дозволить розглядати гідроагрегат як "чорну скриньку". Наведена структура цієї мережі та її математична модель. Розглянуті деякі перерізи амплітудно-частотно-часових спектрів вібросигналів гідроагрегату і на їх основі зроблені попередні діагностичні висновки.

Ключові слова: діагностування дефектів гідроагрегату, штучна нейроподібна мережа, амплітудно-частотно-часовий спектр.

Аннотация. Гидроагрегат является очень сложной динамической гидроэлектро-механической системой и адекватное математическое описание зависимости виброакустического сигнала от всех факторов, вызывающих вибрацию, практически невозможно. Поэтому в этой работе предложено для автоматизированного диагностирования существующих дефектов гидроагрегата использовать трехслойную искусственную нейроподобную сеть, что позволит рассматривать гидроагрегат как "черный ящик". Приведена структура этой сети и ее математическая модель. Рассмотрены некоторые сечения амплитудно-частотно-временных спектров вибросигналов гидроагрегата и на их основе сделаны предварительные диагностические выводы.

Ключевые слова: диагностирование дефектов гидроагрегата, искусственная нейроподобная сеть, амплитудно-частотно-временной спектр.

Abstract. The hydrounit is very difficult dynamic hydroelectromechanical system and adequate mathematical description of dependence vibrating signal by all factors causing vibration, practically is impossible. Therefore in this work it is offered for the automated diagnosing of existing defects of the hydrounit to use three-layer artificial neurosimilar network, that will allow to consider the hydrounit as «black box». The structure of this network and its mathematical model is given. Some sections peak-frequency-hour of spectra of vibrating signals of the hydrounit are considered and on their basis the preliminary diagnostic conclusions are made.

Keyword: diagnosing of defects of hydrounits, artificial neurosimilar network, peak-frequency-hour of spectra.

Вступ

Вітчизняний і закордонний досвід доводить, що впровадження засобів діагностування є одним із найважливіших чинників підвищення економічної ефективності використання устаткування в промисловості. Призначення діагностування – виявлення і попередження відмов та несправностей, підтримка експлуатаційних показників у встановлених межах, прогнозування стану обладнання з метою повного використання ресурсу.

Одним з найпоширеніших видів діагностування є вібродіагностування, оскільки практично миттєва реакція вібросигналу на зміну стану обладнання є незамінною якістю в аварійних ситуаціях, коли визначальним чинником є швидкість постановки діагнозу і прийняття рішення, що черговий раз довела аварія на Саяно-Шушенський ГЕС.

Вібродіагностування – це галузь знань, що включає в собі теорію і методи організації процесів розпізнавання технічних станів машин і механізмів по вихідній інформації, що міститься у віброакустичному сигналі.

Основним фізичним носієм інформації про стан елементів працюючого устаткування у вібродіагностуванні є віброакустичний сигнал – збірне поняття, що містить інформацію про коливальні процеси (вібраційні, гідро- або газодинамічні тощо) і акустичний шум механізму в навколишньому середовищі. Отже вібродіагностуванню може піддаватися будь-яке устаткування, функціонування якого супроводжується коливальними процесами. Всяке відхилення параметрів функціонування устаткування від норми призводить до зміни характеру взаємодії його елементів і до зміни супровідної взаємодії віброакустичних процесів [1,2].

Протягом останніх 20 років у світі (і зокрема на теренах СНД) ведеться активна робота по створенню автоматизованих комп'ютерних систем моніторингу та діагностування дефектів машин та обладнання. Одними з перших були системи, розроблені в АО ВАСТ (м. С.-Петербург), які базувалися на спектральному аналізі вібросигналів, та забезпечували вібродіагностування роторних машин [2,3].

В останні роки створені спеціалізовані комп'ютерні системи для діагностування гідроагрегатів, виробництва ТОВ "Ампер" (комплекс КНК-32) та ДІАМЕХ2000 (комплекс АЛІМА3-7010-ГЕС), які теж базуються на спектральному аналізі вібросигналів.

Недоліком вищезгаданих систем діагностування є те, що спектр сигналу отримується за допомогою звичайного перетворення Фур'є, яке дає адекватний результат лише для стаціонарних сигналів, в той час як вібросигнал гідроагрегату є суттєво нестационарним.

Тому останнім часом проводяться дослідження, які виправляють цей недолік (наприклад, [4,5]), але при цьому нестационарний вібросигнал трансформується в тривимірний амплітудно-частотно-часовий спектр (АЧЧС) з допомогою віконного перетворення Фур'є, що має фіксовану ширину вікна і, внаслідок цього, недостатню роздільну здатність.

Автором було запропоновано спектральний аналіз вібросигналів здійснювати за допомогою дискретного вейвлет-перетворення (ДВП), яке має змінну ширину вікна і тому значно кращу роздільну здатність [6-9].

Разом з тим, навіть за наявності АЧЧС вібросигналу діагностування існуючих дефектів гідроагрегатів є досить складною задачею. Справа в тому, що гідроагрегат є дуже складною динамічною гідроелектромеханічною системою і математичний опис залежності віброакустичного сигналу від всіх чинників, які викликають вібрацію, практично неможливий. Особливо це стосується гідромеханічної складової вібросигналу, яка викликається турбулентністю, кавітацією, виникненням стоячих хвиль тощо [1, 10].

Тому доцільно розглядати гідроагрегат як "чорну скриньку", тобто моделювати не його структуру, а зовнішнє функціонування.

Досить часто таке моделювання здійснюється за допомогою нейромережних технологій [11 – 13].

Найбільш поширеними типами нейронних мереж, що використовуються при моделюванні складних технічних систем є мережі Кохонена, перцептрони, ймовірнісні та нечіткі мережі, нейронні мережі адаптивної резонансної теорії тощо. Вибір типу мережі та її структури залежить як від вхідних потоків даних, так і від вихідних діагностичних висновків.

Для розв'язання задачі вібродіагностування передбачається створення системи автоматизованого діагностування і прогнозування розвитку дефектів гідроагрегатів (САДП-РДГ) [6].

В цій системі основні масиви вхідних даних формуються за допомогою ДВП вібросигналів, які надходять від вібросенсорів ВД1 – ВД4, що встановлені по радіальним вертикальному і горизонтальному напрямкам турбінного і опорно-упорного підшипників гідроагрегату. Додатково до віброакустичних сигналів САДП-РДГ отримує дані про струм навантаження та частоту обертів гідроагрегата, а також рівень води у водосховищі.

Вихідні діагностичні висновки повинні відповідати основним чинникам [1], що викликають вібрацію гідроагрегату, а саме: нерівноваженість ротора; порушення жорсткості опорної системи; дефекти турбінного і опорно-упорного підшипників; вібрація електричних машин електромагнітного походження; порушення гідродинаміки потоку.

Аналізуючи вищесказане, можна дійти до висновку, що неоднорідність вхідних масивів даних, складність задачі, значна невизначеність експертних оцінок призводять до недоцільності використання стандартних типів нейронних мереж і вимагають розробки нестандартної штучної *нейроподібної* мережі (ШНМ).

Актуальність задачі

Задача своєчасного діагностування існуючих дефектів гідроагрегату в режимі квазіреального часу останнім часом набула підвищеної актуальності після всім відомої аварії на Саяно-Шушенській ГЕС, яка викликала людські жертви і завдала значних економічних збитків Російській Федерації.

Мета задачі

Метою задачі є побудова ШНМ для діагностування існуючих дефектів гідроагрегату.

Структура ШНМ САДП-РДГ для діагностування існуючих дефектів гідроагрегату

Для діагностування дефектів гідроагрегату запропонована тришарова ШНМ, яка зображена на рис. 1.

Кількість вхідних нейронів (на рис. 1 вони зображені колами) залежить від кількості частотних смуг в АЧЧС кожного з 4-ох вібросигналів. Якщо кількість частотних смуг в АЧЧС дорівнює M , то кількість вхідних нейронів дорівнює $4(M + 1) + 2$. Вхідні нейрони виконують функцію прийняття числових даних і їх сортування. Розміри числових масивів, які надходять на кожний вхідний нейрон різні. Так, на перший вхідний нейрон кожної з 4-ох груп надходять неприпустимі значення віброзміщення та їх часові відліки, розмір цього масиву даних важко передбачити – він може коливатися від 0 до декількох тисяч.

На другий вхідний нейрон надходять один вейвлет-коефіцієнт першої смуги частот, на третій – k вейвлет-коефіцієнтів другої смуги частот, на четвертий – k^2 вейвлет-коефіцієнтів третьої смуги частот, на останній нейрон кожної групи надходить k^{M-1} вейвлет-коефіцієнтів M -ої смуги частот (k – коефіцієнт стиснення ДВП).

Кількість значень рівня води у водосховищі з їх часовими відліками теж може коливатися у широких межах, оскільки нове значення рівня повинно надходити до вхідного нейрона лише при його зміні. За аналогічним принципом надходять дані і про струм навантаження гідроагрегатора.

Перший шар ШНМ (позначений цифрою 1) містить $4M$ нейронів. Кожен з них отримує вейвлет-коефіцієнти певної смуги частот від відповідного вхідного нейрона і, одночасно, масив значень рівня води у водосховищі і масив значень струму навантаження гідроагрегатора. Нейрони першого шару при-

значені для виділення із АЧЧС кожного з чотирьох вібросигналів спектральних складових, які викликаються гідродинамічними та електродинамічними чинниками. Крім того, ці нейрони повинні для кожної частоти АЧЧС визначити *фонові* спектральні складові.

Ми будемо називати *фоновими* спектральні складові вібросигналу, коли гідроагрегат, для якого проводиться діагностування, з якихось причин зупинений, а інші агрегати Дністровської ГЕС-2 працюють. В цьому випадку вібросигнали генеруються працюючими гідроагрегатами і через будівельні конструкції передаються на зупинений гідроагрегат, де і фіксуються його вібросенсорами.

Всі ці процедури налаштування, які виконує перший шар ШНМ, мають відбуватися протягом дослідної експлуатації САДП-РДГ.

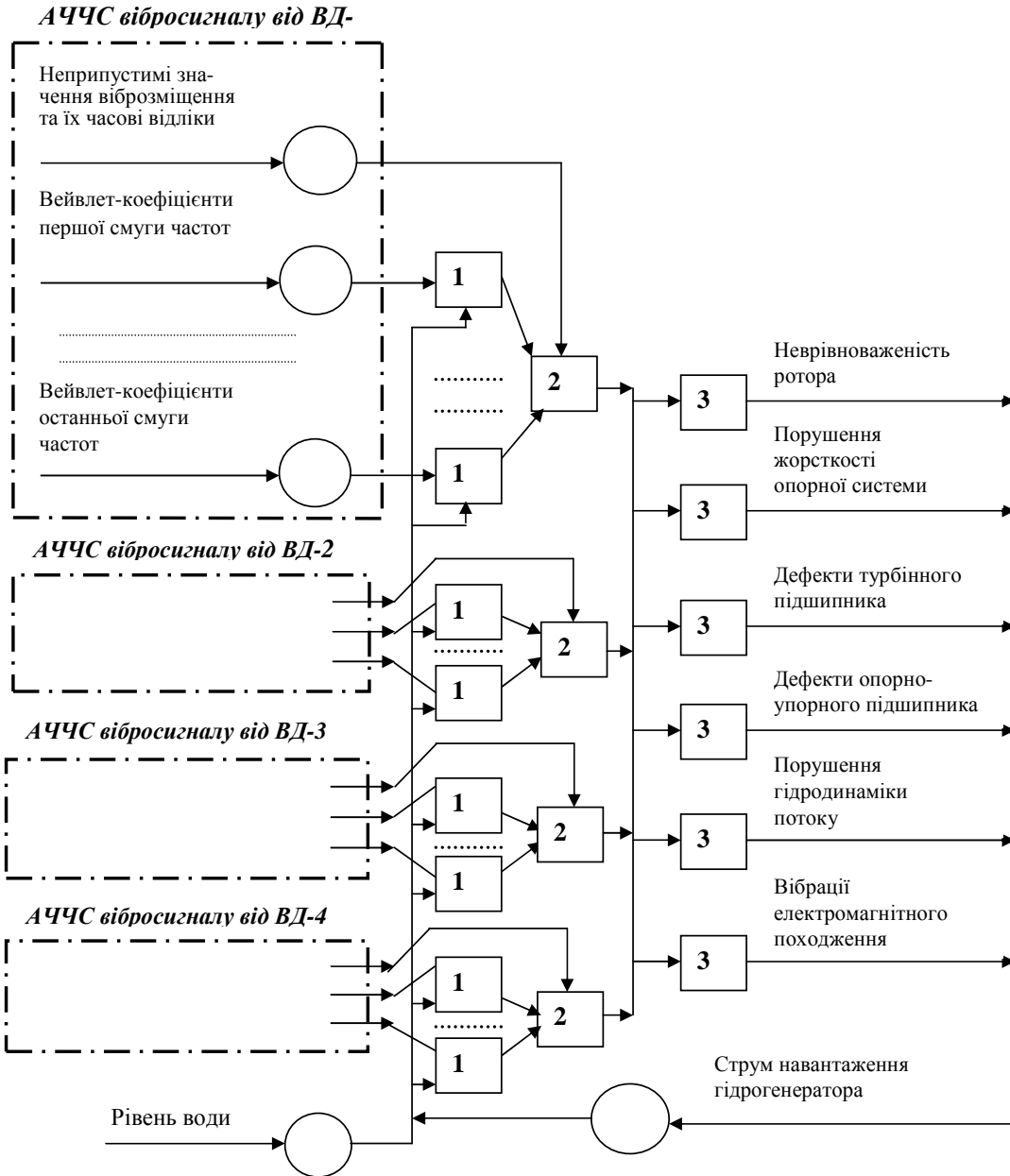


Рисунок 1 – Структура ШНМ для діагностування дефектів гідроагрегату

Другий шар ШНМ (позначений цифрою 2) містить 4 нейрона. Кожен з них отримує амплітудно-частотно-часовий спектр одного з чотирьох віброакустичних сигналів, а також неприпустимі значення віброзміщення та їх часові відліки для цього вібросигналу. Крім того, до кожного нейрону надходять дані про можливу залежність певних смуг частот відповідної АЧЧС від гідродинамічних та електродинамічних чинників і фонові спектральні характеристики, що є результатом налаштування першого шару ШНМ. Кожен нейрон другого шару визначає переріз певного АЧЧС для усіх моментів часу, коли віброзміщення цього вібросигналу перевищує припустиме значення.

Третій шар ШНМ (позначений цифрою 3) містить 6 нейронів, кожен з яких відповідає одному з чинників, які є причинами виникнення вібрацій. Кожний нейрон третього шару отримує всі перерізи певного АЧЧС для моментів часу, коли віброзміщення цього вібросигналу перевищує припустиме значення, а також дані про можливу залежність певних смуг частот відповідної АЧЧС від гідродинамічних та електродинамічних чинників і фонові спектральні характеристики.

Кожен нейрон третього шару визначає рівень вірогідності того, що причиною надмірного віброзміщення певного вібросигналу є характерний чинник вібрації, якій відповідає цьому нейрону.

Очевидно, що ця ШНМ спрацьовує лише у випадку, коли хоча б в одному з вібросигналів міститься надмірне віброзміщення.

Слід відзначити, що можливі ситуації, коли одночасно декілька різних характерних чинників вібрації отримають високі рівні вірогідності. В цьому випадку ШНМ видасть діагноз про необхідність перевірки під час чергового або позачергового технічного огляду одразу декількох дефектів гідроагрегату.

Математична модель, алгоритм і програмна реалізація ШНМ САДП-РДГ для діагностування існуючих дефектів гідроагрегату

Насамперед, відзначимо, що всі процедури цього розділу виконуються з абсолютними значеннями вейвлет-коефіцієнтів.

Як вже було сказано в попередньому підрозділі нейрони першого шару мають виконувати три функції:

- визначати фонові значення вейвлет-коефіцієнтів кожної смуги частот;
- визначати залежність вейвлет-коефіцієнтів кожної смуги частот від струму навантаження гідроагрегату;

- визначати залежність вейвлет-коефіцієнтів кожної смуги частот від рівня води у водосховищі.

Відомо, що рівень турбулентності обернено пропорційний рівню води у водосховищі, тобто ця залежність має гіперболічний характер.

Що ж стосується залежності рівня вібрації від струму навантаження то вона має прямо пропорційний характер.

Звичайно, в процесі дослідної експлуатації САДП-РДГ характер цих залежностей може уточнюватися, але, апіорі, для задачі діагностування можна записати:

$$|d_j(H, I)| = D_{0j} + D_j + v_j I + \frac{1}{p_j + q_j H^2}, \quad (1)$$

де H – рівень води у водосховищі;

I – струм навантаження гідрогенератора;

$d_j(H, I)$ – вектор-функція залежності вейвлет-коефіцієнтів j -ої смуги частот від рівня води у водосховищі та струму навантаження;

D_{0j} – усереднене фонове значення масиву вейвлет-коефіцієнтів j -ої смуги частот;

D_j – усереднене значення масиву вейвлет-коефіцієнтів j -ої смуги частот при максимальному рівні води у водосховищі та режимі холостого ходу гідрогенератора;

v_j – узагальнений числовий коефіцієнт, що характеризує залежність вейвлет-коефіцієнтів j -ої смуги частот від струму навантаження;

p_j, q_j – узагальнені числові коефіцієнти, що характеризують залежність вейвлет-коефіцієнтів j -ої смуги частот від рівня води у водосховищі.

Ця залежність визначається лише для тих смуг частот, для яких вейвлет-коефіцієнти зростають при падінні рівня води чи при зростанні струму навантаження. Для тих смуг частот, вейвлет-коефіцієнти яких ніяк не реагують ні на рівень води, ні на струм з АЧЧС виділяється лише D_{0j} .

Отже, кожен нейрон першого шару передає далі в ШНМ, окрім масиву вейвлет-коефіцієнтів j -ої полоси частот, ще й параметри $H, I, D_{0j}, D_j, v_j, p_j, q_j$.

Процедура, яку виконує кожен нейрон другого шару полягає в тому, що для кожного моменту часу τ , для якого віброзміщення перевищує припустиме значення, визначається переріз АЧЧС відповідного вібросигналу. Для цього спочатку в кожній смузі частот j знаходяться інтервали часу $[t_{j,n} - t_{j,n+1}]$, які відповідають умові:

$$t_{j,n} < \tau < t_{j,n+1}. \quad (2)$$

Після цього для всіх $j = 1, M$ знаходять переріз $S_\tau(M)$, який і передається далі в ШНМ.

Таким чином, на кожен нейрон третього шару надходять вектори $S_\tau(M)$ кожного вібросигналу і параметри $H, I, D_{0j}, D_j, v_j, p_j, q_j$ кожної смуги АЧЧС кожного вібросигналу.

Нагадаємо, що кожний нейрон третього шару відповідає за діагностування певного чинника виникнення вібрації. При формуванні нейронної мережі в кожен нейрон третього шару введена апріорна експертна інформація про перелік номерів смуг частот, на які може впливати чинник цього нейрону.

Для зручності, додатково введемо у вищенаведені параметри індекс $i = 1...4$, який відповідає кожному з 4-ох вібросигналів від сенсорів ВД1 – ВД4, та індекс $k = 1...6$, який відповідає кожному з нейронів третього шару ШНМ.

Тоді, кожен вектор $S_{i\tau}(M)$ містить M елементів $d_{ij\tau} \in S_{i\tau}(M)$.

Далі кожен нейрон має виконати такі процедури:

1. Для кожної смуги частот кожного АЧЧС відокремити фонову, гідродинамічну та електродинамічну складові від складових, які викликані іншими чинниками.

Для перших чотирьох нейронів, які характеризують безпосередньо механічні дефекти гідроагрегату це виконується за формулою

$$\forall k = 1, 4 \forall i = 1, 4 \forall j = 1, M \left(d_{kij\tau}^* = |d_{ij\tau}| - D_{0ij} - v_j I - \frac{1}{p_{ij} + q_{ij} H^2} \right). \quad (3)$$

Для п'ятого нейрона, який повинен діагностувати порушення гідродинаміки потоку можна записати

$$\forall i = 1, 4 \forall j = 1, M \left(d_{5ij\tau}^* = \frac{1}{p_{ij} + q_{ij} H^2} \right). \quad (4)$$

Для шостого нейрона, який повинен діагностувати електродинамічну складову вібрації можна записати

$$\forall i = 1, 4 \forall j = 1, M \left(d_{6ij\tau}^* = v_j I \right). \quad (5)$$

2. Наступним кроком є нормування всіх елементів $d_{kij\tau}^*$, яке виконується з формулою

$$\forall k = 1, 6 \forall i = 1, 4 \forall j = 1, M \left(d_{kij\tau}^{norm} = \frac{d_{kij\tau}^*}{\sum_{k=1}^6 \sum_{i=1}^4 \sum_{j=1}^M d_{kij\tau}^*} \right). \quad (6)$$

3. Для визначення абсолютного показника вірогідності того, що причиною надмірного віброзміщення є характерний чинник вібрації, якій відповідає k -му нейрону, використовують його множину номерів смуг частот Q_k .

Тоді показник вірогідності $PV_{k\tau}$ чинника, який відповідає k -му нейрону, на момент часу τ визначається як

$$\forall k = 1, 6 \forall i = 1, 4 \forall j \in Q_k \left(PV_{k\tau} = \sum_{i,j} w_{ki} d_{kij\tau}^{norm} \right), \quad (7)$$

де w_{ki} – вагові коефіцієнти, які визначають важливість урахування вейвлет-коефіцієнтів смуг частот АЧЧС i -го вібросигналу в рівні вірогідності k -го нейрону.

Апріорі, вагові коефіцієнти визначаються так:

- для першого, другого, п'ятого та шостого нейронів вони приймаються однаковими і дорівнюють 1, тобто

$$\forall i = 1, 4 \wedge (k = 1 \vee k = 2 \vee k = 5 \vee k = 6) \Rightarrow (w_{ki} = 1); \quad (8)$$

- для третього нейрона (дефекти турбінного підшипника) пріоритетним доцільно вважати АЧЧС вібросигналів, отриманих від вібросенсорів ВД1 та ВД2 (встановлених на турбінному підшипнику), тому

$$(i = 1 \vee k = 2) \wedge (k = 3) \Rightarrow (w_{ki} = 1), \quad (9)$$

а для АЧЧС вібросигналів, отриманих від вібросенсорів ВД3 та ВД4 (встановлених на опорно-упорному підшипнику) доцільно ввести знижені вагові коефіцієнти (наприклад, 0.5), тобто

$$(i = 3 \vee k = 4) \wedge (k = 3) \Rightarrow (w_{ki} = 0.5); \quad (10)$$

- аналогічно, для четвертого нейрона можна записати

$$(i = 3 \vee k = 4) \wedge (k = 4) \Rightarrow (w_{ki} = 1), \quad (11)$$

$$(i = 1 \vee k = 2) \wedge (k = 4) \Rightarrow (w_{ki} = 0.5). \quad (12)$$

4. В процесі експлуатації САДП-РДГ випадки перевищення припустимих віброзміщень будуть час від часу повторюватися, тому передбачено накопичення та усереднення отриманих показників вірогідності кожного нейрона, тобто за певний проміжок часу t

$$PV_{kt} = \frac{\sum PV_{k\tau}}{N_t}, \quad (13)$$

де N_t – кількість випадків перевищення припустимого віброзміщення за час t .

5. Таким чином, можна сформуувати множину показників вірогідності всіх нейронів третього шару ШНМ $\{PV_{1t}, PV_{2t}, PV_{3t}, PV_{4t}, PV_{5t}, PV_{6t}\}$.

Аналізуючи вищенаведені вирази неважко помітити, що абсолютні значення показників вірогідності будуть досить малими і це викличе певні незручності при їх аналізі. Тому доцільно ввести відносний рівень вірогідності RV_{kt} за формулою:

$$RV_{kt} = \frac{PV_{kt}}{\max(PV_{1t}, PV_{2t}, PV_{3t}, PV_{4t}, PV_{5t}, PV_{6t})}, \quad (14)$$

6. Нарешті, ми можемо перейти до формулювання діагностичного висновку. При цьому слід пам'ятати, що причиною вібрації є одночасна дія усіх без винятку чинників, тому мова може йти лише про встановлення деякої ієрархії небезпеки цих чинників для гідроагрегату в поточний момент часу.

Відзначимо, що рівень вірогідності RV_{kt} є, хоча і дуже важливою, але не єдиною діагностичною ознакою, які враховуються при формулюванні діагностичного висновку.

Так, наприклад, якщо в якийсь момент часу τ , неприпустимі значення віброзміщення містяться лише у вібросигналах, отриманих від ВД1 та ВД2, а у двох інших вібросигналах вони цілком припустимі, то можна зробити висновок про переважаючу небезпеку дефектів саме турбінного підшипника і прийняти для нього $RV_{kt} = 1$.

У випадку, якщо неприпустимі значення віброзміщення містяться лише у вібросигналах, отриманих від ВД3 та ВД4, а у двох інших вібросигналах вони цілком припустимі, то можна зробити висновок про переважаючу небезпеку дефектів саме опорно-упорного підшипника і прийняти для нього $RV_{kt} = 1$.

Якщо ж неприпустимі значення віброзміщення містяться в усіх вібросигналах, то ситуація стає неоднозначною і потребує більш детального аналізу. Розглянемо це на прикладах, отриманих за допомогою програми, реалізованої за алгоритмом, наведеним вище.

Формування попередніх діагностичних висновків

На рис. 2 зображений переріз АЧЧС (визначеним за допомогою ДВП з $k = 2$) вібросигналу від ВД1, який відповідає надмірному віброзміщенню в момент часу τ_1 .

З цього перерізу видно різкий стрибок спектральних характеристик 6-ої та 7-ої смуг частот. Згідно експертній інформації, це свідчить про переважаючу небезпеку невірноваженості ротора, і тому для цього чинника можна встановити $RV = 1$.

Більш складною є ситуація, зображена на рис. 3 (момент часу τ_2).

Тут помітні стрибки на 5-ій, 6-ій і 9-ій смугах частот. Це вказує на гідродинамічні та електродинамічні чинники, а також на порушення жорсткості опорної системи і на дефекти підшипників.

Цікавим є і наступний переріз АЧЧС вібросигналу від ВД2, зображений на рис. 4 (момент часу τ_3). На ньому можна спостерігати дуже різкий стрибок на п'ятій смузі частот, що свідчить про субгармонійну спектральну складову, яка виникає при порушеннях гідродинаміки потоку та порушеннях жорсткості опорної системи.

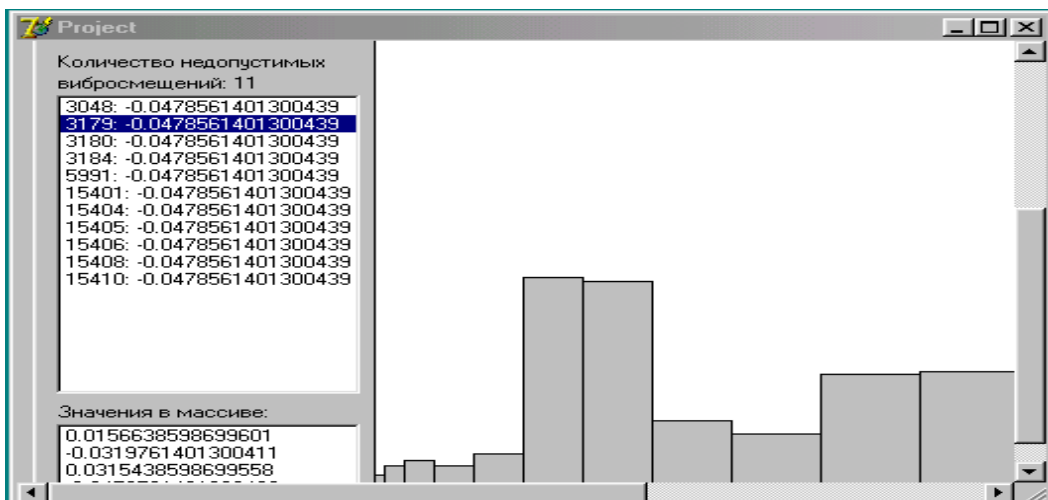


Рисунок 2 – Переріз АЧЧС вібросигналу від ВД1 в момент часу τ_1

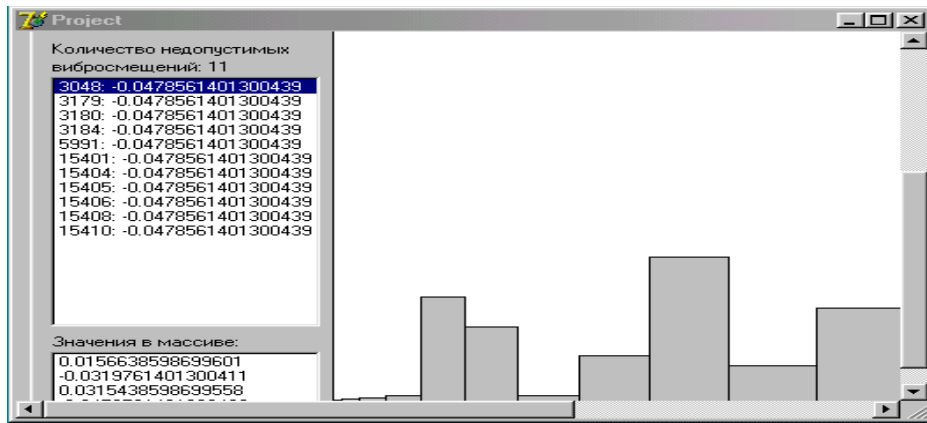


Рисунок 3 – Переріз АЧЧС віброцигналу від ВД1 в момент часу τ_2

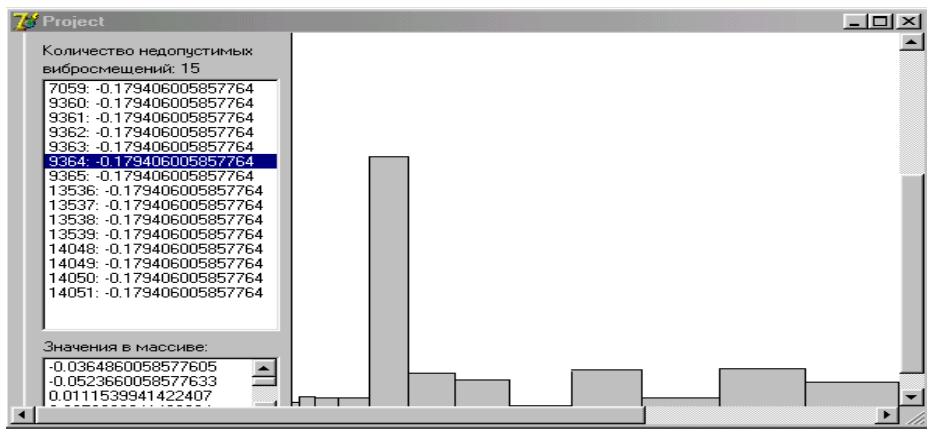


Рисунок 4 – Переріз АЧЧС віброцигналу від ВД2 в момент часу τ_3

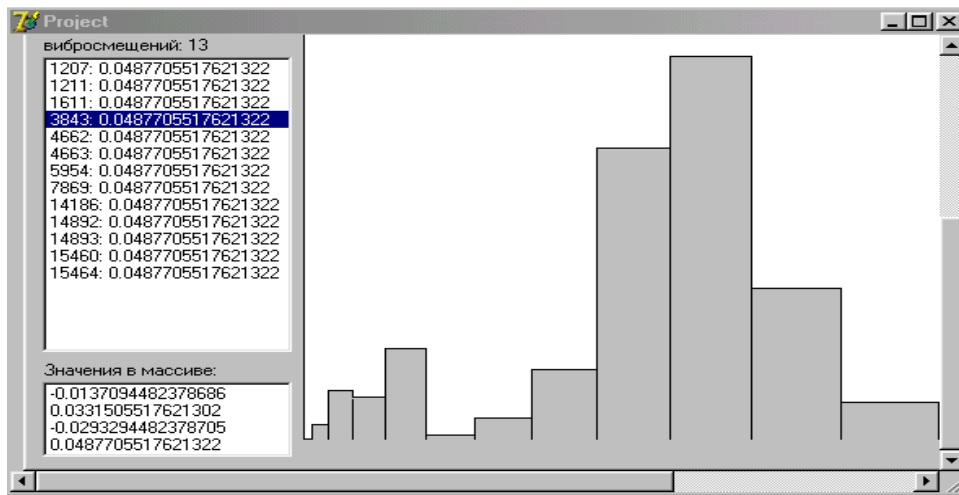
На рис. 5 показаний переріз АЧЧС віброцигналу від ВД3, який встановлений на опорно-упорному підшипнику (момент часу τ_4).

На ньому ми бачимо найбільші спектральні складові на 9-ій та 10-ій смугах. В цих смугах містяться лопаткові частоти і частоти пов'язані з кавітацією, тобто тут явно має місце порушення гідродинаміки потоку.

Таким чином, результуючий діагностичний висновок може бути сформульований як множина значень рівнів вірогідності різних чинників вібрації $\{RV_{1t}, RV_{2t}, RV_{3t}, RV_{4t}, RV_{5t}, RV_{6t}\}$.

Зупинимося окремо на проблемі настроювання ШНМ.

Кожен гідроагрегат є виключно складною електрогідромеханічною системою і його вібраційні характеристики містять багато індивідуальних особливостей. Це означає, що експертне визначення його множин номерів смуг частот Q_k і вагових коефіцієнтів w_{ki} до початку дослідної експлуатації САДП-РДГ може бути лише дуже приблизним.

Рисунок 5 – Переріз АЧЧС вібро сигналу від ВДЗ в момент часу τ_4

Тому представляється доцільною така стратегія настроювання:

- протягом дослідної експлуатації САДП-РДГ щодобово архівується інформація про множини значень рівнів вірогідності різних чинників вібрації $\{RV_{1t}, RV_{2t}, RV_{3t}, RV_{4t}, RV_{5t}, RV_{6t}\}$ і про перерізи $S_{\tau}(M)$, що відповідають усім випадкам надмірного віброзміщення;
- під час кожного чергового або позачергового технічного огляду гідроагрегату відповідальними особами обслуговуючого персоналу ГЕС (які в цьому випадку виконують функцію експертів) визначається рівень зношеності основних вузлів гідроагрегату;
- далі особа, яка має відповідати за функціонування САДП-РДГ, порівнює висновки експертів щодо наявних дефектів гідроагрегату з архівними даними рівнів вірогідності різних чинників вібрації $\{RV_{1t}, RV_{2t}, RV_{3t}, RV_{4t}, RV_{5t}, RV_{6t}\}$ та перерізами $S_{\tau}(M)$, і приймає рішення щодо уточнення множин номерів смуг частот Q_k і вагових коефіцієнтів w_{ki} .

Висновки

1. Для діагностування дефектів гідроагрегату запропонована тришарова ШНМ, основною вхідною інформацією для якої є АЧЧС віброакустичних сигналів гідроагрегату.
2. Виключна складність гідроагрегату як динамічної гідроелектромеханічної системи обумовлює значну невизначеність апріорних експертних оцінок щодо залежності АЧЧС віброакустичних сигналів від чинників, які викликають вібрацію.
3. Тому процедури настроювання ШНМ повинні проводитися протягом усього часу дослідної експлуатації САДП-РДГ.

Список літератури

1. Ширман А.Р. Практическая вибродиагностика и мониторинг состояния механического оборудования / Ширман А.Р., Соловьев А.Б. – Москва, 1996. – 276 с.
2. Барков А.В. Мониторинг и диагностика роторных машин по вибрации / Барков А. В., Баркова Н. А., Азовцев А. Ю. – Л.: АО ВАСГ, 1997. – 170 с.
3. Пакет прикладных программ для диагностирования и прогнозирования состояния подшипников качения. Инструкция по эксплуатации. Части 3,4. Изд. АО ВАСГ, г. Санкт-Петербург, 1992г.
4. Воронцов О.Г. Розробка методу віброакустичної діагностики тихохідних роторних машин / Воронцов О.Г., Дегтяренко І. В. // Наукові праці Донецького державного технічного університету. Серія: Обчислювальна техніка та автоматизація. – Донецьк: РВА ДонДТУ – 2002 – Вип. 41. – с.212-215.
5. Воронцов О.Г. Метод повышения достоверности вибродиагностики роторных машин по комплексу диагностических параметров / Воронцов О.Г. // Наукові праці Донецького державного технічного університету. Серія: Електротехніка і енергетика. – Донецьк: РВА ДонДТУ – 2002 – Вип. 38. – с.250-256.

6. Кухарчук В.В. Система автоматизованого діагностування і прогнозування розвитку дефектів гідрогенераторів / Кухарчук В.В., Каців С.Ш., Говор І.К., Мадьяров В.Г., Ніколаєв В.Я., В.Л. Маліцький // Вісник Інженерної Академії України. – 2009 – №2. – с.126-132.
7. Кухарчук В.В. Аналіз вібросигналів гідроагрегату за допомогою дискретного вейвлет-перетворення з коефіцієнтом стискання 2 / В. В. Кухарчук, С. Ш. Каців, І. К. Говор, С. О. Биковський // Вісник Інженерної Академії України. – 2011. – №1. – с. 124-129.
8. Кухарчук В.В. Дослідження результатів дискретного вейвлет-перетворення вібросигналу при коефіцієнті стискання 2 та різних материнських вейвлетах / В. В. Кухарчук, С. Ш. Каців, В. Г. Мадьяров, С. О. Биковський // Вісник Інженерної Академії України. – 2011. – №2. – с. 66-69.
9. Кухарчук В.В. Порівняльний аналіз результатів дискретного вейвлет-перетворення вібросигналів з довільними цілими коефіцієнтами стискання / В. В. Кухарчук, С. Ш. Каців, І. К. Говор, С. О. Биковський // Вісник Інженерної Академії України. – 2011. – №2. – с. 172-176.
10. Тарасов А.А. Физические процессы в гидрогенераторах с реактивной турбиной [Электронный ресурс] / Тарасов А.А. // Режим доступа : <http://diamech.ru/notes03112009.html>
11. Осовский С. Нейронные сети для обработки информации / Осовский С., пер. с польск. – М.: Финансы и статистика, 2002. – 344 с., ил. – ISBN 5-279-02567-4.
12. Зажицкий А. В. Классификация состояния объекта виброакустической диагностики с использованием нейротехнических структур / Зажицкий А. В., Бурау Н. И., Тяпченко А. Н. // *Авіаційно-космічна техніка і технологія*. – Харків, 2002. – Вип. 31. – с. 181-185.
13. Зажицкий О. Аналіз ефективності розпізнавання технічного стану об'єкту віброакустичної діагностики нейронними мережами / Олександр Зажицький, Надія Бурау // *Методи та прилади контролю якості*. – 2003. – № 11. – с. 21-26.

Відомості про авторів

Каців Самоїл Шулімович – к. т. н., доц., доцент кафедри теоретичної електротехніки та електричних вимірювань, Вінницький національний технічний університет, тел. (0432) 598444, kaciv@ineem.vntu.edu.ua, Хмельницьке шосе, 95, м. Вінниця, 21021.

УДК 681.3.07

А.І.КАРДАШ, С.М.ЛЕВИЦЬКА, А.Т.ДУДИКЕВИЧ

Львівський національний університет імені Івана Франка, Львів

ЗАДАЧА РОЗПІЗНАВАННЯ ЛЮДСЬКИХ ОБЛИЧ МЕТОДАМИ ШТУЧНОГО ІНТЕЛЕКТУ

Анотація. Об'єктом даного дослідження стало виявлення людських облич на цифрових зображеннях. Ми ознайомилися з різними підходами до розв'язання таких проблем, а також реалізували та проаналізували роботу одного з власних підходів.

Ключові слова: розпізнавання облич, метод Віоли-Джонса, інтегральне представлення зображень, скануюче вікно, ознаки Гаара, бустинг.

Аннотация. Объектом данного исследования стало выявление человеческих лиц на цифровых изображениях. Мы ознакомились с разными подходами к решению таких проблем, а также реализовали и проанализировали работу одного из собственных подходов.

Ключевые слова: распознавание лиц, метод Виолы-Джонса, интегральное представление изображений, сканированное окно, признак Гаара, бустинг.

Abstract. The object of the face recognition on digital images. The authors have studied different approaches to solving such problems and applied and analyzed the work of one of their own approaches.

Key words: face recognition, Viola - Jones method, integral representation of images, scanning window, Haar features, boosting.

Вступ

Розпізнавання образів є популярною проблемою сучасної інформатики. Образи можуть бути як візуальні так і звукові чи інші, залежно від типів інформації, з якими має працювати система. Багато задач є дуже складним для комп'ютера (хоча й елементарними для людей з використанням вроджених і набутих умінь). Саме тому часто воно відбувається з використанням засобів штучного інтелекту. Оскільки людині (розробнику системи чи науковцю) важко проаналізувати проблемну область і знайти якісь чіткі правила або риси, які однозначно визначатимуть шуканий об'єкт серед шуму, то в більшості випадків використовується навчання з учителем. Будеться система, яка здатна аналізувати дані з предметної області.

Дуже важливим напрямком у розпізнаванні образів є розпізнавання зображень. Оскільки найбільшу частку інформації людина сприймає через зір, то природно, що інтелектуальна система буде певним чином імітувати вміння людини опрацювати власне зображення (або відео). Для того, щоб проаналізувати об'єкт, потрібно спочатку знайти його на зображенні, тобто проглянути зображення та розпізнати (виявити) шукані об'єкти.

Розпізнавання зображень особливо важливе для такого полідисциплінарного напрямку, як взаємодія людини з комп'ютером (Human-computer Interaction), над яким активно працює багато корпорацій, які виготовляють цифрову техніку та розробляють програмне забезпечення. Найбільше розробляються системи з наступними задачами розпізнавання: виявлення об'єктів на зображенні (наприклад пішоходів чи автомобілів на дорозі – очевидне застосування у розробці автоматично керованих автомобілів або при аналізі відео з дорожніх камер), виявлення облич (ця функція присутня майже у всіх сучасних фотоапаратах, допомагає в налаштуванні автофокусу, одна з перших дій при отриманні характеристик для розпізнавання емоцій та міміки лица), розпізнавання жестів (для керування комп'ютером без використання мишки). Все це можна застосувати при створенні людиноподібних роботів-помічників для розробки зручних засобів керування ними [1-3].

У даній роботі розглядається розпізнавання людських облич на зображенні (найчастіше це фотографії). Застосуванням розробленої системи може бути, як робота з фотографіями у складі інших програм для виявлення облич на фотографіях і використання інформації про їх розташування в завданнях, які ставляться перед цими програмами (це може бути ідентифікація облич або виділення лиця з тла у відеорозмовах для зменшення обсягу передавання даних), так і виділення облич для подальшого розпізнавання емоцій на них.

Мета та постановка задачі

Метою роботи було проаналізувати існуючі підходи до розв'язання цієї проблеми та запрограмувати і проаналізувати роботу одного з найкращих методів, а також створити систему, яка знаходить людські обличчя на зображенні. Було розглянуто метод Віоли-Джонса [2,3]. Для роботи алгоритму використовується інтегральне представлення зображення, яке дозволяє швидко порахувати сумарну яскравість ділянки на зображенні. При цьому воно має таку ж розмірність, як і висхідне, але замість кожного пікселя вказана сума яскравостей всіх пікселів, розташованих ліворуч та вище.

Як прості класифікатори, які використовуються для побудови та розпізнавання облич складного класифікатора, взято ознаки Гаара, які описують риси зображення, яке потрібно знайти. Прямокутні ознаки Гаара добре працюють з інтегральним представлення зображень, оскільки для їх обчислення достатньо додати кілька елементів матриці зображення.

Для побудови дієвого класифікатора використовується навчальна вибірка прикладів з обличчями та без них. Для аналізу роботи методу Віоли-Джонса випробувано два типи класифікатора: повільний, по-

будований процедурою AdaBoost та швидкий каскадний, який побудований ітеративним алгоритмом, який використовує AdaBoost для навчання каскадів, з яких складається.

Методи та алгоритми досліджень

Метод Віоли-Джонса був вперше запропонований у 2001 році і завдяки своїй швидкодії швидко став популярним. У літературі можна знайти багато спроб його удосконалення, проте більшість з них спрямована на модифікацію бустингу, тим не менше, такі основні риси, як інтегральне представлення зображень та використання ознак Гаара залишаються незмінними, хіба що спеціально підбираються конкретні ознаки з наміром показати покращення роботи побудованого з них класифікатора.

Однією з сильних сторін підходу є те, що він може бути застосований до пошуку різних об'єктів на зображеннях [4]. Прямокутні (і не тільки) ознаки можна обчислити дуже швидко, зі складністю $\Theta(1)$, використовуючи проміжне представлення зображення в інтегральній формі. Таке зображення у точці (x, y) , на відміну від вихідного зображення, містить не значення пікселя, а суму всіх пікселів, що розташовані не нижче та не правіше від точки (x, y) :

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y'),$$

тут $ii(x, y)$ – це значення пікселя в інтегральному представленні, а $i(x', y')$ – значення пікселя у вихідному зображенні. Можна використовувати наступні рекурентні формули:

$$s(x, y) = s(x, y-1) + i(x, y),$$

$$ii(x, y) = ii(x-1, y) + s(x, y).$$

Тут $s(x, y)$ є кумулятивною сумою стовпчика. Умови $s(x, -1) = 0$ та $ii(-1, y) = 0$ роблять обчислення скінченними. Таким чином, інтегральне представлення зображення будується за один прохід через вихідне зображення.

Використовуючи інтегральне представлення, сума будь якої прямокутної області може бути обчислена через 4 звертання до масиву. Очевидно, що різниця між сумами двох прямокутних ділянок обчислена за 8 звертань. А ознаки, які складаються з двох суміжних прямокутників, можуть бути обчислені за 6 звертань до масиву, у випадку трьох суміжних прямокутників – за 8, чотирьох – за 9. Завдяки чому, ознаки Гаара знаходяться швидко.

Математично, інтегральне представлення є подвійним інтегралом функції, яка описує зображення (спочатку по рядках, а потім по стовпцях).

Неважко побачити, що маючи розширення скануючого вікна 24×24 , повна множина різних прямокутних ознак є дуже великою: 45396 елементів. Зауважимо що на відміну від базису Гаара, така множина не буде лінійно незалежною. Тут розмірність підпростору буде 576, що набагато менше ніж в [3].

Фройнд і Шапір [5] довели, що навчальна помилка сильного класифікатора прямує до нуля експоненційно по кількості ітерацій. Ще більш важливо, в інших дослідженнях згодом було продемонстровано та обгрунтовано добрі узагальнюючі здатності даного підходу. Ключовим моментом тут є те, що узагальнююча здатність залежить від границі для набору прикладів, а AdaBoost досягає великих границь досить швидко.

На рисунку 1 показано прямокутні ознаки відносно скануючого вікна. Ознаки з двох прямокутників наведені в (A) і (B). Малюнок (C) показує ознаку з трьох прямокутників, і (D) – ознака з чотирьох прямокутників. Справа – решта ознак.

Звичайна процедура AdaBoost може бути легко інтерпретована як жадібний процес вибору ознак. Розглянемо загальну задачу бустингу, в якій великий набір функцій класифікації об'єднуються за допомогою функції зваженого голосування. Проблема полягає в тому, щоб пов'язати великі ваги з добрими функціями класифікації, а малі з поганими. AdaBoost – це агресивний механізм відбору невеликого набору гарних функцій класифікації, які, тим не менш, сильно відрізняються між собою. Проводячи аналогію між слабкими класифікаторами та ознаками, AdaBoost є ефективною процедурою для пошуку з невеликого числа "хороших", які тим не менш, мають значне розмаїття.

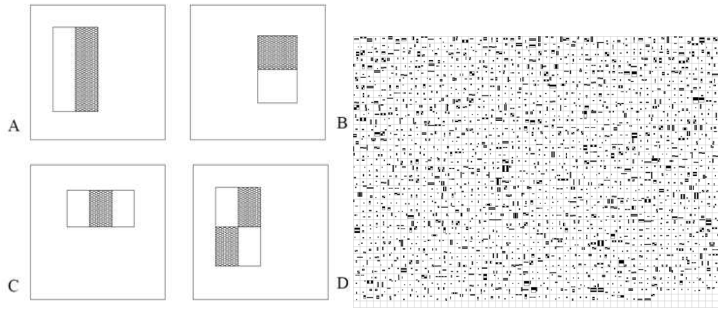


Рисунок 1– Приклади прямокутних ознак

Розглянемо типовий алгоритм навчання із застосуванням бустингу. Будуються гіпотези, кожна з яких використовує одну ознаку. Остаточна гіпотеза є зваженою лінійною комбінацією з гіпотез, де ваги обернено пропорційні помилкам навчання [3, 6].

1. Нехай дано навчальні приклади зображень. Подамо їх парами $(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)$, де $y_1 = 0, 1$ для негативних та позитивних прикладів відповідно.
2. Ініціалізуємо ваги наступним чином: $w_{1,i} = 1/(2m)$ або $w_{1,i} = 1/(2l)$ для $y_1 = 0, 1$ відповідно. Тут m та l – це кількість негативних та позитивних прикладів у навчальній множині.
3. Для кожної гіпотези $t = 1, \dots, T$ виконати наступне:

- a. Нормалізувати ваги: $w_{t,i} = w_{t,i} / \sum_{j=1}^n w_{t,j}$.

- b. Для кожної ознаки j навчати класифікатор h_j , який обмежений лише одною ознакою.

Помилка оцінюється для w_t $e_j = \sum_{i=1}^n w_j |h_j(x_i) - y_i|$.

- c. Вибрати класифікатор h_t з найменшим значенням помилки e_t .

- d. Оновити ваги: $w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$, де $\beta_t = e_t / (1 - e_t)$.

4. Остаточний сильний класифікатор матиме вигляд:

$$h(x) = \begin{cases} 1, & \text{якщо } \sum_{t=1}^T a_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T a_t, \\ 0, & \text{інакше} \end{cases}$$

де $a_t = \log(1/\beta_t)$.

Висновки

Результатом виконаної роботи є створення працюючої системи для розпізнавання облич на зображенні з використання методу Віоли-Джонса та з модифікацією методу сканування зображення.

Отримана система працює швидко та може опрацювати близько 15 зображень розміром 400×300 пікселів за секунду на звичайному комп'ютері.

Як результат проведеного дослідження з двома способами побудови класифікатора напрашується висновок, що адаптивні каскадні класифікатори є кращими для розпізнавання облич, оскільки майже без втрат в ефективності досягається набагато вища швидкість роботи. Також вони є кращими порівняно із

перцептронами, бо в процесі їх роботи задіюються лише ті частини системи, які виконують корисну роботу.

Побудована система може слугувати основою для використання у складніших задачах аналізу обличчя, як перший етап знаходження обличчя, а зважаючи на швидкодію, може працювати з відео.

На перспективу можна, використовуючи такі ж алгоритми побудови класифікаторів, побудувати систему для виділення частин лица та аналізу емоцій.

Список літератури

1. Стюарт Рассел. Искусственный интеллект: современный подход, 2-е изд./ Стюарт Рассел, Питер Норвіг; пер.с англ. – М., Вільямс, 2006 - 1408 с.
2. P. Viola. Rapid Object Detection using a Boosted Cascade of Simple Features,/ P. Viola and M.J. Jones, proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2001), 2001.
3. P. Viola. Robust real-time face detection,/ P. Viola and M.J. Jones, International Journal of Computer Vision, vol. 57, no. 2, 2004., pp.137–154.
4. Stan Z. Li (Editor). Handbook of Face Recognition. Springer,/ Stan Z. Li (Editor), Anil K. Jain (Editor), 2004.
5. Yoav Freund. A decision-theoretic generalization of online learning and an application to boosting./ Yoav Freund and Robert E. Schapire. In Computational Learning Theory: Eurocolt '95, pages 23–37. Springer-Verlag, 1995.
6. Y.Freund. A short introduction to boosting./ Y.Freund, R.E.Schapire. Journal of Japanese Society for Artificial Intelligence 14, 1999 , pp. 771–780.

Відомості про авторів

Кардаш Андрій Іванович – кандидат фіз.- мат. наук, доцент кафедри програмування, Львівський національний університет імені Івана Франка.

Левицька Софія Михайлівна – старший викладач кафедри програмування, Львівський національний університет імені Івана Франка.

Дудикевич Анна Теодорівна – кандидат фіз.- мат. наук, доцент кафедри обчислювальної математики, Львівський національний університет імені Івана Франка.

УДК 681.31

Б.М. ШЕВЧУК

Інститут кібернетики ім. В.М. Глушкова НАН України, м. Київ

СИСТЕМНИЙ ПІДХІД ДО ВИРІШЕННЯ ПРОБЛЕМ ОПТИМІЗАЦІЇ ОБЧИСЛЕНЬ ЗАСОБАМИ ОБ'ЄКТНИХ СИСТЕМ СЕНСОРНИХ МЕРЕЖ

Анотація. З урахуванням комплексу завдань, пов'язаних з обробленням і кодуванням даних моніторингу та формуванням компактних і захищених пакетів інформації засобами об'єктних систем сенсорних мереж запропонована реалізація оперативних обчислень в процесі стиску сигналів і відеосигналів з допустимими втратами інформації, стиску-захисту даних без втрат, оперативного визначення інформаційних станів об'єктів тривалого моніторингу. Запропоновані методи оперативного завадостійкого кодування даних та формування компактних кодово-сигнальних послідовностей пакетів інформації.

Ключові слова: об'єктні системи, сенсорні мережі, стиск сигналів і даних, обчислення, сигнально-кодові послідовності.

Аннотация. С учетом комплекса задач, связанных с обработкой и кодированием данных мониторинга, а также формированием компактных и защищенных пакетов информации средствами объектных систем сенсорных сетей предложена реализация оперативных вычислений в процессе сжатия сигналов и видеосигналов с допустимыми потерями информации, сжатия-защиты данных без потерь, оперативного определения информационных состояний объектов длительного мониторинга. Предложены методы оперативного помехоустойчивого кодирования данных и формирования компактных сигнально-кодовых последовательностей пакетов информации.

Ключевые слова: объектная система, сенсорные сети, сжатие сигналов и данных, вычисления, сигнально-кодовые последовательности.

The Abstract. Taking into account a complex of the tasks connected with processing and coding of data of monitoring, and also formation of the compact and protected packages of information by means of object systems of touch networks offered realization of expeditious calculations in the course of compression of signals and video signals with admissible losses of information, compression - data security without loss, expeditious definition of information conditions of objects of long monitoring. Methods of expeditious noiseproof coding of data and formation of compact alarm and code sequences of packages of information are offered.

Keywords: object system, touch networks, compression of signals and data, calculations, alarm and code sequences.

Вступ

Сучасні сенсорні мережі (СМ) із самоорганізацією передачі пакетів інформації є основою для побудови "всюди проникаючих мереж". При цьому на об'єктах контролю різноманітної природи (промислових і технічних об'єктах, мобільних роботах, об'єктах екомоніторингу, операторах людино-машинних комплексах, пацієнтах, параметри яких контролюються телемедициними системами та ін.) встановлюються об'єктні системи СМ, які вирішують комплекс завдань [1-3], пов'язаних з введенням, обробленням, кодуванням моніторингової інформації та передаванням пакетів даних до засобів міжмережевої взаємодії. З метою підвищення інформаційної ефективності функціонування СМ в ISM діапазоні частот (ISM – Industrial, Scientific, Medical: 433, 868, 902-928 (для США), 2400МГц) в [2,3] запропоновані методи та алгоритми обробки і кодування даних моніторингу засобами об'єктних систем (ОС). Об'єктні системи СМ, як правило, орієнтовані на тривалий час роботи від автономного джерела живлення та характеризуються використанням мікропотужних процесорів (ARM-мікроконтролерів, сигнальних процесорів) з обмеженою обчислювальною продуктивністю. Суттєвим недоліком сучасних СМ є зменшення, приблизно на порядок, швидкості передачі даних в коміркових мережах. Тому серед невирішених проблем ефективного застосування СМ в різноманітних галузях людської діяльності доцільно виділити орієнтацію СМ на передачу низькошвидкісних вимірювальних сигналів, низьку захищеність передачі даних з використанням відкритих каналів зв'язку, включаючи низький криптозахист даних, який базується на принципах симетричної криптографії, та низький захист даних від каналних завад, пов'язаний з обмеженою базою каналних сигналів та відсутністю ефективного завадостійкого кодування даних. Також суттєвим недоліком засобів СМ є той факт, що вони не орієнтовані на передачу відеоданих.

Актуальність

З урахуванням обмеження на частотні ресурси для функціонування СМ актуальним завданням є розміщення в місцях виникнення (зародження) інформаційних потоків об'єктних систем СМ, які, при використанні процесорів з обмеженою продуктивністю, з мінімальними обчислювальними витратами забезпечують формування та передачу компактних та захищених (криптостійких і завадостійких) пакетів інформації різноманітного характеру.

Мета

З метою підвищення швидкості передачі пакетів інформації в СМ з урахуванням системних вимог до засобів об'єктних систем, які орієнтовані на реалізацію комплексу алгоритмів оброблення, кодування і шифрування вхідних потоків моніторингових даних та формування компактних і захищених пакетів інформації, в статті запропоновані взаємодоповнюючі алгоритми оперативного опрацювання даних на інформаційному рівні. Для реалізації засобами ОС відповідної операції запропонована послідовність виконання обчислень, оптимізованих за швидкістю і точністю кодування даних. При цьому вхідними моніторинговими даними, що підлягають обробці і кодуванню, можуть бути вимірювальні сигнали,

відеосигнали (дані рухомих і нерухомих зображень), масиви даних.

Методологічні та алгоритмічні основи оптимізації обчислень в процесі оброблення, кодування та передавання даних в СМ.

Для ефективного застосування СМ в необхідних місцях об'єктів тривалого моніторингу встановлюються ОС, для яких вхідними даними є аналогові сигнали, відеосигнали або масиви моніторингових даних, отриманих з виходів датчиків, сенсорів та відеосенсорів. Слід зазначити, що для широкого застосування СМ необхідно орієнтувати засоби ОС на передачу як вимірювальних сигналів так і відеоданих. Як правило, моніторингові дані накопичуються в буферному накопичувачі ОС, і в залежності від прикладної орієнтації СМ, отримані дані в пакетному режимі, згідно з протоколами передачі пакетів інформації, передаються на спеціалізовані сервери. Відповідно, ОС можна вважати перетворювачем вхідних моніторингових потоків даних у вихідні інформаційні пакети (ІП), які підлягають передачі та ретрансляції абонентами СМ. Тому кожна ОС в місцях виникнення інформаційних потоків повинна мінімізувати кількість пакетів, що підлягають передачі. Оскільки каналом зв'язку вважають всі ресурси і засоби, які знаходяться між відправником і отримувачем інформації [4] то підвищення ефективності функціонування ОС СМ досягається шляхом підвищення ефективності функціонування апаратно-програмних засобів інформаційного рівня ОС (рівня введення, оброблення, кодування даних та формування ІП) та засобів радіотехнічного рівня. При заданих величинах робочої смуги частот F і ймовірності помилкового прийому кодової послідовності P_n ефективність функціонування мережі передачі інформації характеризується показником ефективності системи $\eta = R/C$, де $R_{\max} = f(F, P_n, (E_{is}/J_0)_n, K_c, L/B)$ – поточна максимальна швидкість передачі інформації, C – пропускна здатність каналу зв'язку (теоретична максимальна швидкість передачі інформації), $(E_{is}/J_0)_n$ – необхідне енергетичне співвідношення сигнал/шум в каналі зв'язку, $E_{is} = S \cdot T_{is}$, S – потужність сигналу, T_{is} – тривалість інформаційного символу (кодової послідовності), $J_0 = J/F$, J – середня потужність сумарних завад в каналі ($(E_{is}/J_0)_n \approx (S/J) \cdot B$ [4]), де $B = F \cdot T_{is}$ – база сигналу (коефіцієнт розширення спектру сигналу), K_c – сумарний коефіцієнт стиску даних, $L \leq B/4$ [5] – кількість ортогональних сигналів, які асиметрично можна передавати в загальному радіоканалі (величина L відповідає кількості незалежних кодових моно каналів в смузі частот F). Таким чином досягнення максимальної поточної швидкості передачі інформації R_i за умови підтримки необхідного енергетичного співвідношення $(E_{is}/J_0)_n$ у радіоканалі здійснюється шляхом адаптивного вибору мінімально необхідної бази каналних сигналів B_{\min} . При цьому швидкість передачі інформації $R_i = K_c \cdot L/k_s \cdot T_b \cdot B_{\min}$, де $K_c = K_i \cdot K_n$, K_i – коефіцієнт стиску даних на інформаційному рівні, який враховує стиск даних (сигналів, відеосигналів) з допустимими втратами та стиск двійкових даних без втрат, K_n – коефіцієнт стиску даних на радіотехнічному рівні (враховує підвищення швидкості передачі даних за рахунок багатопозиційних методів маніпуляції даних та ущільнення каналів передачі даних), $k_s \geq 1.4-1.8$ – коефіцієнт, що враховує якість відновлення фронтів цифрових сигналів [4], T_b – тривалість бітової послідовності. З урахуванням використання спрощених радіотехнічних засобів ($K_n = 1$), що характерно для ОС сенсорних мереж, ефективність передачі інформації можливо підвищити за рахунок оптимізації величини $K_c \rightarrow \max K_c$, де $K_c = k_1 \cdot k_2 \cdot k_3$, k_1 – коефіцієнт стиску даних з допустимими втратами інформації, який обмежується особливостями прикладних досліджень і завдань, k_2 – коефіцієнт стиску даних без втрат, k_3 – коефіцієнт стиску (ущільнення) даних в процесі формування сигнально-кодових послідовностей ІП.

Таким чином, досягнення високої інформаційної ефективності передачі даних в СМ забезпечується за рахунок використання об'єктними системами високопродуктивних процесорів, спеціалізованих пристроїв (кодеків, цифрових прийомопередавачів сигнально-кодових послідовностей) та програмно-апаратній реалізації комплексу взаємодоповнюючих алгоритмів стиску-захисту даних, завадостійкого кодування та формування компактних сигнально-кодових послідовностей ІП. При передачі вимірювальних сигналів і зображень для формування компактних ІП засобами ОС виконується наступна послідовність операцій:

- 1) визначення суттєвих відліків (СВ) обвідних сигналів (відеосигналів) та компактне кодування амплітудно-часових параметрів СВ;

- 2) оперативний стиск-захист даних без втрат з використанням словника та гаміювання даних з псевдовипадковими послідовностями, згенерованими з використанням абонентських закритих ключів;
- 3) оперативне завадостійке кодування даних, формування перевіркових кодів;
- 4) формування компактних кодово-сигнальних послідовностей (КСП) ІІ, перемішування КСП.

В процесі швидкодуючого стиску даних з допустимими втратами основними обчислювальними операціями є усереднення поточних відліків сигналів в адаптивно визначеному вікні усереднення даних [6] та визначення різницевих амплітудних значень $\Delta X_i = X_i^F - X_{i-1}^F$ між сусідніми відфільтрованими

відліками X_i^F та різницевих амплітудних значень сусідніх СВ [6,7]. Основою стиску та якісного відновлення сигналів (відеосигналів) є збереження характеристик обвідних сигналів з урахуванням вимог та особливостей прикладних завдань, галузей застосування адаптивних алгоритмів кодування даних. При цьому в процесі стиску параметрів СВ (глобальних і локальних екстремумів, точок перегину або точок зміни руху кривої) з допустимими (контрольованими) втратами інформації доцільно оперативно визначити найбільш інформативні (чисті від шумів) ділянки сигналу, на яких параметри СВ кодуються максимально точно та неінформативні ("зашумлені") ділянки, на яких досягається максимальний стиск даних. Після стиску даних з допустимими втратами в масивах, як правило, присутні збиткові двійкові послідовності даних (довгі послідовності однотипних бітів, n – бітові послідовності в різних комбінаціях, які часто повторюються, $n = 3, 4, 5 \dots$). Ці масиви підлягають подальшому стиску без втрат та криптозахисту з використанням псевдовипадкових послідовностей (ПВП), кодові ключі генерації яких відомі тільки абоненту-відправнику та абоненту-приймачу інформації. Беззбиткові та криптостійкі масиви даних, які фактично є псевдохаотичними даними, підлягають завадостійкому кодуванню, а при передачі пакетів інформації, з урахуванням якості каналу зв'язку (рівня шумів в радіоканалі), формуються відповідні завадостійкі кодово-сигнальні послідовності.

При реалізації алгоритму компактного кодування відліків сигналу, оптимізованого за швидкістю і точністю кодування, обробка і кодування даних здійснюється послідовними вибірками. Підвищення коефіцієнту стиску даних та швидкодії компактного кодування досягається за рахунок оперативного вибору максимально допустимого інтервалу опиту та аналізу поточних відліків сигналу, зменшенням кількості СВ, переш за все на недостовірних та пологих ділянках сигналу, визначенням ТП тільки на чистих від шумів нединамічних ділянках сигналу, коли $\Delta X_i^F \leq \Delta F_d$, де ΔX_i^F – поточний приріст сусідніх відліків відфільтрованого сигналу, ΔF_d – попередньо задана дослідником допустима величина поточної крутизни сигналу, а також шляхом опиту неінформативних ділянок сигналу з максимально допустимим інтервалом, який гарантує досягнення мінімально необхідного коефіцієнту стиску даних $K_{c \min}$. Чистою від шумів вважається ділянка сигналу яку утворюють два і більше сусідніх СВ, для яких виконується умова: $\Delta X_{CBi}^N < \delta_d^N$, де $\Delta X_{CBi}^F = |X_{CBi}^F - X_i^F|$ – оперативно визначена оцінка показника вхідного співвідношення сигнал/шум для i -го відфільтрованого СВ X_{CBi}^F , X_i^F – амплітудне значення вхідного сигналу, часовий відлік якого відповідає СВ X_{CBi}^F , δ_d^N – допустима величина рівня вхідних шумів в околиці СВ. При цьому на недостовірних ділянках сигналу амплітудні значення СВ кодуються з використанням мінімального значення кількості біт q_{\min} , а на чистих від шумів – з використанням кількості біт q_{\max} . Слід зазначити, що для визначення показника ΔX_{CBi}^F в околиці СВ можливі різноманітні обчислення оцінки вхідного співвідношення сигнал/шум, які задає дослідник з урахуванням особливостей прикладних досліджень. В швидкодуючому алгоритмі компактного кодування СВ вхідна вибірка даних проріджується зі сталим кроком прорідження відліків, з яким здійснюється попередня фільтрація сигналу та пошук параметрів первинних (приблизно визначених) екстремумів СВ-Е_p. На основі отриманих даних визначаються межі чистих від шумів та "зашумлених" ділянок сигналу, СВ яких кодуються з використанням службових даних та різницевих кодів амплітудних і часових даних. При цьому, в залежності від наявного часу обробки і кодування даних, амплітудно-часові параметри СВ (СВ-Е, СВ-ТП) на чистих від шумів ділянках сигналу уточнюються та кодуються більш точно.

Підвищення точності та швидкодії компактного кодування відліків сигналу з допустимими втратами інформації досягається за рахунок максимально допустимого прорідження відліків вхідної вибірки даних, адаптивного вибору вікна усереднення l_u в процесі фільтрації сигналу, де $l_u = f(\Delta X_{CB}^F)$ та шляхом реалізації адаптивної медіанної фільтрації і вибору оптимального (максимально допустимого) інтервалу опиту (відбору) відліків сигналу на чистих та зашумлених ділянках, враховуючи, що частота

дискретизації сигналу вибирається збитковою, тобто $f_d = k \cdot f_{\max}$, де f_{\max} – максимальна частота спектру сигналу, $k > 5-8$. З метою прискорення і оптимізації (мінімізації обчислювальних операцій) обробки та кодування даних інтервал відбору збиткових відліків поточної вибірки сигналу t_v визначається адаптивно з урахуванням залежності: $t_v = f(\Delta X_{CB_i}^N, \{\Delta X_{CB_j}^F\})$, де $\{\Delta X_{CB_j}^F\} = \Delta X_{CB_j}^F / d_j$ – нормована крутизна сигналу між поточними сусідніми СВ, $\Delta X_{CB_j}^F = X_{CB_j}^F - X_{CB_{j-1}}^F$, $d_j = 1, \dots, r$ – кількість відліків дискретизації сигналу між СВ_{j-1} і СВ_j, r – максимальна кількість відліків між сусідніми СВ.

Для передачі відеоданих по низькошвидкісним каналам СМ, в залежності від продуктивності абонентського процесора, прикладної орієнтації та галузі застосування відеокодека здійснюється прорідження відеокадрів. В процесі компактного кодування вибірок відеосигналів поточних кадрів визначаються СВ та показники їх інформативності [6,7]. На основі отриманих даних в кожному кадрі, що підлягає аналізу, визначають інформативність рядків та поточного кадру, після чого визначаються ключові кадри, компактні дані яких підлягають накопиченню в буфері та передачі по каналам зв'язку. Для реалізації якісного відеомоніторингу об'єктів центральна станція комірки СМ керуючими пакетами визначає послідовність та тривалість передачі відповідних фрагментів відеоданих з мінімальними затримками.

В процесі стиску-захисту даних без втрат захист інформації ґрунтується на виконанні операції додавання за модулем 2 над відповідними бітами двох послідовностей: послідовності бітів первинного масиву даних $X = x_1, \dots, x_i, \dots, x_n$ і послідовності випадкових бітів поточного секретного ключа $K = k_1, \dots, k_i, \dots, k_n$. В результаті виконання операцій додавання за модулем 2 отримуємо криптограму $Y = y_1, \dots, y_i, \dots, y_n$, для якої справедливий вираз $Y = X \oplus K$, де $y_1 = x_1 \oplus k_1, \dots, y_i = x_i \oplus k_i, \dots, y_n = x_n \oplus k_n$, а $X \oplus 0 = X$, $X \oplus X = 0$. Суттєвою вимогою при виконанні операцій шифрування даних з одноразовим ключем є дотримання вимоги, щоб при виконанні кожної наступної операції шифрування (додавання за модулем 2) використовувався інший, незалежно згенерований секретний ключ. Відповідно, для j -тої операції шифрування, парою абонентів, які приймають участь в передачі/прийомі ІІ, генерується поточна послідовність випадкових бітів $K_j = k_{1+j}, \dots, k_{i+j}, \dots, k_{n+j}$. Таким чином базовими операціями захисту масивів даних ІІ є використання абонентами мережі операцій генерації довготривалих ПВП, гаміювання відповідних масивів даних, формування перевіркових кодів пакетів даних. Величина ступеня захисту інформації P_z пропорційна величинам масивів даних, що підлягають гаміюванню, тобто $P_z \cong \max [2^m]$, де m – мінімально необхідна довжина поточної ПВП, яка використовується для надійного захисту інформації ($m \geq 1048$ біт).

Первинні потоки моніторингових даних визначаються мінімальними та максимальними значеннями амплітудних і частотних параметрів сигналів, відповідно A_{\min} , A_{\max} и f_{\min} , f_{\max} , що підлягають контролю. Для суттєвого зменшення інформаційних потоків в СМ на кожній ОС доцільно організувати оперативну обробку відповідних сигналів з визначенням показників, які підлягають апертурному або зонному контролю [1]. Ефективним способом зменшення інформаційних потоків на ОС є визначення інформаційних станів об'єктів моніторингу шляхом періодичного формування булевих змінних вектора $Q_s = (q_{1s}, q_{2s}, \dots, q_{vs}, \dots, q_{rs})$ по результатам контролю обчислених величин, де Q_s - вектор s -го інформаційного стану об'єкту моніторингу, q_{vs} - значення v -го булевого елемента Q_s - го вектора. Найбільш оперативно булеві змінні q_{vs} визначаються по результатам апертурного контролю вимірних величин або обчислених показників у відповідності до виразу

$$q_{vs} \begin{cases} 1, F(X_{pm}) - Z_v < \varepsilon_v / 2 \\ 0, F(X_{pm}) - Z_v \geq \varepsilon_v / 2 \end{cases},$$

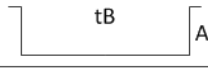
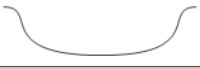


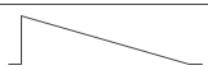

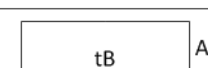

де $F(X_{pm})$ – виміряна або обчислена величина (функція) p -го сигналу з довжиною вибірки m відліків, Z_v – опорна величина (умовна норма) параметру, що підлягає контролю, ε_v – величина апертури.

З інженерної практики відомо, що при візуальному вивченні обвідних сигналів дослідник в першу чергу звертає увагу на екстремуми та точки перегину (зміни опуклості) кривої. Інформативність екстремумів і точок перегину важлива при дослідженні нединамічних (гладких) ділянок сигналу, а при дослідженні швидкодіючих процесів точками перегину можна знехтувати. Для експрес-аналізу різної тривалості сигналів в [8] запропонований метод оперативного спектрального аналізу сигналів з урахуванням амплітудно-часових параметрів екстремальних значень сигналу. До інформативних відліків, які впливають на спектральну характеристику реалізації сигналу, також слід віднести точки перегину, для яких природи різницевих значень сусідніх відліків $\Delta(\Delta X_i^F)$ по абсолютній величині перевищують задану величину. Тобто, окрім екстремумів враховуються інформативні точки, в яких характер огинаючої знає значних змін. Таким чином на інтервалі вибірки сигналу $x(t)$ визначають послідовність N інформативних відліків x_1, x_2, \dots, x_N і вимірюють часові інтервали $\{\Delta t_i\}$ між сусідніми i -м і $(i+1)$ -м інформативними відліками, а також визначають відповідні їм амплітудні прирости Δx_i , $i = \overline{1, N}$. Інтервали між сусідніми інформативними відліками містять інформацію про тривалість півхвилі j -ї частотної компоненти: $\Delta t_i = \Delta t_{i+1} - \Delta t_i = 1/2f_i$, де f_i – частота компоненти сигналу, що підлягає аналізу. Із приростів $\{\Delta x_i\}$, де $\Delta x_i = |x_{i+1} - x_i|$, отримуємо дані про амплітуду півхвиль. Шляхом сумування і усереднення $\{\Delta x_i\}$, які відносяться до однакових інтервалів Δt_i , отримуємо оцінку амплітуди спектральної складової j -ї частоти, тобто $a_j = \frac{1}{2Q} \sum_{i=1}^Q \Delta x_i$, де Q – кількість виміряних однакових по тривалості інтервалів між сусідніми інформативними відліками. Спектральна обробка інформативних відліків здійснюється на протязі k етапів, де $k \leq \log_2 N$, при цьому на кожному етапі виконуються наведені вище обчислення. Після усереднення значень $\{\Delta x_i\}$ отримуємо характеристику $S(t)$, яка, з інженерної точки зору, в повній мірі відображає спектральні значення процесу $x(t)$.

Формування завадостійких кодово-сигнальних послідовностей III

Враховуючи обмеження на робочу ширину смуги частот для функціонування СМ, що в свою чергу обмежує каналну швидкість передачі інформації, важливо на кожній ОС організувати мінімізацію кількості передач III обмеженої тривалості. Вирішення цього завдання здійснюється на основі експрес-обробки моніторингових даних засобами ОС з метою визначення найбільш інформативних даних, які підлягають першочерговій доставці (передачі) в локальну та центральну базу даних, а також за рахунок компактного кодування двійкових послідовностей III. Для зменшення тривалості III відповідні двійкові послідовності доцільно кодувати та передавати у вигляді сигнальних послідовностей мінімальної довжини. Основою формування компактних III засобами ОС є підхід, коли n – бітові послідовності, $n = 2, 3, 4, \dots$ кодуються (заміщуються) відповідними сигнальними послідовностями, мінімальна тривалість яких t_b відповідає тривалості бітової послідовності T_b . В результаті виконання такого кодування формуються вихідні КСП, які подаються на модулятор (маніпулятор) радіопередавача ОС. В залежності від способу радіопередачі інформації (на основі маніпуляції або з використанням модуляції несучої радіопередавача) формуються дворівневі КСП або комбіновані КСП з використанням дворівневих, лінійних та нелінійних сигнальних послідовностей. В таблиці 1 наведені сигнальні послідовності, які забезпечують зменшення тривалості пакету даних в 2 рази. При цьому, як правило, в передавачах використовують частотну модуляцію. На рис. 1 наведені приклади заміни двійкових даних сигнальними послідовностями. При наявності, наприклад, більше трьох сигнальних ознак "одиночного" або "нульового" рівнів доцільно після третьої сигнальної ознаки використовувати спеціалізовану сигнальну ознаку, яка перериває тривалу "одиночну" або "нульову" послідовність. При $n > 2$ доцільно визначати частоту зустрічі n – бітових послідовностей і тим послідовностям, які найчастіше зустрічаються – присвоюють менші по тривалості інтервальні сигнали. При використанні лінійних та нелінійних сигнальних послідовностей для заміщення n – бітових послідовностей, де $n = 2, 3$, коефіцієнт стиску даних k_3 , відповідно, досягає величини $k_3 \geq 2$. В усіх випадках на приймальній стороні необхідно використовувати спеціалізовані приймачі, які при вибраній кількості сигнальних ознак, здійснюють аналіз та цифрову обробку вихідних даних демодулятора та розпізнавання прийнятої

Таблиця 1

Код	Сигнал	Сигнал після фільтрації
00		
01		
10		
11		

поточної сигнальної ознаки. Запропоноване підвищення швидкості передачі даних орієнтоване на використання в безпроводових сенсорних мережах у випадку виявлення чистого від шумів радіоканалу.

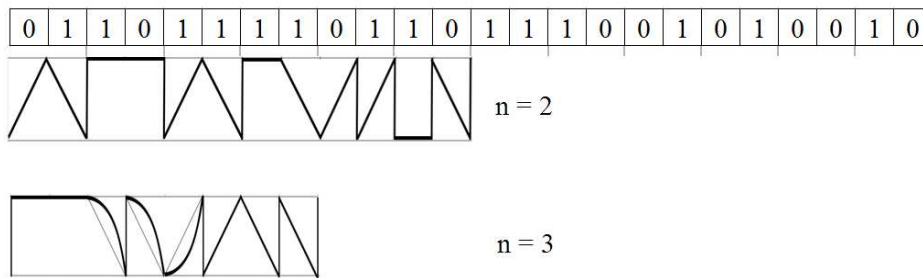


Рис.1. Приклади заміщення n – бітових послідовностей сигнальними послідовностями.

Підвищення завадостійкості передачі пакетів даних досягається за рахунок використання додаткових факторів в процесі кодування даних, що підлягають передачі, включаючи прорахунок нулів і одиниць первинного масиву даних (стислого та захищеного масиву даних) [9], тобто виконується підміна відповідних бітових послідовностей попередньо відомими бітами псевдовипадкових послідовностей, забезпечується перемішування даних, виконується формування відповідних КСП, які передаються на модулятор передавача. Також використовується формування перевірок кодів CRC даних пакетів. Коди CRC можуть передаватися декілька разів. Також ці коди можуть передаватися у вигляді шумоподібних сигналів. На приймальній стороні, з урахуванням вірно прийнятих сигнально-кодових послідовностей, введених факторів та способів кодування даних, забезпечується відновлення тих двійкових послідовностей, які були спотвореними канальними завадами. З метою реалізації завадостійкої передачі пакетів даних з підвищеною швидкістю при використанні мінімальної кількості сигнальних ознак, що значно спрощує декодер приймача ОС, актуальним завданням є розробка швидкодіючого методу завадостійкого кодування/декодування даних та формування КСП для передачі відповідних n – бітових послідовностей. Основою для зменшення кількості сигнальних ознак при формуванні КСП є уникнення тривалих однотипних m – бітових послідовностей в масиві даних, що підлягає передачі, $m = 2,3$. Для реалізації швидкодіючого завадостійкого кодування/декодування даних доцільне попереднє внесення залежностей між сусідніми бітами масиву даних без формування додаткової службової інформації. При цьому важливо змінювати (регулювати) кількість залежних сусідніх двійкових елементів масиву даних та характеристики залежності між відповідними елементами даних. Для оперативного відновлення спотворених n – бітових послідовностей на передавальній стороні, згідно прийнятих за основу поточних залежностей між сусідніми бітами, здійснюється багаточиклове кодування відповідних бітів даних по Грею. Суть кодування двійкових даних по Грею полягає в тому, що для двобітової послідовності x_k, x_l, x_k – старший бітовий розряд, формується код Грея, старший біт якого x_{kg} відповідає старшому розряду x_k , а наступний біт коду Грея x_{lg} дорівнює сумі по модулю два наступного біту x_l і попереднього біту x_k ,

тобто $x_{lg} = x_l \oplus x_k$ (рис. 2,а). Обернене перетворення є наступним: старші розряди співпадають, тобто $x_k = x_{kg}$, а наступний двійковий код x_l дорівнює сумі по модулю два отриманого попереднього розряду двійкового коду x_k і наступного коду Грея x_{lg} , тобто $x_l = x_k \oplus x_{lg}$ (рис. 2,б).

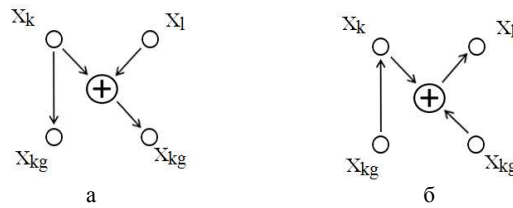


Рис. 2. Кодування та декодування двійкових даних по Грею

На рис. 3 наведені зв'язки між бітами масиву даних з виконанням логічних операцій сумування відповідних бітів по модулю два. При цьому на першому циклі кодування взаємозв'язками охоплюються поточні крайні біти із чотирьох сусідніх бітів, тобто в результаті їх сумування по модулю два визначаються біти X_i^1 , де $X_i^1 = X_{i-3} \oplus X_i$, $i = 1, p$. Перші три біти масиву даних окремо охоплюються зв'язками, утворюючи сусідні залежні пари бітів. На другому циклі кодування взаємозв'язками охоплюються поточні крайні біти із трьох сусідніх бітів, при цьому $X_i^2 = X_{i-2}^1 \oplus X_i^1$, а перші два біти окремо охоплюються зв'язками. На третьому циклі кодування взаємозв'язками охоплюються пари сусідніх бітів, при цьому $X_i^3 = X_{i-1}^2 \oplus X_i^2$. Таким чином, в результаті трьох циклового кодування даних для кожного поточного біту даних вноситься відповідна залежність з попередніми бітами. В даному випадку, починаючи з першого біту X_1 до сьомого біту X_7 , збільшується залежність до 7-ми сусідніх бітів, яка підтримується на протязі тривалості масиву даних, що підлягає кодуванню. При цьому характерною рисою даного трьох циклового кодування бітів даних по Грею є внесення залежностей між віддаленими сусідніми бітами масиву даних, а на останньому циклі взаємозалежними є сусідні біти, значення яких визначені на попередніх циклах.

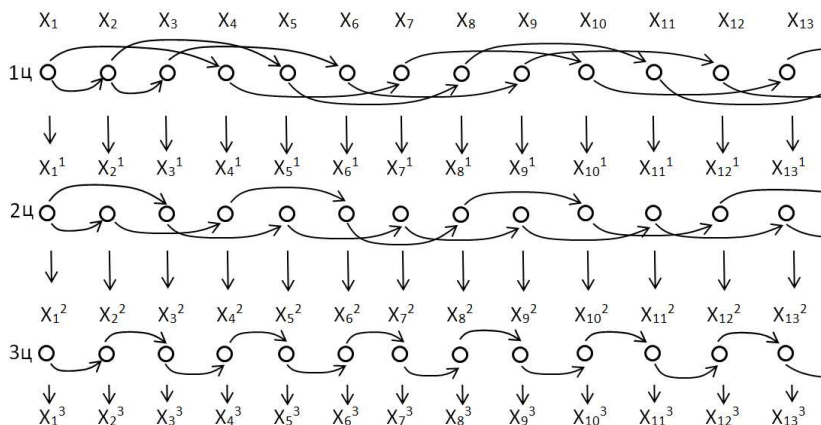


Рис. 3. Приклади трьох циклового кодування двійкових даних по Грею

В результаті виконання багато циклового кодування відповідних бітів даних по Грею на виході кодера завадостійкого кодування даних формується масив взаємозалежних бітів даних, який в процесі передачі П перетворюється в послідовності КСП. Для підвищення завадостійкості передачі П відповідні КСП підлягають перемішуванню. На приймальній стороні, після деперемішування прийнятих сигнально-кодівих послідовностей, у випадку враження шумами радіоканалу тих чи інших КСП в прийнятому потоці двійкових даних будуть присутні невідомі n – бітові послідовності. Для їх оперативного відновлення шляхом перебору можливих комбінацій n – бітових даних здійснюється відповідне декодування даних по Грею. При цьому для кожної із можливих комбінацій враженої n – бітової послідов-

ності, з урахуванням значень достовірно прийнятих бітів, які відповідають вірно прийнятим сигнально-кодним послідовностям, після декодування визначаються можливі значення відновлених даних на короткому інтервалі. В подальшому, для різних версій отриманих можливих значень відновлених даних на приймальної стороні здійснюється відповідне кодування даних по Грею. По результатах співпадання визначених даних по Грею з достовірно прийнятими бітами приймається рішення про достовірне відновлення відповідної версії можливих значень відновлених даних, яке підтверджується співпаданням перевіркових кодів CRC.

Висновки

1. Досягнення високої інформаційної ефективності передачі даних в СМ забезпечується за рахунок розміщення в місцях зародження інформаційних потоків моніторингових даних (сигналів, відеосигналів, масивів даних) високопродуктивних ОС, засоби яких реалізують комплекс алгоритмів оперативного оброблення, кодування та шифрування даних, формування компактних і захищених пакетів інформації, при цьому кожна ОС повинна мінімізувати кількість передач пакетів.

2. Основою для зменшення кількості передач пакетів є реалізація засобами ОС оперативних методів компактного кодування даних, оптимізованих за швидкодією і точністю кодування даних та формування компактних кодово-сигнальних послідовностей пакетів, при цьому основою якісного та оперативного стиску і відновлення сигналів і відеосигналів є збереження амплітудно-часових характеристик суттєвих та вагомих відліків обвідних кривих (відповідних екстремумів та точок перегину), а основними обчислювальними операціями є усереднення поточних відліків сигналів в адаптивно визначеному вікні усереднення даних та визначення різницевої амплітудних значень відфільтрованих відліків і суттєвих відліків.

3. Для оперативного формування компактних та завадостійких КСП ІІ між бітами даних кадру вносяться відповідні взаємозалежності по Грею.

Література

1. Шевчук Б.М. Технологія багатофункціональної обробки і передачі інформації в моніторингових мережах/ Б.М. Шевчук, В.К. Задірака, Л.О. Гнатів, С.В. Фраєр// – К.: Наук. думка, 2010. – 370 с.
2. Шевчук Б.М. Оброблення, кодування та передавання даних засобами абонентських систем інформаційно-ефективних радіомереж/ Б.М. Шевчук// Комп'ютерні засоби, мережі та системи. – 2010. - №9. – С. 130-139.
3. Шевчук Б.М. Оперативне формування і передавання компактних, криптистійких та завадостійких пакетів інформації у радіомережах/ Б.М. Шевчук// Комп'ютерні засоби, мережі та системи. – 2011. - №10. – С. 143-152.
4. Скляр Б. Цифровая связь. Теоретические основы и практическое применение/ Б.Скляр// 2-е изд.: Пер. с англ. – М.: Издательский дом "Вильямс", 2003. – 1104 с.
5. Голяницкий И.А. Математические модели и методы в радиосвязи/ И.А. Голяницкий// – М.: Эко-трендз, 2005. – 440 с.
6. Шевчук Б.М. Високоінформативна обробка та кодування сигналів і зображень в мережах дистанційного моніторингу станів об'єктів/ Б.М. Шевчук, В.П. Зінченко, С.В. Фраєр// Наукові вісті НТУУ "КПІ". – 2008. №2. – С. 89-96.
7. Шевчук Б.М. Эффективные методы и алгоритмы оперативного сжатия сигналов и изображений, ориентированные на построение объектных систем сенсорных сетей/ Б.М. Шевчук// Компьютерная математика. – 2012. - №2. – С. 100-106.
8. Пономарёва И.Д. Сверхбыстрый спектральный анализ/ И.Д. Пономарёва, Г.В. Цепков// Проблемы управления и информатики. – 1998. - №1. – С. 107-114.
9. Николайчук Я.М. Дослідження ефективності формування сигнальних кодів/ Я.М. Николайчук Т.М. Гринчишин, О.М. Заставний, А.Р. Воронич// Комп'ютерні системи та компоненти. – Науковий вісник Чернівецького університету, 2009. – Вип. 479. – С. 114 – 125.

Відомості про автора

Шевчук Богдан Михайлович – к.т.н., старший науковий співробітник Інституту кібернетики ім. В.М. Глушкова НАН України, м. Київ, incors@ukr.net.

РАДІОВИМІРЮВАЛЬНІ ПРИЛАДИ

УДК 621.383

В. С. ОСАДЧУК, О. В. ОСАДЧУК, В. В. МАРТИНЮК, О. П. БІЛИЛІВСЬКА

Вінницький національний технічний університет, м. Вінниця

МІКРОЕЛЕКТРОННИЙ ПЕРЕТВОРЮВАЧ МАГНІТНОГО ПОЛЯ З ЧАСТОТНИМ ВИХОДОМ З МАГНІТОЧУТЛИВИМ ЕЛЕМЕНТОМ ХОЛЛА

Анотація. Запропонована мікроелектронна схема перетворювача магнітної індукції у частотний сигнал з магніточутливим елементом Холла. На основі математичної моделі, що складається з рівнянь Кірхгофа, які описують поведінку схеми, отримані функції перетворення та рівняння чутливості. Чутливість перетворювача складає 2-18 кГц/мТл в діапазоні зміни магнітної індукції від 0,01 мТл до 300 мТл.

Ключові слова: магнітне поле, частотний вихід, елемент Холла, математична модель, чутливість.

Аннотация. Предложена микроэлектронная схема преобразователя магнитной индукции в частотный сигнал с магниточувствительным элементом Холла. На основе математической модели, базирующейся на уравнениях Кирхгофа, которые описывают поведение схемы, получены функции преобразования и уравнение чувствительности. Чувствительность преобразователя составляет 2-18 кГц/мТл в диапазоне изменения магнитной индукции от 0,01 мТл до 300 мТл.

Ключевые слова: магнитное поле, частотный выход, элемент Холла, математическая модель, чувствительность

Abstract. The microelectronic circuit of the transducer of magnetic induction into frequency signal with magnetosensitive Hall element is proposed. On the basis of the mathematical model basing equations Kirgoff which present behavior of the circuit, function of transformation and the equation of sensitivity are received. The sensitivity has 2-18 kHz/mT in range of magnetic induction from 0,01 mT up to 300 mT.

Key words: magnetic field, frequency output, Hall element, mathematical model, sensitivity.

Вступ

На сьогодні магнітні сенсори широко використовуються як для безпосереднього вимірювання магнітних величин, так і в якості проміжних перетворювачів для реєстрації частоти обертання, положення, відстані, кутових та лінійних переміщень, вимірювання струму, виявлення дефектів у магнітних матеріалах та виробках тощо. Разом з тим, вимоги до чутливості, роздільної здатності та інших важливих параметрів вимірювальних систем постійно зростають. Тому виникає потреба у розробці нових перетворювачів магнітного поля, які б відповідали цим вимогам.

Актуальність

Елементи Холла є найпоширенішими перетворювачами магнітного поля в електричний сигнал. Інтегральні сенсори Холла виробляють такі фірми, як Honeywell, Melexis, Allegro Microsystems, Micronas Intermetall, Siemens, Analog Devices та інші. Як правило, необхідна попередня обробка сигналу з таких вимірювачів полягає в підсиленні, температурній компенсації тощо [1-3]. Оскільки сучасні сенсори магнітного поля на основі ефекту Холла є переважно аналоговими пристроями, їх недоліками є низький рівень вихідного сигналу (десяті долі – одиниці мілівольт), низька завадостійкість (і відповідно, низька точність), необхідність аналогово-цифрового перетворення та підсилювальної апаратури.

Використання транзисторних структур з від'ємним опором для побудови мікроелектронних перетворювачів магнітного поля з частотним виходом виключає використання аналого-цифрових перетворювачів при обробці сигналів, що знижує собівартість систем контролю і управління. Крім того, частотні магнітні перетворювачі поєднують простоту і універсальність, які властиві аналоговим пристроям, із точністю та завадостійкістю, що характерні для перетворювачів з кодовим виходом, мають високу чутливість, малу масу, габарити, конструктивну та технологічну сумісність з мікроелектронними засобами обробки інформації. Що і є перевагою їх перед існуючими сенсорами магнітної індукції [4, 5].

Мета

Метою дослідження є розгляд можливості побудови мікроелектронного перетворювача магнітної індукції у частотний сигнал на основі нового підходу – використання реактивних властивостей транзисторних структур з від'ємним опором.

Задачі

1. Розробити математичну модель магнітного перетворювача.
2. Отримати аналітичні залежності функції перетворення і рівняння чутливості.
3. Розробити схему перетворювача магнітної індукції у частоту.
4. Здійснити технічну реалізацію і провести експериментальні дослідження.
5. Порівняти теоретичні і експериментальні залежності основних параметрів перетворювачів.

Математична модель

З метою вивчення властивостей частотного перетворювача магнітної індукції на основі біполярних транзисторів необхідно розробити математичну модель, яка дозволила б отримати аналітичні залежності активної і реактивної складової повного опору на електродах колектор-колектор структури, функцій

перетворення і чутливості від дії магнітної індукції. Електрична схема перетворювача приведена на рис. 1. Це гібридна інтегральна схема, що складається з трьох біполярних транзисторів, які є основою автогенераторного пристрою, у ланцюг зворотного позитивного зв'язку якого включений сенсор Холла.

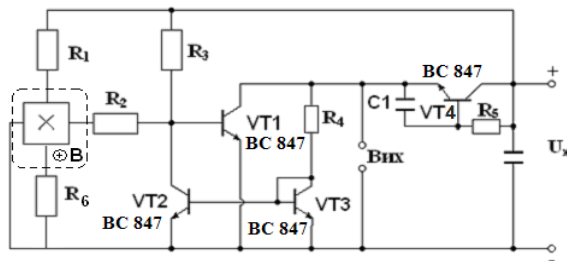


Рисунок 1 – Електрична схема частотного перетворювача магнітної індукції

На електродах колектор-емітер транзисторів VT1 і VT2 існує повний опір, активна складова якого має від'ємне значення, а реактивна – ємнісний характер. Підключення активної індуктивності до колектора VT1 створює коливальний контур, втрати енергії в якому компенсуються від'ємним опором [3]. Резонансна частота коливального контуру залежить від дії магнітної індукції на сенсор Холла. Еквівалентна схема частотного перетворювача магнітної індукції представлена на рис. 2. Вона враховує нелінійні властивості схеми, тому що автогенератор може працювати як у лінійному, так і нелінійному режимі.

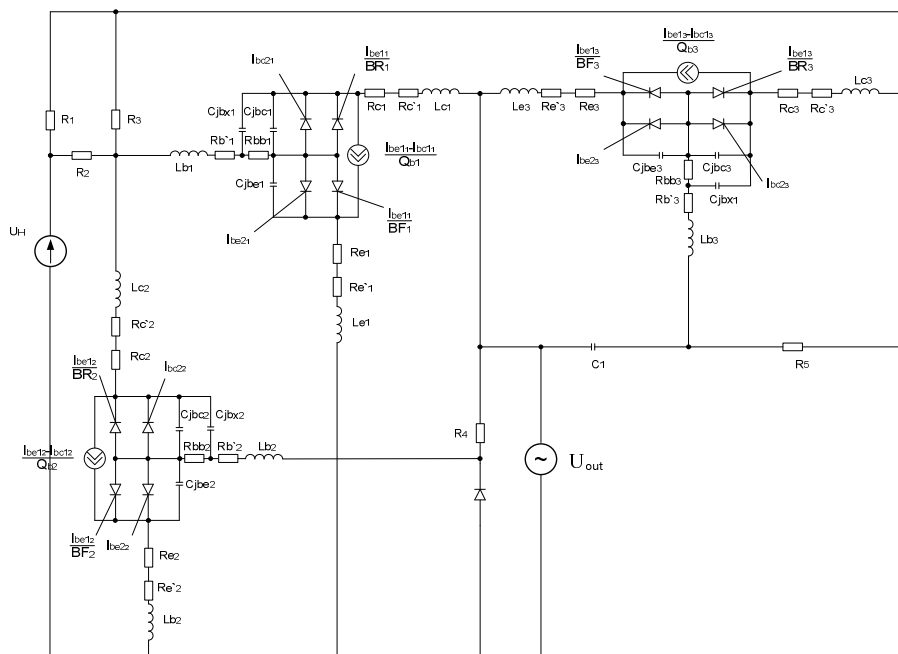


Рисунок 2 – Еквівалентна схема частотного перетворювача магнітної індукції

Для зручності розрахунків еквівалентну схему подано у вигляді рис. 3. Вольт-амперна характеристика має ділянку від'ємного опору, що дозволяє компенсувати втрати енергії в коливальному контурі, який утворений еквівалентною ємністю транзисторів VT1 та VT2 (рис. 1) і активною індуктивністю.

Для визначення функції перетворення необхідно знайти залежність частоти генерації від магнітної індукції. Це можливо зробити, розв'язавши систему рівнянь Кірхгофа, яка складена для змінного струму на основі еквівалентної схеми (рис. 3). Вирішення системи рівнянь (1) дозволяє отримати величину повного опору на електродах колектор-емітер перетворювача. При розкладі повного опору на дійсний і уявний складові можна визначити еквівалентну ємність коливального контуру, яка залежить від магнітної індукції. Система рівнянь Кірхгофа, згідно напрямкам контурних струмів, має вигляд:

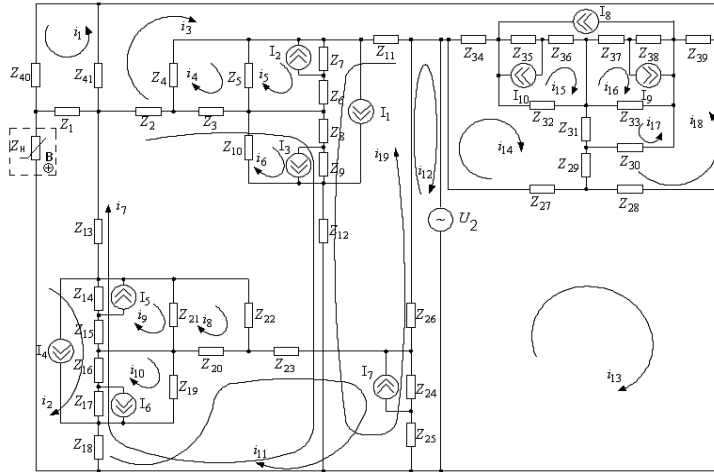


Рисунок 3 – Перетворена нелінійна еквівалентна схема частотного перетворювача на основі трьох біполярних транзисторів

$$\begin{cases}
 0 = (Z_1 + Z_{40} + Z_{41})i_1 + Z_1 i_2 + Z_{41} i_3; \\
 0 = (Z_1 + Z_{13} + Z_{14} + Z_{15} + Z_{16} + Z_{17} + Z_{18} + Z_H)i_2 + Z_1 i_1 - Z_{13} i_7 + Z_{14}(I_5 - i_9 - I_4 - i_7) - \\
 - Z_{15}(I_4 + i_9 + i_7) - Z_{16}(I_4 + i_{10} + i_7) - Z_{17}(I_4 + I_6 + i_{10} + i_7) - Z_{18} i_7; \\
 0 = (Z_4 + Z_2 + Z_{41} + Z_{39} + Z_{38} + Z_{37} + Z_{36} + Z_{35} + Z_{34} + Z_1)i_3 + Z_{41} i_1 - Z_2 i_7 - Z_4 i_4 + Z_{11} i_{19} - \\
 - Z_{34} i_{14} - Z_{35}(I_8 + I_{10} - i_{15}) - Z_{36}(I_8 + i_{15}) - Z_{37}(I_8 + i_{16}) + Z_{38}(I_9 - I_8 - i_{16}) + Z_{39} i_{18}; \\
 0 = (Z_3 + Z_4 + Z_5)i_4 + Z_4 i_3 - Z_3 i_7 - Z_5 i_5; \\
 0 = (Z_5 + Z_7 + Z_6)i_5 - Z_5 i_4 + Z_7(I_2 + i_{19} - I_1) + Z_6(i_{19} - I_1); \\
 0 = (Z_8 + Z_9 + Z_{10})i_6 + Z_8(i_7 + i_{19} - I_1) + Z_9(i_{19} - I_1 - I_3 + i_7); \\
 0 = (Z_2 + Z_3 + Z_8 + Z_9 + Z_{12} + Z_{18} + Z_{17} + Z_{16} + Z_{15} + Z_{14} + Z_{13})i_7 - Z_2 i_3 - Z_3 i_4 + Z_8(i_6 - \\
 - I_1 + i_{19}) + Z_9(i_6 - I_3 - I_1 + i_{19}) + Z_{12} i_{19} + Z_{18}(i_{11} - i_2) + Z_{17}(i_{10} + I_6 - i_2 + I_4) + \\
 + Z_{16}(i_{10} + I_4 - i_2) + Z_{15}(i_9 - i_2 + I_4) + Z_{14}(i_9 - I_5 - i_2 + I_4) - Z_{13} i_2; \\
 0 = (Z_{20} + Z_{21} + Z_{22})i_8 - Z_{21} i_9 - Z_{20} i_{11}; \\
 0 = (Z_{14} + Z_{15} + Z_{21})i_9 + Z_{14}(i_7 - i_2 - I_5 + I_4) + Z_{15}(i_7 - i_2 + I_4) - Z_{21} i_8; \\
 0 = (Z_{16} + Z_{17} + Z_{19})i_{10} + Z_{16}(i_7 - i_2 + I_4) + Z_{17}(i_7 - i_2 + I_4 + I_6) - Z_{19} i_{11}; \\
 0 = (Z_{18} + Z_{20} + Z_{19} + Z_{23} + Z_{24} + Z_{25})i_{11} - Z_{19} i_{10} - Z_{20} i_8 + Z_{24}(I_7 - i_{12} - i_{19}) - \\
 - Z_{25}(i_{12} + i_{19}) + Z_{18}(i_7 - i_2); \\
 \dot{U}_2 = (Z_{25} + Z_{24} + Z_{26})i_{12} + Z_{25}(i_{19} - i_{11}) + Z_{24}(i_{19} - i_{11} - I_7) + Z_{26} i_{19}; \\
 \dot{U}_2 = i_{13}(Z_{27} + Z_{28}) - i_{14} Z_{27} + i_{18} Z_{28}; \\
 0 = i_{14}(Z_{34} + Z_{32} + Z_{31} + Z_{29} + Z_{27}) - i_{15} Z_{32} + i_{17} Z_{31} + i_{18} Z_{29} - i_{13} Z_{27} - i_3 Z_{34}; \\
 0 = i_{15}(Z_{35} + Z_{36} + Z_{32}) + I_{10} Z_{35} + I_8 Z_{35} + I_8 Z_{36} - i_{14} Z_{32} - i_3(Z_{35} + Z_{36}); \\
 0 = i_{16}(Z_{37} + Z_{38} + Z_{33}) - I_9 Z_{38} + I_8 Z_{37} + I_8 Z_{38} + i_{17} Z_{33} - i_3(Z_{37} + Z_{38}); \\
 0 = i_{17}(Z_{31} + Z_{33} + Z_{30}) - i_{18} Z_{30} + i_{14} Z_{31} + i_{16} Z_{33}; \\
 0 = i_{18}(Z_{29} + Z_{39} + Z_{30} + Z_{28}) - i_{14} Z_{29} - i_{17} Z_{30} + i_3 Z_{39} + i_{13} Z_{28}; \\
 0 = i_{19}(Z_{11} + Z_7 + Z_6 + Z_8 + Z_9 + Z_{12} + Z_{25} + Z_{24} + Z_{26}) + I_2 Z_7 + i_5 Z_7 - i_1 Z_7 + \\
 + i_5 Z_6 - I_1 Z_6 + i_6 Z_8 - I_1 Z_8 + i_7 Z_8 - I_3 Z_9 + i_6 Z_9 - I_1 Z_9 + i_7 Z_9 + i_7 Z_{12} + i_{12} Z_{25} - \\
 - i_{11} Z_{25} + i_{12} Z_{24} - I_7 Z_{24} + i_{12} Z_{26} + i_3 Z_{11} - Z_{24} i_{11},
 \end{cases} \quad (1)$$

де $Z_1 = R_2$, $Z_2 = R_{b1} + j\omega L_{b1}$, $Z_3 = R_{bb1}$, $Z_4 = -j/(\omega C_{bx1})$, $Z_5 = -j/(\omega C_{jbc1})$, $Z_6 = R_8$,
 $Z_7 = -j/(\omega C_{bc21})$, $Z_8 = R_9$, $Z_9 = -j/(\omega C_{be21})$, $Z_{10} = -j/(\omega C_{jbe1})$, $Z_{11} = R_{c1} + R'_{c1} + j\omega L_{c1}$,

$Z_{12} = R_{e1} + R'_{e1} + j\omega L_{e1}$, $Z_{13} = R_{c2} + R'_{c2} + j\omega L_{c2}$, $Z_{14} = -j/(\omega C_{bc22})$, $Z_{15} = R_8$, $Z_{16} = R_9$,
 $Z_{17} = -j/(\omega C_{be22})$, $Z_{18} = R_{e2} + R'_{e2} + j\omega L_{e2}$, $Z_{19} = -j/(\omega C_{jbe2})$, $Z_{20} = R_{bb2}$,
 $Z_{21} = -j/(\omega C_{jbc2})$, $Z_{22} = -j/(\omega C_{bx2})$, $Z_{23} = R_{b2} + j\omega L_{b2}$, $Z_{24} = -j/(\omega C_d)$, $Z_{25} = R_d$,
 $Z_{26} = R_4$, $Z_{27} = -j/(\omega C_1)$, $Z_{28} = R_5$, $Z_{29} = R_{b3} + j\omega L_{b3}$, $Z_{30} = -j/(\omega C_{bx3})$, $Z_{31} = R_{bb3}$,
 $Z_{32} = -j/(\omega C_{jbe3})$, $Z_{33} = -j/(\omega C_{jbc3})$, $Z_{34} = R_{e3} + R'_{e3} + j\omega L_{e3}$, $Z_{35} = -j/(\omega C_{be23})$,
 $Z_{36} = R_{11}$, $Z_{37} = R_{10}$, $Z_{38} = -j/(\omega C_{bc23})$, $Z_{39} = R_{c3} + R'_{c3} + j\omega L_{c3}$, $Z_{40} = R_1$, $Z_{41} = R_3$,
 $Z_H = R_H / (1 + w^2 R_H^2 C_H^2) - jR_H^2 w C_H / (1 + w^2 R_H^2 C_H^2)$, R'_b - омичний опір базового електроду, R_b -
 опір бази, L_b - індуктивність базового електроду, C_{jbc} - ємність колекторного переходу, C_{jbe} - єм-
 ність емітерного переходу, C_{bx} - вхідна ємність транзистора, R'_c - омичний опір колекторного електро-
 ду, R_c - опір колекторного переходу, L_c - індуктивність колекторного електроду, R'_e - омичний опір
 емітерного електроду, R_e - опір емітерного переходу, R_H - опір холлівського елемента, C_H - ємність
 холлівського елемента, L_e - індуктивність емітерного електроду, R_1, R_2, R_3, R_4 - опори дільника.

Ємність емітерного переходу [4]:

$$C_{be} = C_{едиф} + C_{ебар}, C_{едиф} = \tau / (m_e \cdot \Phi_T) (I_{едиф} - I_{be0}), C_{ебар} = C_{ебар0} [U_{кон.е} / (U_{кон.е} - U_e)]^{n_e},$$

де $C_{едиф}$, $C_{ебар}$ - дифузійна та бар'єрна ємності емітерного переходу, τ - ефективний час життя неосновних носіїв бази при нормальному включенні біполярного транзистора, $C_{ебар0}$ - бар'єрна ємність емітерного переходу при нульовій напрузі зміщення, $U_{кон.е}$ - контактна різниця потенціалу емітерного переходу, n_e - коефіцієнт, який залежить від характеру переходу.

Ємність колекторного переходу [4]:

$$C_{bc} = C_{сдиф} + C_{сбар}, C_{сдиф} = \frac{\tau_i}{m_c \cdot \Phi_T} (I_{сдиф} - I_{bc0}), C_{сбар} = C_{сбар0} [U_{кон.с} / (U_{кон.с} - U_c)]^{n_c},$$

де $C_{сдиф}$, $C_{сбар}$ - дифузійна та бар'єрна ємності колекторного переходу, τ_i - ефективний час життя неосновних носіїв бази при інверсному включенні біполярного транзистора, $C_{сбар0}$ - бар'єрна ємність переходу при нульовій напрузі зміщення, $U_{кон.с}$ - контактна різниця потенціалу колекторного переходу, n_c - коефіцієнт, який залежить від характеру колекторного переходу.

Джерела струмів $I_1, I_2, I_3, I_4, I_5, I_6, I_7, I_8, I_9, I_{10}$ описуються виразами [5, 6]:

$$I_1 = \frac{I_{be1} - I_{bc1}}{Q_{b1}}, I_4 = \frac{I_{be2} - I_{bc2}}{Q_{b2}}, I_8 = \frac{I_{be3} - I_{bc3}}{Q_{b3}},$$

$$I_2 = I_{bc1} / \beta_i, I_3 = I_{be1} / \beta_N, I_5 = I_{bc2} / \beta_N, I_6 = I_{bc2} / \beta_i, I_7 = I_b / \beta_i, I_9 = I_{bc3} / \beta_i, I_{10} = I_{be3} / \beta_N,$$

де I_{bc} , I_{be} - струм база-колектор та струм база-емітер, β_i - статистичний коефіцієнт передачі струму в інверсному режимі, β_N - статистичний коефіцієнт передачі струму в нормальному режимі, $I_{едиф}$ - дифузійна складова емітерного струму, $I_{сдиф}$ - дифузійна складова колекторного струму.

Струми база-емітер I_{be} і база-колектор I_{bc} описуються виразами [5]:

$$I_{be} = I_{be0} [\exp(U_{be}/(m_e \phi_T)) - 1], \quad I_{bc} = I_{bc0} [\exp(U_{bc}/(m_c \phi_T)) - 1],$$

де I_{be0} , I_{bc0} - теплові струми емітерного і колекторного переходів, U_{be} , U_{bc} - напруги на переходах база-емітер та база-колектор, m_e , m_c - коефіцієнти неідеальності транзисторів, ϕ_T - температурний потенціал.

Дифузійна складова емітерного та колекторного струмів описуються виразами [7]:

$$I_{e\text{диф}} = I_{be0} [\exp(U_{be}/(m_e \phi_T)) - 1], \quad I_{c\text{диф}} = I_{bc0} [\exp(U_{bc}/(m_c \phi_T)) - 1].$$

Опір бази визначається як сума двох складових

$$R_b = R_{b0} + R_{b\text{д}}, \quad (2)$$

де R_{b0} - омичний опір матеріалу бази, $R_{b\text{д}}$ - дифузійний опір.

Омичний опір матеріалу бази визначається формулою

$$R_{b0} = \rho \left[\frac{1}{2W} + \frac{l}{aW_1} \right], \quad (3)$$

де ρ - питомий опір матеріалу бази, a , l , W , W_1 - ширина, довжина та товщина активної та пасивної частини базової області транзистора.

Дифузійний опір $R_{b\text{д}}$ визначається формулою

$$R_{b\text{д}} = \frac{kT}{2qI_e(1-h_{12b})}. \quad (4)$$

З врахуванням (3) та (4) формула (2) набуває вигляду

$$R_b = \rho \left[\frac{1}{2W} + \frac{l}{aW_1} \right] + \frac{kT}{2qI_e(1-h_{12b})}. \quad (5)$$

Опір емітера визначається згідно [5]

$$R_e = \frac{kT}{qI_e}.$$

Опір колекторного переходу визначається виразом [5]

$$R_c = \frac{WU_c}{I_e d_c (1-h_{21b})},$$

де W - ширина бази, d_c - шар об'ємного заряду колекторного переходу, h_{21b} - коефіцієнт передачі струму.

Для спрощення запису системи рівнянь Кірхгофа введемо наступні позначення:

$$\begin{cases}
 B_1 = Z_1 + Z_{40} + Z_{41}; \\
 B_2 = Z_1 + Z_{13} + Z_{14} + Z_{15} + Z_{16} + Z_{17} + Z_{18} + Z_H; \\
 B_3 = Z_4 + Z_2 + Z_{41} + Z_{39} + Z_{38} + Z_{37} + Z_{36} + Z_{35} + Z_{34} + Z_{11}; \\
 B_4 = Z_3 + Z_4 + Z_5; \\
 B_5 = Z_5 + Z_7 + Z_6; \\
 B_6 = Z_8 + Z_9 + Z_{10}; \\
 B_7 = Z_2 + Z_3 + Z_8 + Z_9 + Z_{12} + Z_{18} + Z_{17} + Z_{16} + Z_{15} + Z_{14} + Z_{13}; \\
 B_8 = Z_{20} + Z_{21} + Z_{22}; \\
 B_9 = Z_{14} + Z_{15} + Z_{21}; \\
 B_{10} = Z_{16} + Z_{17} + Z_{19}; \\
 B_{11} = Z_{18} + Z_{19} + Z_{20} + Z_{23} + Z_{24} + Z_{25}; \\
 B_{12} = Z_{25} + Z_{24} + Z_{26}; \\
 B_{13} = Z_{27} + Z_{28}; \\
 B_{14} = Z_{34} + Z_{32} + Z_{31} + Z_{29} + Z_{27}; \\
 B_{15} = Z_{35} + Z_{36} + Z_{32}; \\
 B_{16} = Z_{37} + Z_{38} + Z_{33}; \\
 B_{17} = Z_{31} + Z_{33} + Z_{30}; \\
 B_{18} = Z_{29} + Z_{30} + Z_{39} + Z_{28}; \\
 B_{19} = Z_{11} + Z_7 + Z_6 + Z_8 + Z_9 + Z_{12} + Z_{25} + Z_{24} + Z_{26}.
 \end{cases} \quad (6)$$

Тоді система (3.1) набуває вигляду :

$$\begin{cases}
 0 = B_1 i_1 + Z_1 i_2 + Z_{41} i_3; \\
 0 = B_2 i_2 + Z_1 i_1 - Z_{13} i_7 + Z_{14} (I_5 - i_9 - I_4 - i_7) - Z_{15} (I_4 + i_9 + i_7) - Z_{16} (I_4 + i_{10} + i_7) - \\
 - Z_{17} (I_4 + I_6 + i_{10} + i_7) - Z_{18} i_7; \\
 0 = B_3 i_3 + Z_{41} i_1 - Z_2 i_7 - Z_4 i_4 + Z_{11} i_{19} - Z_{34} i_{14} - Z_{35} (I_8 + I_{10} - i_{15}) - \\
 - Z_{36} (I_8 + i_{15}) - Z_{37} (I_8 + i_{16}) + Z_{38} (I_9 - I_8 - i_{16}) + Z_{39} i_{18}; \\
 0 = B_4 i_4 + Z_4 i_3 - Z_3 i_7 - Z_5 i_5; \\
 0 = B_5 i_5 - Z_5 i_4 + Z_7 (I_2 + i_{19} - I_1) + Z_6 (i_{19} - I_1); \\
 0 = B_6 i_6 + Z_8 (i_7 + i_{19} - I_1) + Z_9 (i_{19} - I_1 - I_3 + i_7); \\
 0 = B_7 i_7 - Z_2 i_3 - Z_3 i_4 + Z_8 (i_6 - I_1 + i_{19}) + Z_9 (i_6 - I_3 - I_1 + i_{19}) + Z_{12} i_{19} + Z_{18} (i_{11} - i_2) + \\
 + Z_{17} (i_{10} + I_6 - i_2 + I_4) + Z_{16} (i_{10} + I_4 - i_2) + Z_{15} (i_9 - i_2 + I_4) + Z_{14} (i_9 - I_5 - i_2 + I_4) - Z_{13} i_2; \\
 0 = B_8 i_8 - Z_{21} i_9 - Z_{20} i_{11}; \\
 0 = B_9 i_9 + Z_{14} (i_7 - i_2 - I_5 + I_4) + Z_{15} (i_7 - i_2 + I_4) - Z_{21} i_8; \\
 0 = B_{10} i_{10} + Z_{16} (i_7 - i_2 + I_4) + Z_{17} (i_7 - i_2 + I_4 + I_6) - Z_{19} i_{11}; \\
 0 = B_{11} i_{11} - Z_{19} i_{10} - Z_{20} i_8 + Z_{24} (I_7 - i_{12} - i_{19}) - Z_{25} (i_{12} + i_{19}) + Z_{18} (i_7 - i_2); \\
 U_2 = B_{12} i_{12} + Z_{25} (i_{19} - i_{11}) + Z_{24} (i_{19} - i_{11} - I_7) + Z_{26} i_{19}; \\
 \dot{U}_2 = i_{13} B_{13} - i_{14} Z_{27} + i_{18} Z_{28}; \\
 0 = i_{14} B_{14} - i_{15} Z_{32} + i_{17} Z_{31} + i_{18} Z_{29} - i_{13} Z_{27} - i_3 Z_{34}; \\
 0 = i_{15} B_{15} + I_{10} Z_{35} + I_8 Z_{35} + I_8 Z_{36} - i_{14} Z_{32} - i_3 (Z_{35} + Z_{36}); \\
 0 = i_{16} B_{16} - I_9 Z_{38} + I_8 Z_{37} + I_8 Z_{38} + i_{17} Z_{33} - i_3 (Z_{37} + Z_{38}); \\
 0 = i_{17} B_{17} - i_{18} Z_{30} + i_{14} Z_{31} + i_{16} Z_{33}; \\
 0 = i_{18} B_{18} - i_{14} Z_{29} - i_{17} Z_{30} + i_3 Z_{39} + i_{13} Z_{28}; \\
 0 = i_{19} B_{19} + I_2 Z_7 + i_5 Z_7 - i_1 Z_7 + i_5 Z_6 - I_1 Z_6 + i_6 Z_8 - I_1 Z_8 + i_7 Z_8 - I_3 Z_9 + i_6 Z_9 - \\
 - I_1 Z_9 + i_7 Z_9 + i_7 Z_{12} + i_{12} Z_{25} - i_{11} Z_{25} + i_{12} Z_{24} - I_7 Z_{24} + i_{12} Z_{26} + i_3 Z_{11} - Z_{24} i_{11}.
 \end{cases} \quad (7)$$

На основі розв'язку системи рівнянь (7) за допомогою пакета прикладних програм "Matlab 6.2" був визначений повний опір на електродах транзисторів VT1 та VT2. Розрахунки повного опору структури дозволяють отримати всі необхідні теоретичні залежності. На рис. 4. показана теоретична та експериментальна залежність активної та реактивної складової повного опору. При проведенні експериментальних досліджень в якості еталонного засобу вимірювань використано гаусметр 1-ST із роздільною здатністю 0,1 Гс (0,01 мТл) в діапазоні магнітної індукції від 0,1 Гс до 19 999,9 Гс (від 0,01 мТл до 1999, 99 мТл).

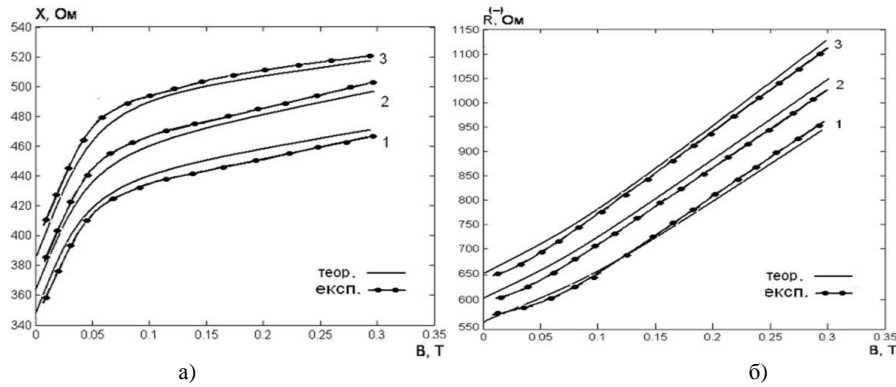


Рисунок 4 – Залежності реактивної (а) та активної (б) складової повного опору при напругах живлення: 1-5В, 2-5,5В, 3- 6В

Відповідно до методу Ляпунова визначена функція перетворення пристрою, що являє собою залежність частоти генерації від величини магнітної індукції. Аналітична залежність функції перетворення:

$$F = \frac{1}{4} \frac{V_1 R_H(B) + \sqrt{V_1^2 R_H^2(B) - 4R_H^2(B)(C_H^2(B)C_{be3} + C_H^2(B)C_{bc1} + C_H(B)C_{be3}C_{bc1})V_2}}{\pi R_H^2(B)(C_H^2(B)C_{be3} + C_H^2(B)C_{bc1} + C_H(B)C_{be3}C_{bc1})}; \quad (8)$$

де $V_1 = C_{bc1}C_{be3}$; $V_2 = C_{bc1} + C_{be3}$.

Графічна залежність функції перетворення представлена на рис. 5 Чутливість перетворювача визначається на підставі виразу (9) і описується рівнянням

$$S_B^F = -0.5((2C_H(B)C_{be3}(\frac{\partial C_H(B)}{\partial B}) + 2C_H(B)C_{bc1}(\frac{\partial C_H(B)}{\partial B}) + (\frac{\partial C_H(B)}{\partial B})C_{be3}C_{bc1})V_2) + (V_1^2 - 4(C_H^2(B)C_{be3} + C_H^2(B)C_{bc1} + C_H(B)C_{be3}C_{bc1})V_2)^{1/2} R_H(B)\pi(C_H^2(B)C_{be3} + C_H^2(B) \times C_{bc1} + C_H(B)C_{be3}C_{bc1})) - \frac{1}{4}((V_1^2 - 4(C_H^2(B)C_{be3} + C_H^2(B)C_{bc1} + C_H(B)C_{be3}C_{bc1})V_2)^{1/2}) \times (\frac{\partial R_H(B)}{\partial B}) / (R_H^2(B)\pi(C_H^2(B)C_{be3} + C_H^2(B)C_{bc1} + C_H(B)C_{be3}C_{bc1})) - \frac{1}{4}(V_1 + (V_1^2 - 4(C_H^2(B) \times C_{bc1} + C_H(B)C_{be3}C_{bc1})V_2)^{1/2})(2C_H^2(B)C_{be3}(\frac{\partial C_H(B)}{\partial B}) + 2C_H(B)C_{bc1} \times (\frac{\partial C_H(B)}{\partial B})) + C_{be3}C_{bc1}(\frac{\partial C_H(B)}{\partial B}))) / (\pi R_H(B)(C_H^2(B)C_{be3} + C_H^2(B)C_{bc1} + C_H(B)C_{be3}C_{bc1})^2). \quad (9)$$

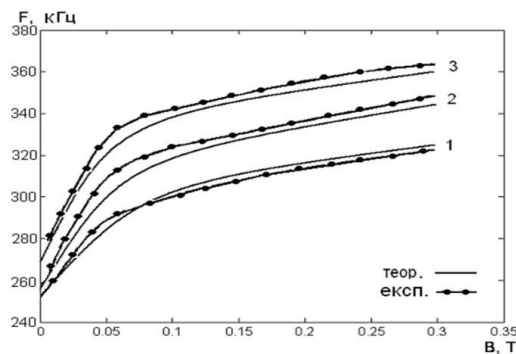


Рисунок 5 – Залежність частоти генерації частотного перетворювача від величини магнітної індукції при напругах живлення: 1-5В, 2-5,5В, 3-6В

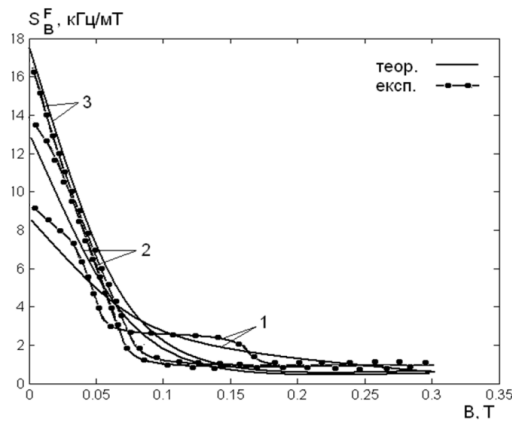


Рисунок 6 – Залежність частоти генерації частотного перетворювача від величини магнітної індукції при напругах живлення: 1-5В, 2-5,5В, 3-6В

Графік залежності чутливості представлений на рис.3.6. Як видно з графіка, найбільша чутливість пристрою лежить у діапазоні від 0,01 мТл до 300 мТл і складає 2-18 кГц/мТл.

Висновки

1. Запропонована мікроелектронна схема перетворювача магнітної індукції у частотний сигнал з магніточутливим елементом Холла.
2. Розроблена математична модель перетворювача на основі рівнянь Кірхгофа, що описують поведінку схеми, на основі якої отримані аналітичні вирази функції перетворення та рівняння чутливості.
3. Чутливість перетворювача складає 2-18 кГц/мТл в діапазоні магнітної індукції від 0,01 мТл до 300 мТл.

Список літератури

1. Popovic R.S. Hall effect devices. - Bristol ; Philadelphia : Institute of Physics, 2004. - 419 p.
2. Бараночников М.Л. Микромагнитоэлектроника. - М. : ДМК Пресс, 2001. - Т.1. - 364 с.
3. HAL805. Programmable Linear Hall Effect Sensor. Preliminary Data Sheet "Micronas", 2000. - 24 p.
4. Новицкий П.В., Кноринг В.Г., Гутников В.С. Цифровые приборы с частотными датчиками. - Л.: Энергия, 1970. - 424 с.
5. Осадчук О.В. Микроэлектронные частотные преобразователи на основе транзисторных структур с отрицательным сопротивлением. - Винница: « Универсум-Винница», 2000. - 303 с.
6. Осадчук В.С., Осадчук О.В. Реактивные свойства транзисторов и транзисторных схем. - Винница: «Универсум-Винница», 1999. - 275 с.
7. Ферри Д., Эйкерс Л., Гринич Э. Электроника ультрабольших интегральных схем: Пер. с англ. - М.: Мир, 1991.- 327 с.
8. PSPICE User's guide. MicroSim Corporation. La Cadena Drive, Laguna Hills, 1989. - 450 p.
9. Разевиг В.Д. Применение программ P-CAD и Pspice для схемотехнического моделирования на ПЗВМ. Выпуск 2. Модели компонент аналоговых устройств. - М.: Радио и связь, 1992. - 72 с.
10. Разевиг В.Д. Применение программ P-CAD и Pspice для схемотехнического моделирования на ПЗВМ. Выпуск 3. Моделирование аналоговых устройств. - М.: Радио и связь, 1992. -120 с.

Відомості про авторів

Осадчук Володимир Степанович – доктор технічних наук, професор, заслужений діяч науки і техніки України, професор кафедри електроніки ВНТУ, (0432)-58-20-45, osadchuk69@mail.ru.

Осадчук Олександр Володимирович – доктор технічних наук, професор, завідувач кафедри радіотехніки ВНТУ (0432)- 59-84-81, osadchuk69@mail.ru.

Мартинюк Володимир Валерійович – кандидат технічних наук, доцент, доцент кафедри електроніки ВНТУ, (0432)-59-80-13, toihtoide@i.ua.

Білілівська Ольга Петрівна – аспірант кафедри електроніки ВНТУ, (0432)-59-80-13, stovbcha@mail.ru.