БІОЛОГІЧНІ ТА МЕДИЧНІ ПРИЛАДИ І СИСТЕМИ

УДК 681.518.3: 535.243.2

В. Г. ПЕТРУК[1], О. Є. КВАТЕРНЮК[1], О. І. МОКАНЮК[2], С. М. КВАТЕРНЮК[1]

[1]Вінницький національний технічний університет
[2]Вінницький національний медичний університет ім. М.І.Пирогова

ОЦІНЮВАННЯ ВАЛІДНОСТІ ДІАГНОСТУВАННЯ ПОШКОДЖЕНЬ БІОТКАНИН ЗА ДОПОМОГОЮ ЦИФРОВОЇ КОЛОРИМЕТРІЇ З ВИКОРИСТАННЯМ НЕЙРОМЕРЕЖІ

Анотація: У роботі розроблено та проаналізовано експертну системи підтримки прийняття діагностичного рішення при дослідженні поверхневих пошкоджень біотканин за допомогою цифрової колориметрії з використанням нейромережі для обробки результатів. Перевірено можливості застосування архітектури нейромережі на основі багатошарового персептрона з одним прихованим шаром, багатошарового персептрона з двома прихованими шарами, імовірнісної нейромережі та мережі з радіальною базисною функцією. Після навчання нейромереж на основі тестових послідовностей досліджено результати продуктивності та помилок при їх роботі. Розроблено спеціалізоване програмне забезпечення для введення анкетних даних пацієнтів при судово-медичній експертизі біотканин шкіри людини у спільну базу даних з цифровими фотографіями ушкоджень. Проаналізовано операційні параметри діагностичного тесту, що дозволяють оцінити точність діагностики поверхневих пошкоджень біотканин на основі вдосконаленого методу цифрової колориметрії.

Ключові слова: цифрова колориметрія, біотканини, судово-медична експертиза

Аннотация: В работе разработана и проанализирована экспертная системы поддержки принятия диагностического решения при исследовании поверхностных повреждений биотканей с помощью цифровой колориметрии с использованием нейросети для обработки результатов. Проверены возможности применения архитектуры нейросети на основе многослойного персептрона с одним скрытым слоем, многослойного персептрона с двумя скрытыми слоями, вероятностной нейросети и сети с радиальной базисной функцией. После обучения нейронных сетей на основе тестовых последовательностей исследованы результаты производительности и ошибок при их работе. Разработано специализированное программное обеспечение для ввода анкетных данных пациентов при судебно-медицинской экспертизе биотканей кожи человека в общую базу данных с цифровыми фотографиями повреждений. Проанализированы операционные параметры диагностического теста, позволяющие оценить точность диагностики поверхностных повреждений биотканей на основе усовершенствованного метода цифровой колориметрии.

Ключевые слова: цифровая колориметрия, биоткани, судебно-медицинская экспертиза

Abstract: In this paper, we designed and analyzed the expert system of diagnostic adoption of sup-port solutions in the study of surface damage biological tissues by means of digital colorimetry using a neural network to process the results. The researchers tested the possibility of using neural network architecture based on multilayer perceptron with one hidden layer; multilayer perceptron with two hidden layers, probabilistic neural networks and network with radial basis function. After learning neural network based on test sequences we have studied performance results and errors in their work. We have developed specialized software for administration personal data of patients in forensic examination of biological tissues of human skin and combination of data in common database with digital photos injuries. The researchers analyzed the operating parameters of a diagnostic test to assess the diagnostic accuracy of surface damage biological tissues based on advanced of digital colorimetry method.

Key words: digital colorimetry, biological tissue, forensic examination

Вступ

Вдосконалений спосіб цифрової колориметрії поверхневих пошкоджень біотканин у судовій медицині дозволяє документально зареєструвати ушкодження біотканини шкіри судово-медичним експертом та використати результати досліджень у якості доказової бази [8]. Спеціалізоване алгоритмічно-програмне забезпечення засобу діагностики біотканин шкіри людини дозволяє визначити гістограму кольорів зображення патологічної біотканини шкіри при механічному ураженні тупим предметом. Отримана гістограма кольорів разом з іншими біомедичними параметрами пацієнта подається на вхід експертної системи підтримки прийняття діагностичного рішення. Таким чином, засіб діагностики дозволяє формалізувати відомі методики візуального експертного оцінювання терміну ушкодження біотканини шкіри людини внаслідок травми тупим предметом, що необхідно для зменшення суб'єктивності проведення досліджень у судово-медичній діагностиці.

Актуальність роботи обумовлена необхідністю підвищення точності діагностики параметрів поверхневих пошкоджень біотканин шкіри людини з урахуванням їх оптико-фізичних параметрів за рахунок вдосконалення методу цифрової колориметрії.

Метою дослідження є розробка та аналіз роботи експертної системи підтримки прийняття діагностичного рішення при дослідженні поверхневих пошкоджень біотканин за допомогою цифрової колориметрії з використанням нейромережі для обробки результатів, а також аналіз параметрів валідності діагностики.

Експертна система підтримки прийняття діагностичного рішення при дослідженні поверхневих пошкоджень біотканин

Для підтримки прийняття діагностичного рішення стосовно поверхневих пошкоджень біотканин у судовій медицині розроблено експертну систему для визначення часу, що минув з моменту нанесення травми тупим предметом на основі результатів обробки параметрів кольору ушкодженої ділянки у вигляді гістограми $N_{j\%}$ з відносною кількістю пікселів, яку займають елементи зображення з

координатами кольору близькими до кожного з елементів шкали зразків, а також інших біомедичних параметрів з анкет пацієнтів, які є змінними множинного типу відповідних множин з фіксованою кількістю елементів, а саме:

- вік: 0-20 років ($x1_1$), 20-30 років ($x1_2$), 30-40 років ($x1_3$), 40-50 років ($x1_4$), 50-60 років ($x1_5$), 60 років ($x1_6$) $X1 = \{x1_1, x1_2, \ldots, x1_6\}$;
- стать: чоловік ($x2_1$), жінка ($x2_2$) $X2 = \{x2_1, x2_2\}$;
- індекс Rees-Eisenck: до 96 пікнічний тип ($x3_1$), 96-106 нормастенічний тип ($x3_2$), понад 106 астенічний тип ($x3_3$) $X3 = \{ x3_1, x3_2, x3_3\}$;
- об’єкт експертизи: жива особа ($x4_1$)/ труп ($x4_2$) $X4 = \{x4_1, x4_2\}$;
- локалізація: голова, шия ($x5_1$), плеча, надпліччя ($x5_2$), тулуб ($x5_3$), стегно, сідниця ($x5_4$), передпліччя, кисть ($x5_5$), гомілка, стопа($x5_6$) $X5 = \{x5_1, x5_2, \ldots, x5_6\}$;
- форма: округла ($x6_1$), овальна($x6_2$), неправильна овальна ($x6_3$) $X6 = \{x6_1, x6_2, x6_3\}$.

На виході експертна система повинна сформувати діагностичне рішення у вигляді тривалості часу, що минув з моменту нанесення травми тупим предметом, який є елементом наступної множини:
Тривалість: до 1 години (t_1), від 1 до 3 годин (t_2), від 6 до 12 годин (t_3), від 12 до 24 годин (t_4), від 24 до 48 годин (t_5), від 48 до 72 годин (t_6), від 72 до 96 годин (t_7), понад 96 годин (t_8) $T = \{ t_1, t_2, \ldots, t_8\}$.

Розроблено спеціалізоване програмне забезпечення Multima-Expert для введення анкетних даних пацієнтів при судово-медичній експертизі біотканин шкіри людини та поєднання їх у єдиній базі даних з цифровими фотографіями ушкоджень. Розроблена програма має можливість експортувати анкетні дані пацієнтів для друку та долучення до медичних карток.

Для побудови експертної системи підтримки прийняття діагностичного рішення про тривалість часу, що минув з моменту нанесення травми тупим предметом у судово-медичній експертизі за допомогою нейромережі використаємо пакет STATISTACA 6.0 Neural Networks [1].

Для побудови нейромережі перевіримо можливості застосування архітектури мережі на основі:

- багатошарового персептрона з одним прихованим шаром – БШП1 (рис.1);
- багатошарового персептрона з двома прихованими шарами – БШП2;
- імовірнісної нейромережі - ІНМ;
- мережа з радіальною базисною функцією – РБФ.

Для навчання нейромережі використовуємо дані тривалості часу, який було обчислено судово-медичним експертом на основі результатів опитування пацієнтів $T_{in} = \{ t_1, t_2, \ldots, t_8\}$. Після навчання нейромереж протягом 500 ітерацій на основі тестових послідовностей отримано наступні результати продуктивності та помилок їх роботи (див. табл. 1).

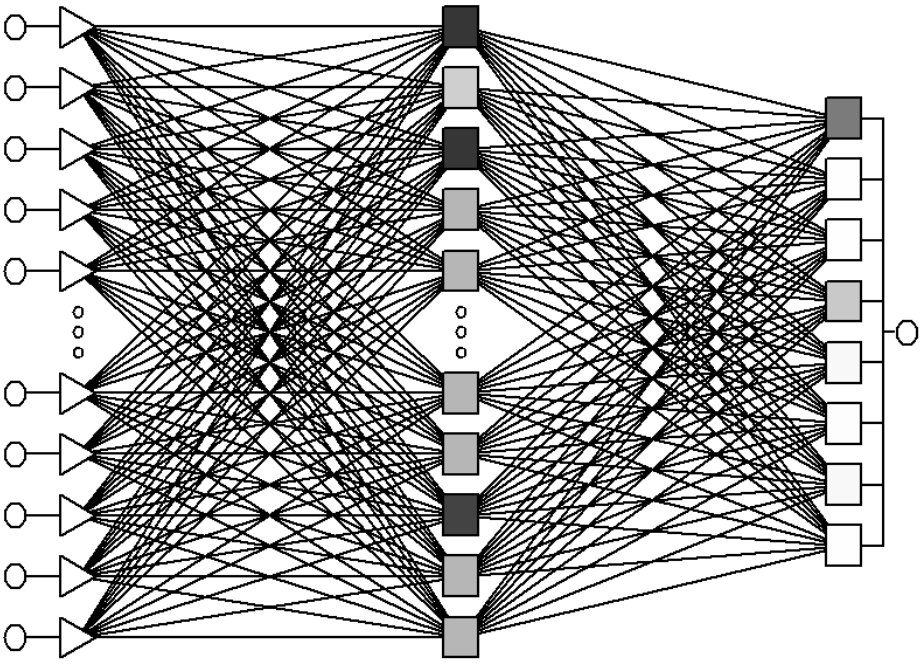


Рисунок 1 – Архітектура нейромережі на основі багатошарового персептрона з одним прихованим шаром

Таблиця 1 – Результати продуктивності та помилок нейромереж

Архітектура	Продуктив. навч.	Продуктив. контр.	Продуктив. тест.	Помилка при навчанні, %	Контрольна помилка, %	Тестова помилка, %
БШП1	0,986111	0,485714	0,371429	0,372299	5,978848	6,183768
БШП2	0,986111	0,485714	0,400000	0,153704	5,539583	5,836819
ІНМ	0,722222	0,371429	0,314286	0,203663	0,2883422	3,082088
РБФ	0,555556	0,428571	0,285714	0,255504	0,2848522	3,031888

Оцінювання валідності біомедичного діагностування поверхневих пошкоджень біотканин на основі вдосконаленого методу цифрової колориметрії

Проаналізуємо точність діагностики поверхневих пошкоджень біотканин на основі вдосконаленого методу цифрової колориметрії використавши набір даних тестування: T_{in} – тривалість часу, який було обчислено судово-медичним експертом на основі результатів опитування пацієнтів (використовується у якості істинного значення часу), T_{out} – тривалість часу отримана на виході експертної системи. T_{in} та T_{out} є елементами множини T = { t_1, t_2,…, t_8}.

У результаті діагностування необхідно визначити чи перевищує тривалість часу певне фіксоване значення T_{const}. У цьому випадку при діагностиці можливі такі варіанти:

$$\begin{cases} P_1 = 1 \quad якщо \quad T_{in} \geq T_{const}, \quad інакше \quad P_1 = 0, \\ P_2 = 1 \quad якщо \quad T_{in} < T_{const}, \quad інакше \quad P_2 = 0, \\ P_3 = 1 \quad якщо \quad T_{out} \geq T_{const}, \quad інакше \quad P_3 = 0, \\ P_4 = 1 \quad якщо \quad T_{out} < T_{const}, \quad інакше \quad P_4 = 0, \end{cases} \tag{1}$$

де P_1, P_2 – змінні множинного типу {0;1}, які відповідають дійсному значенню тривалості T_{in},
P_3, P_4 – змінні множинного типу {0;1}, які відповідають значенню тривалості сформованому на виході експертної системи.

При цьому можливо розрахувати змінні A, B, C, D, які відповідають усім можливим подіям у процесі діагностики:

$$\begin{cases} A = 1 \quad якщо \left(\left(P_1 = 1 \right) \wedge \left(P_3 = 1 \right) \right), \quad інакше \quad A = 0, \\ B = 1 \quad якщо \left(\left(P_2 = 1 \right) \wedge \left(P_3 = 1 \right) \right), \quad інакше \quad B = 0, \\ C = 1 \quad якщо \left(\left(P_1 = 1 \right) \wedge \left(P_4 = 1 \right) \right), \quad інакше \quad C = 0, \\ D = 1 \quad якщо \left(\left(P_2 = 1 \right) \wedge \left(P_4 = 1 \right) \right), \quad інакше \quad D = 0, \end{cases} \tag{2}$$

де A, B, C, D – змінні множинного типу {0;1},
A – відповідає події, коли у процесі діагностування отримано істинно-позитивний результат (у результаті діагностування отримано позитивний результат діагностування, а дійсне значення теж було позитивним),
B – відповідає події, коли у процесі діагностування отримано хибно-позитивний результат (у результаті діагностування отримано позитивний результат діагностування, а дійсне значення ознаки було негативним),
C – відповідає події, коли у процесі діагностування отримано хибно-негативний результат (у результаті діагностування отримано негативний результат діагностування, а дійсне значення ознаки було позитивним),
D – відповідає події, коли у процесі діагностування отримано істинно-негативний результат (у результаті діагностування отримано негативний результат діагностування, а дійсне значення ознаки теж було негативним).

Розраховуємо змінні P_1, P_2, P_3, P_4 та, відповідно, A, B, C, D за формулами (1), (2) для набору даних тестування. Результати аналізу подій для біомедичного діагностування поверхневих пошкоджень біотканин наведено у табл. 2.

Таблиця 2 – Аналіз подій при діагностуванні поверхневих пошкоджень біотканин

N	T_{in}	T_{out}	Дійсне значення		Результати діагностики		A	B	C	D
			P_1 (+)	P_2 (-)	P_3 (+)	P_4 (-)				
1.	1	1	0	1	0	1	0	0	0	1
2.	1	1	0	1	0	1	0	0	0	1
3.	5	5	1	0	1	0	1	0	0	0
4.	2	4	0	1	1	0	0	1	0	0
5.	8	8	1	0	1	0	1	0	0	0
6.	7	2	1	0	0	1	0	0	1	0
	...									
583	8	8	1	0	1	0	1	0	0	0

Відповідно кількість подій кожного типу для біомедичного діагностування на основі цифрової колориметрії буде $a = \sum_{i=0}^{n} A_i$, $b = \sum_{i=0}^{n} B_i$, $c = \sum_{i=0}^{n} C_i$, $d = \sum_{i=0}^{n} D_i$. Отримані значення *a*, *b*, *c*, *d* вносимо в таблицю спряженості (2x2 contingency table) для порівняння двох груп за бінарною ознакою (табл. 3.).

Таблиця 3 – Порівняння результатів діагностування та дійсного значення ознаки

Результати діагностування	Дійсне значення ознаки («золотий» стандарт)		Разом
	Позитивне значення (+)	Негативне значення (-)	
Позитивний результат діагностування (+)	істинно-позитивний результат (true positive (TP)) a = 350	хибно-позитивний результат (false positive (FP)) b = 11	$a + b$ = 361
Негативний результат діагностування (-)	хибно-негативний результат (false negative (FN)) c = 8	істинно-негативний результат (true negative (TN)) d = 214	$c + d$ = 222
Разом	$a + c$ = 358	$b + d$ = 225	a + b + c + d=583

На підставі отриманих даних розраховуємо показники валідності діагностичного тесту. Прогнозна валідність – інформація про методику (тест), що характеризує ступінь точності і обґрунтованості судження про якість діагностики за її результатом через певний час після вимірювань. Справжня цінність (міра довіри) дослідження визначається співвідношенням чутливості та специфічності. У деяких практичних випадках доцільна (за високої специфічності) надлишкова чутливість, що пов’язана зі зростанням параметра гіпердіагностики. У іншому випадку, при зменшенні чутливості зростає параметр гіподіагностики.

На основі порівняльної таблиці визначаємо операційні параметри діагностичного тесту [2-7]:

– діагностична чутливість (sensitivity або true positive rate (*TPR*)), частка осіб з позитивним результатом діагностування серед осіб з справжнім позитивним значенням ознаки

$$TPR = a/\left(a+c\right) = 350/(350+8) = 0,978\,, \tag{4}$$

– діагностична специфічність (specificity (*SPC*) або true negative rate), частика осіб з негативним результатом діагностування серед осіб з справжнім негативним значенням ознаки

$$SPC = d/\left(d+b\right) = 214/\left(214+11\right) = 0,951\,, \tag{5}$$

– діагностична ефективність, середнє між *TPR* та *SPC*

$$E = \left(TPR + SPC\right)/2 = (0,978 + 0,951)/2 = 0,964\,, \tag{6}$$

– позитивне прогнозне значення тесту (precision or positive predictive value (*PPV*)) – ймовірність наявності позитивного значення ознаки при позитивному результаті діагностування

$$PPV = a/(a+b) = 350/(350+11) = 0,97, \quad (7)$$

– негативне прогнозне значення тесту (negative predictive value (*NPV*)) – ймовірність наявності негативного значення ознаки при негативному результаті діагностування

$$NPV = d/(\tilde{n}+d) = 214/(8+214) = 0,964. \quad (8)$$

– загальна статистична валідність (validity або accuracy (ACC))

$$ACC = (a+d)/(a+b+c+d) = (350+214)/583 = 0,967. \quad (9)$$

– індекс F1 (F1 score)

$$F1 = 2a/(2a+b+c) = 2\cdot 350/(2\cdot 350+11+8) = 0,973. \quad (10)$$

Ймовірність помилки першого роду, показник гіпердіагностики (α errors, type I error probability, fall-out or false positive rate (FPR)):

$$\alpha = b/(b+d) = 11/(11+214) = 0,049. \quad (11)$$

Ймовірність помилки другого роду, показник гіподіагностики (β errors, type II errors probability, false negative rate (FNR)):

$$\beta = c/(a+c) = 8/(350+8) = 0,022. \quad 12)$$

Достовірність (adequacy):

$$D = 1-\alpha-\beta = 1-0,049-0,022 = 0,929. \quad (13)$$

Відношення правдоподібності позитивного результату (likelihood ratio positive):

$$LR_+ = TPR/(1-SPC) = 20. \quad (14)$$

Відношення правдоподібності негативного результату (likelihood ratio negative):

$$LR_- = (1-TPR)/SPC = 0,023. \quad (15)$$

Відношення шансів для діагностичного тесту (diagnostic odds ratio (DOR))

$$DOR = LR_+/LR_- = 20/0,023 = 851,14. \quad (16)$$

Висновки

У роботі розроблено та проаналізовано роботу експертної системи підтримки прийняття діагностичного рішення при дослідженні поверхневих пошкоджень біотканин за допомогою цифрової колориметрії з використанням нейромережі для обробки результатів. Перевірено можливості застосування архітектури нейромережі на основі багатошарового персептрона з одним прихованим шаром, багатошарового персептрона з двома прихованими шарами, імовірнісної нейромережі та мережі з радіальною базисною функцією. Після навчання нейромереж протягом 500 ітерацій на основі тестових послідовностей досліджено результати продуктивності та помилок їх роботи. Найменшу кількість помилок 0,288% забезпечила імовірнісна нейромережа. Проаналізовано операційні параметри діагностичного тесту, що дозволяють оцінити точність діагностики поверхневих пошкоджень біотканин на основі вдосконаленого методу цифрової колориметрії. Отримано показники діагностичної чутливості

0,978, діагностичної специфічності 0,951, загальної статистичної валідності 0,967, що достатні для використання розробленого засобу діагностики для потреб судово-медичної експертизи.

Список використаних джерел

1. Нейронные сети. STATISTACA Neural Networks: Методология и технологии современного анализа данных / Под редакцией В.П. Боровикова. – 2-е изд., перераб. и доп. – М.: Горячая линия. – Телеком, 2008. – 392 с.
2. Loong T-W. Understanding sensitivity and specificity with the right side of the brain / T-W. Loong // British Medical Journal. – 2003. – V.327. – P. 716-719.
3. Altman D.G. Statistics Notes: Diagnostic tests 1: sensitivity and specificity / D. G. Altman, J. M. Bland // British Medical Journal. – 1994. – V.308. – P. 1552.
4. ISO 3534-1:2006 Statistics - Vocabulary and symbols - Part 1: General statistical terms and terms used in probability. Committee ISO/TC 69/SC 1, Edition: 1, Publication date: 2010-10-25. – 105 p.
5. Реброва О. Ю. Статистический анализ медицинских данных. Применения пакета прикладных программ STATISTICA. – М.: МедиаСфера, 2002. – 312 с.
6. Альтман Д. Г. Пересмотренный вариант единых стандартов представления результатов рандомизированных контролируемых испытаний (CONSORT): разъяснения и перспективы дальнейшего усовершенствовании / [Альтман Д. Г., Шульц К.Ф., Мохер Д. и др.] // Международный журнал мед. практики. – 2001. – №4. – С.5-34.
7. Мінцер О.П. Інформаційні технології в охороні здоров’я і практичній медицині: У 10 кн. Кн. 5. Оброблення клінічних і експериментальних даних у медицині: Навч. посіб. / О.П. Мінцер, Ю.В. Вороненко, В.В. Власов. – К.: Вища школа, 2003. – 350 с.
8. Спосіб визначення і реєстрації кольору та розмірів ушкоджень в судово-медичних дослідженнях / [В. Г. Петрук, О. І. Моканюк; О. Є. Кватернюк та ін.]// Патент України №70759 МПК7 G01N 21/21 / заявл. 05.12.2011; опубл. 25.06.2012; Бюл. № 12. – 6 с.

Відомості про авторів

Петрук Василь Григорович – д.т.н., професор, директор Інституту екологічної безпеки та моніторингу довкілля, Вінницький національний технічний університет, Хмельницьке шосе, 95, м. Вінниця;

Кватернюк Олена Євгенівна – здобувач кафедри екології і екологічної безпеки, Вінницький національний технічний університет;

Моканюк Олександр Іванович – к.м.н., доцент кафедри патологічної анатомії, судової медицини та права, Вінницький національний медичний університет ім. М. І. Пирогова, Вінниця;

Кватернюк Сергій Михайлович – к.т.н., с.н.с. кафедри екології і екологічної безпеки; вінницький національний технічний університет.

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

УДК 004.93:159.95

О. В. БІСІКАЛО, О. В. ЯХИМОВИЧ

Вінницький національний технічний університет, м. Вінниця

ЗНАХОДЖЕННЯ КЛЮЧОВИХ СЛІВ АНГЛОМОВНОГО ТЕКСТУ ЗА ДОПОМОГОЮ ІНСТРУМЕНТАЛЬНИХ ЗАСОБІВ ПАКЕТУ DKPRO CORE

Анотація. Запропоновано новий підхід до визначення ключових слів, який базується на знаходженні зв'язків між словоформами англомовного тексту за допомогою інструментальних можливостей пакету DKPro Core. Проілюстрований аналізом прикладів застосування підхід спрямовано на розв'язання задач ефективної обробки текстових документів – індексування, реферування, кластеризації та класифікації. В результаті теоретичного та експериментального дослідження встановлено, що розроблений підхід забезпечує кращу релевантність знаходження ключових слів у порівнянні з аналогами за критеріями точності та повноти.

Ключові слова: ключові слова, англійська мова, релевантність, лінгвістичний пакет, DKPro Core, синтаксичний аналіз

Аннотация. Предложен новый поход к определению ключевых слов, основанный на нахождении связей между словоформами англоязычного текста с помощью инструментальных возможностей пакета DKPro Core.Проиллюстрированный анализом примеров применения подход направлен на решение задач эффективной обработки текстовых документов – индексирования, реферирования, кластеризации и классификации. В результате теоретического и экспериментального исследования установлено, что разработанный подход обеспечивает более высокую по сравнению с аналогами релевантность по критериям точности и полноты.

Ключевые слова: ключевые слова, английский язык, релевантность, лингвистический пакет, DKPro Core, синтаксический анализ

Abstract. It is proposed a new approach to determining the keywords based on finding connections between word forms of the English text with the instrumental capabilities of package DKPro Core. The approach, which illustrated with examples of analysis, aimed at solving problems of efficient processing of text documents – indexing, abstracting, clustering and classification. As a result of theoretical and experimental studies it is found that the developed approach ensures better relevance of the keywords compared to similar criteria for accuracy and completeness.

Key words: keywords, English, method, relevance, linguistic package, DKPro Core, syntactic analysis.

Вступ

Якість виділення ключових слів з тексту грає важливу роль у бібліотечній справі, лексикографії та термінознавстві, а також в задачах інформаційного пошуку. В даний час обсяги і динаміка інформації, яка підлягає обробці в цих областях, роблять особливо актуальною задачу автоматичного визначення ключових слів, які можуть використовуватися для створення і розвитку термінологічних ресурсів, а також для ефективної обробки документів: індексування, реферування, кластеризації та класифікації.

Актуальність

У переважній більшості доступних лінгвістичних систем, що орієнтовані на обробку природномовних текстів, присутні функції автоматичного виділення ключових слів. В основу реалізації таких функцій покладено відомі методи визначення ключових слів, які діляться на лінгвістичні та статистичні. Лінгвістичні методи ґрунтуються на значеннях слів, зокрема використовують онтології та семантичні дані про слово. На жаль, методи такого класу ресурсоємні на ранніх етапах: розробка онтологій, наприклад, вельми трудомісткий процес [1]. З іншого боку, статистичні методи супроводжуються значними обсягами «вербального шуму», який суттєво впливає на якість визначення ключових слів. Тому найбільш перспективними для дослідження, на думку авторів, є гібридні методи, для яких швидкість статистичної обробки тексту підсилюється можливостями сучасних лінгвістичних пакетів.

Мета

Мета роботи полягає у підвищенні релевантності визначення ключових слів англомовного тексту за рахунок врахування інформації щодо синтаксичних зв'язків між словами у реченнях тексту.

Задачі

1. Обгрунтування підходу до визначення ключових слів тексту на основі інформації про синтаксичні зв'язки.
2. Програмна реалізація підходу та експериментальна оцінка складових релевантності.

Інформаційна оцінка синтаксичного аналізу тексту для задачі визначення ключових слів

Розглянемо задачу визначення ключових слів тексту як певну інформаційну технологію, що має на вході текст, а на виході – множину з *l* ключових слів $W^k = \{w_1^k, ..., w_l^k\}$. Без применшення загальності будемо вважати, що текст *T* складається з *m* різних слів, а в окреме його *j*-те речення з *k* налічує *n* слів з *m* можливих, причому $m >> n$ та $m >> l$. Більшість відомих методів визначення ключових слів тексту беруть за основу частотний словник тексту, який фактично є списком або упорядкованою множиною пар $D = \{< w_i,\ f_i >\},\ i = \overline{1,m}$, де w_i – одне слово з *m*, а f_i – його частота ($f_i \geq f_{i+1}$,

$i = \overline{1, m-1}$), що визначена для *T*. За певною фільтрацією окремих незначущих категорій слів ключовими вважають перші *l* слів зі списку *D*, тобто, дещо спрощено маємо $W^k = \{w_1, ..., w_l\}$.

Проте результати парсерингу природних мов за допомогою сучасних лінгвістичних пакетів дозволяють на доступному програмному рівні [2] оперувати синтаксичними зв'язками між словами окремого речення. Окрім того, можливості цих пакетів дозволяють суттєво зменшити значення *m* шляхом об'єднання слів у словоформи, а останні – у леми та стеми. Отже, необхідно з'ясувати, які формальні переваги для визначення $W^k = \{w_1^k, ..., w_l^k\}$ надасть нам програмно-лінгвістичне забезпечення процедури синтаксичного аналізу всіх речень тексту *T*.

З інформаційної точки зору розуміння сенсу речення окремим суб'єктом супроводжується розпізнаванням а) окремих слів, з яких воно складається та б) зв'язків між парами цих слів з відповідною побудовою дерева таких зв'язків [3]. Вважатимемо, що всі ці процеси відбуваються шляхом порівняльного аналізу та залучення інформації з деякої загальнолінгвістичної бази знань суб'єкта розуміння. Якщо кожен з цих етапів супроводжується збільшенням інформації, то приймаємо робочу гіпотезу:

- Рівень загального розуміння тексту *T* може змінюватися від мінімально можливого до максимального в залежності від обсягу та інших параметрів загальнолінгвістичної бази знань суб'єкта;
- Якість визначення $W^k = \{w_1^k, ..., w_l^k\}$ пропорційна рівню загального розуміння тексту, що має підтверджуватися формальними ознаками.

Нехай будь-яке *j*-те речення з *k* складається з *n* різних слів, що не є досить жорстким обмеженням. Тоді зв'язне дерево парних залежностей такого речення налічує або *n–1* гілок, якщо не брати до уваги зворотну залежність між підметом та присудком, або *n* – якщо брати. Відповідно загальна кількість слів цього речення для подальшого поглибленого аналізу збільшується або до *2×n–2* або до *2×n*. Проте таке збільшення відбувається нерівномірно – для всіх не термінальних (кінцевих) вузлів дерева частоти відповідних слів не змінюються, а для термінальних (проміжних) можуть зрости суттєво. В таблиці 1 показані випадки зміни частот слів, які позначаються літерами a, b, …, f за порядком застосування, з урахуванням парних залежностей для різних типів речення.

Таблиця 1 – Аналіз збільшення частоти значимих слів унаслідок урахування парних залежностей для різних типів речення

№ з/п	Склад речення / кількість слів	Тип речення та граф його дерева залежностей	Частотна формула	Кінцева частота
1.	Ab / 2	Словосполучення (Коріння дерева)	A+b	2
2.	Abc / 3	Лінійна трійка (Бережи скарби природи)	A+2b+c	4
3.	Abcd / 4	Лінійна четвірка (Отримав переклад слова дивного)	A+2b+2c+d	6
4.	Abcde / 5	Розгалудження (Густий ліс нізвідки завершився провалля)	A+2b+c+3d+e	8
5.	Abcdef / 6	Група підмета (Сині примружені очі коханого говорили багато)	A+b+4c+d+2e+f	10
6.	Abcdef / 6	Група присудка (Досвідчений кінь борозну швидко відчує нюхом)	A+2b+c+d+4e+f	10
7.	Abcdef / 6	Обидві групи (Старий дід Еол зобрав всіх вітрів)	A+3b+c+2d+e+2f	10

Проведений аналіз на рівні одного речення показує, що збільшуються частоти тих слів (підкреслені у 3-му стовпчику таблиці), які потенційно можуть належати до множини ключових. Проведемо формальну оцінку такого збільшення для накладених обмежень щодо наявності виключно різних слів у реченні та не врахуванням зворотної залежності між підметом та присудком:

1. Мінімальне збільшення відсутнє за умови знаходження *i*-го слова з *m* серед не термінальних (кінцевих) вузлів дерева кожного речення, де це слово зустрічається, тобто $f_i^{\min} = 0,\ f_i^{new} = f_i$ $i = \overline{1, m}$.

2. Якщо *i*-те слово знаходиться у кожному з *k* речень тексту та, окрім того, відповідає у кожному реченні найбільш розгалуженому термінальному вузлу, то максимальне збільшення частоти складає $f_i^{\max} = \sum_{j=1}^{k}(n_j - 2), i = \overline{1,m}$. Відповідно $f_i^{new} = f_i + f_i^{\max} = k + \sum_{j=1}^{k}(n_j - 2) = \sum_{j=1}^{k}(n_j - 1)$.

3. В загальному та більш реальному випадку $f_i = z \mid z \le k$, тобто *i*-те слово знаходиться у *z* реченнях з *k* маємо $f_i^{new} = z + \sum_{j=1}^{z}(n_j - 2) = \sum_{j=1}^{z}(n_j - 1)$ як оцінку зверху збільшення частоти *i*-го слова.

Очікується, що експериментальні дослідження мають підтвердити справедливість отриманих формальних оцінок процесу визначення ключових слів тексту у встановлених межах.

Програмна реалізація запропонованого підходу в DKPro Core та проведення експерименту

Для експериментальної перевірки результатів теоретичного аналізу було розроблене програмне забезпечення на основі DKPro Core.

DKPro Core – це набір програмних компонентів для обробки природної мови, що базується на Apache UIMA framework. Він був побудований з метою підвищення продуктивності дослідників, які працюють з автоматичним аналізом мови. Підхід DKPro Core полягає в тому, що дослідники повинні мати можливість зосередитися на своїх реальних наукових питаннях, а не на розробці технологій [4, 5].

Визначення ключових слів відбувається за кількома етапами:

а) створення багаторівневої розмітки тексту;

б) синтаксична розмітка, що враховує складні залежності між парами лем;

в) заміна займенників в отриманих парах на відповідні до них іменники;

г) розбиття пар на окремі слова і визначення кількості зв'язків;

д) вибір перших n слів з найбільшою кількістю зв'язків, де n – кількість потрібних ключових слів.

Аналогами розробленої програми можуть бути сайти SEO оптимізації, де є функція визначення ключових слів. Для даного експерименту вибрані сервіси: advego.ru/text/seo/, rise-top.com/keywordstext.php та seotool.by/analiz/seo/key-wordstext.php.

Для проведення експерименту було взято текст статті з 1460 слів «A new pattern for historical geography: working with enthusiast communities and public history» [6].

Упорядкований список ключових слів з їх позиціями x_i, що були задані автором: Participation (1), Public (2) history (3), Enthusiast (4) communities (5), Museums (6), Heritage (7). Результати знаходження ключових слів для власної розробки і аналогів наведено в таблиці 2. Перед правильно знайденими ключовими словами вказана позиція, на якій знаходиться це слово у авторському списку.

Таблиця 2 – Результати пошуку ключових слів

Слова задані автором		власна розробка		rise-top		advego		seotool	
1	Participation		work		historical		historical		historical
2	Public	5	community	4	enthusiast	4	enthusiast	4	enthusiast
3	history		geography	5	communities		for	5	communities
4	Enthusiast	1	participation	1	participation	5	community	1	participation
5	communities	4	enthusiast		geography		this		work
6	Museums		geographer		work	6	museum		geography
7	Heritage	6	museum		research		geography		new

Кількісними характеристиками релевантності отриманих результатів обрано повноту (за Жаккаром і абсолютну) і точність (за евклідовою і манхеттенською відстанями). Проведено інтерпретацію обраних критеріїв до умов задачі визначення ключових слів.

Повноту за Жаккаром визначено як частку від ділення кількості знайдених ключових слів на різницю кількості можливих ключових слів заданих автором і знайдених програмно (в даному випадку по 7) і кількості знайдених ключових слів. Абсолютна повнота знаходиться як відношення кількості правильно знайдених ключових слів до загальної кількості ключових слів. На рисунку 1 наведена гістограма повноти за Жаккаром і абсолютної для власної розробки і аналогів.

Точність за евклідовою відстанню визначається за формулою:

$$d_{\text{e}} = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}, \quad (1)$$

де n – кількість правильно визначених ключових слів;

x_i – позиція і-го ключового слова з n в авторському списку;

y_i – позиція і-го ключового слова з n в альтернативному списку, що знайдено програмно.

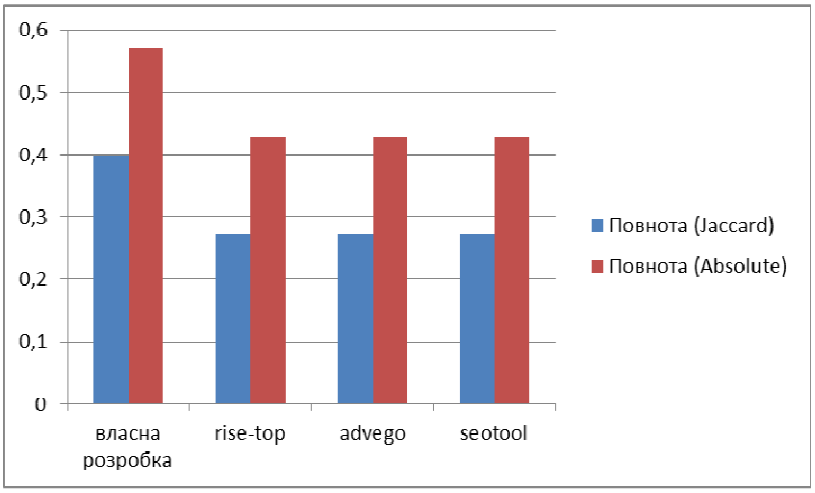


Рисунок 1 – Гістограма повноти за Жаккаром і абсолютної

Манхеттенська відстань визначається за формулою:

$$d_{\text{m}} = \sum_{i=1}^{n}\left|x_i - y_i\right|. \quad (2)$$

На рис. 2 наведена гістограма евклідової та манхеттенської відстані для власної розробки і аналогів.

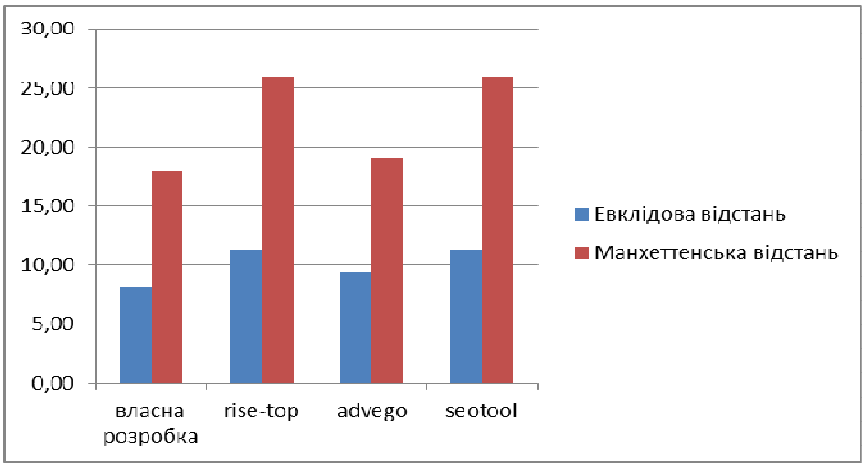


Рисунок 2 – Гістограма евклідової та манхеттенської відстані

Повнота знаходження ключових слів має бути якомога більшою, а відстань між позиціями ключових слів заданих автором і визначених програмно – якомога меншою. Як видно з гістограм власна розробка має кращі кількісні характеристики за різними метриками у порівнянні з аналогами – від 25% до 31,8% за повнотою, а також від 5,3% до 14% за точністю для тестового прикладу [6].

Висновки

1. Оскільки краща якість обробки тексту досягається лінгвістичними методами або ж при їх комбінації зі статистичними, систему автоматичного визначення ключових фраз з тексту природною мовою слід розробляти з використанням морфологічного словника (лексикону) і синтаксичних правил. Ці дані визначаються попередньо і зберігаються в базі даних. Текст підлягає обробці аналізатором, який виробляє інформацію про розділення тексту на абзаци, речення та окремі слова, що необхідно для подальшого оброблення. Кожне слово, виділене аналізатором, піддається морфологічному аналізу з метою побудови морфологічної інтерпретації, визначення основи слова і формування леми. На основі наявної інтерпретації тексту виконується побудова та наповнення синтаксичних груп і виявлення відношень між ними.

2. В роботі запропоновано підхід до визначення ключових слів, що базується на використанні додаткової інформації про складні залежності між членами англомовного речення. Для функціональної реалізації аналізатора тексту обрано популярний лінгвістичний пакет DKPro Core. Проведені експериментальні дослідження теоретичного обґрунтування підходу підтвердили його якісні та кількісні переваги у порівнянні з відомими аналогами. Для англомовного тексту обсягом 1460 слів отримано збільшення повноти визначення ключових слів (на 31,8% за Жаккаром та на 25% за абсолютним значенням) і покращення точності (на 14% за евклідовою і на 5,3% манхеттенською відстанями) у порівнянні з аналогами.

3. Якість отриманих результатів потенційно можна підвищити через окремий аналіз частин мови, оскільки ймовірність релевантності ключового слова, наприклад, іменника і прислівника буде відрізнятися. Окрім цього, варто оцінити збільшення частотних показників для ключових слів шляхом реалізації наявних в DKPro Core компонентів для визначення кореференційних зв'язків.

Список літератури

1. Ershov, Yu. S. (2014). Vydelenie kliuchevyh slov v russkoiazychnyh tekstah. Molodezhnyi nauchno-tehnicheskii vestnik. M.: FGBOU VPO "MGTU im. N. E. Baumana". Available: http://sntbul.bmstu.ru/file/out/730754. Last accessed 21.01.2015.

2. Bisikalo, O. V. (2013). Kontseptualna model systemy obraznoho analizu i syntezu pryrodno-movnykh konstruktsii. Matematychni mashyny i systemy, № 2, 184–187. ISSN 1028-9763.

3. Bisikalo, O. V. (2013). Formalni metody obraznoho analizu ta syntezu pryrodno-movnykh konstruktsii. Vinnytsia: VNTU, 316. ISBN 978-966-641-528-1.

4. Natural Language Processing: Integration of Automatic and Manual Analysis. (2014). Technischen Universität Darmstadt. Available: http://tuprints.ulb.tu-darmstadt.de/4151/1/rec-thesis-final.pdf. Last accessed 21.01.2015.

5. Gurevych, I., Muhlhauser, M., Muller, Ch., Steimle, J., Weimer, M., Zesch, T. (2007, February 9). Darmstadt Knowledge Processing Repository Based on UIMA. Available: https://www.ukp.tu-darmstadt.de/fileadmin/user_upload/Group_UKP/publikationen/2007/gldv-uima-ukp.pdf. . Last accessed 21.01.2015.

6. Geoghegan, H. (2014). A new pattern for historical geography: working with enthusiast communities and public history. Journal of Historical Geography, № 46. Available: http://ac.els-cdn.com/S0305748814001029/1-s2.0-S0305748814001029-main.pdf?_tid=d45ec9e6-ba7b-11e4-b562-00000aab0f01&acdnat=1424600353_48bb4ef54fffbc3b800698d175c3c052. Last accessed 21.01.2015.

Відомості про авторів

Бісікало Олег Володимирович – доктор технічних наук, професор, декан ФКСА, кафедра автоматики та інформаційно-вимірювальної техніки, Вінницький національний технічний університет, Хмельницьке шосе 95, м. Вінниця, Україна, 21000.

Яхимович Олександр Вікторович – кафедра автоматики та інформаційно-вимірювальної техніки, Вінницький національний технічний університет, Хмельницьке шосе 95, м. Вінниця, Україна, 21000.

УДК 519.7.007.004.9

Т. О. САВЧУК, А. В. КОЗАЧУК

Вінницький національний технічний університет

АВТОМАТИЗОВАНЕ ПРИЙНЯТТЯ РІШЕНЬ ЩОДО МАСШТАБУВАННЯ ХМАРНОГО ЗАСТОСУНКУ

Анотація. В статті запропоновано алгоритм автоматизованого прийняття рішень щодо масштабування хмарного застосунку на основі набору правил реактивного масштабування та використання оціночної функції ефективності застосування можливих варіантів масштабування. В результаті проведених досліджень була розроблена оціночна функція варіантів масштабування, що дозволяє отримати оцінку ефективності застосування інформаційної технології, яка базується на вартості утримання інфраструктури хмарного застосунку та на оцінці кількості користувачів, що перестануть користуватися хмарним застосунком через його перевантаженість. Розроблений алгоритм може бути використаний в якості складової інформаційної технології масштабування хмарного застосунку зі змінними піками навантаження.

Ключові слова: хмарні обчислення, масштабування хмарного застосунку.

Аннотация. В статье предложен алгоритм автоматизированного принятия решений по масштабированию облачного приложения на основе набора правил реактивного масштабирования и использования оценочной функции эффективности применения возможных вариантов масштабирования. В результате проведенных исследований была разработана оценочная функция вариантов масштабирования, позволяющая получить экономическую оценку эффективности на основаннии стоимости содержания инфраструктуры облачного приложения и оценке количества пользователей, которые перестанут пользоваться облачным приложением из-за его перегруженности. Прогноз количества потерянных пользователей производиться на основании расчета времени выполнения сетевых запросов. Разработанный алгоритм может быть использован в качестве составной части информационной технологии масштабирования облачного приложения с периодическими пиками нагрузки.

Ключевые слова: облачные вычисления, масштабирование облачных приложений.

Abstract. Article describes the algorithm of automatic decision making on scaling of a cloud application based on a set of rules of reactive scaling and on a function that estimates each possible way of scaling of a cloud application using calculation of lost user count and price of cloud hosting equipment. Article introduces evaluation function of scaling options that allows obtaining economic evaluation of the efficiency based on the cost of maintaining infrastructure in the hosting cloud and estimation of the number of users that stop using cloud applications due to overheating-intensity. The forecast number of orphaned users is made on the basis of calculating the network requests execution time. Developed algorithm can be used as part of a cloud application scaling technology.

Key words: cloud computing, scaling of a cloud application.

Вступ

Завантаженість сучасних веб-застосунків суттєво змінюється протягом доби. Однією з причин розміщення застосунків у хмарних хостингах є можливість проведення швидкого масштабування веб-застосунку в пікові періоди для уникнення перебоїв у його функціонуванні за рахунок виділення додаткових обчислювальних потужностей. В найпростішому випадку масштабування здійснюється в ручному режимі, при цьому оператору необхідно періодично перевіряти завантаженість хмарного застосунку та приймати рішення щодо доцільності проведення масштабування. Автоматизація процесу прийняття рішення щодо масштабування дозволяє швидше реагувати на зміну завантаженості обчислювальних потужностей на усуває необхідність в постійному ручному моніторингу роботи хмарного застосунка.

Постановка задачі

Нехай $\langle S, N, P \rangle$ – поточний стан інфраструктури хмарного застосунку, де S розмір віртуальної машини, N – кількість виділених віртуальних машин P – вартість утримання віртуальних машин протягом 1 хвилини, X – вектор прогнозів кількості мережевих запитів на наступні h хвилин, час після минулого масштабування $\tau_{масш}$, G – граф станів інфраструктури хмарного застосунку, $t_{спок}$ – проміжок часу після масштабування протягом якого нове масштабування не здійснюється, C – інформація про середній час виконання мережевого запиту при різних станах роботи хмарного застосунку.

Необхідно побудувати алгоритм, що на основі даних про стан хмарного застосунку повертає рішення щодо проведення масштабування хмарного застосунку $e = \langle S', \Delta N \rangle$, де S' – розмір віртуальної машини після масштабування, ΔN – приріст кількості віртуальних машин відносно стану інфраструктури хмарного застосунку до початку масштабування.

Аналіз існуючих рішень

Найпоширенішими підходами до автоматизованого масштабування хмарного застосунку є використання методу реактивного масштабування та масштабування за розкладом [1]. Недоліком цих методів є неможливість проведення масштабування до виникнення перезавантаження хмарного застосунку та можливі неточності розкладу, відповідно, що робить описані підходи неефективними при певних умовах функціонування хмарного застосунку.

Алгоритм прийняття рішення щодо масштабування хмарного застосунку

Введемо алгоритм прийняття рішень щодо масштабування хмарного застосунку, що базується на використанні методів реактивного або проактивного масштабування. Метод реактивного масштабування представляє собою набір правил, які виконуються коли існує гостра потреба в проведенні масштабування, проактивне масштабування здійснюється на основі прогнозу роботи хмарного застосунку, який може

бути неточним, тому при виборі остаточного рішення пріоритет надається результатам роботи методу реактивного масштабування.

Результатом роботи алгоритму є рішення про необхідність масштабування. При цьому масштабування може бути вертикальним (змінюється значення N) або горизонтальним (змінюється значення S). Позначимо результати масштабування хмарного застосунку парою $e = \langle S', \Delta N \rangle$. Додатні значення ΔN відповідають масштабуванню вгору, від'ємні – масштабуванню вниз. Результатом алгоритму може бути пара $\langle S, 0 \rangle$, що відповідає відсутності необхідності у масштабуванні. Вважатимемо, що такий результат еквівалентний пустій операції масштабування: $\langle S, 0 \rangle = null$.

Робота алгоритму прийняття рішення щодо масштабування хмарного застосунку базується на використанні під-алгоритмів, що представляють методи реактивного та проактивого масштабування. Вхідними даними цих алгоритмів є підмножини вхідних даних головного алгоритму, а результат їхньої роботи – рішення щодо проведення масштабування $\langle S', \Delta N \rangle$.

Алгоритм складається з наступних кроків:

Крок 1. Початок роботи.

Крок 2. Введення інформації про стан хмарного застосунку.

Крок 3. Присвоїти поточному рішенню щодо масштабування *e* результат роботи алгоритму реактивного масштабування.

Крок 4. Якщо *e* є пустим – виконати крок 5, інакше – крок 6.

Крок 5. Присвоїти поточному рішенню щодо масштабування *e* результат роботи алгоритму проактивного масштабування.

Крок 6. Вивести значення поточного рішення щодо масштабування *e*.

Крок 7. Кінець.

Схема алгоритму зображена на рис. 1.

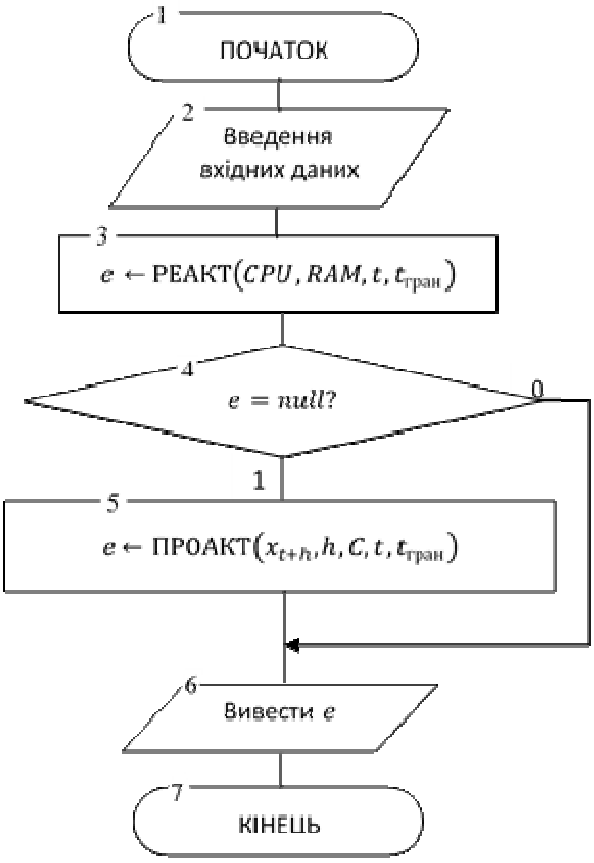


Рисунок 1 – Алгоритм прийняття рішень щодо масштабування хмарного застосунку

Метод реактивного масштабування

Класичним підходом до прийняття рішення щодо масштабування хмарного застосунку є застосування методу реактивного масштабування який передбачає визначення граничних показників використання пам'яті та процесора при досягненні яких відбувається масштабування [1]. Цей метод масштабу-

вання також називають статичним пороговим методом [2]. Ядро методу реактивного масштабуваня складає набір правил $r \in R$ вигляду

$$\text{якщо } X_r \text{ і } t \geq t_{\text{грак}},$$
$$\text{то здійстини масштабування } \langle f_r(S), \Delta N \rangle, t_{\text{грак}} \leftarrow t + \Delta t$$

де t – поточний момент часу, $t_{\text{грак}} = \tau_{\text{масш}} + \Delta t$ – момент часу в який закінчується період стабілізації після минулої операції масштабування, Δt – проміжок часу спокою після масштабування впродовж якого нові операції масштабування не здійснюють, необхідний для того щоб почати знімати показники завантаженості застосунку після зміни конфігурації хмарної інфраструктури, f – функція, що на основі поточного розміру віртуальної машини S повертає необхідний розмір віртуальної машини S', X_r – умова проведення масштабування, що може мати вигляд

завантаженість процесора > 80% або частка вільної пам'яті < 20%.

В більшості випадків [1] правила створюються попарно: для проведення масштабування вгору і вниз. При створенні правил слід враховувати обмеження хостингу хмарного застосунку: слідкувати, щоб кількість та розмір віртуальних машин були у дозволених межах, не використовувати переходи, що відсутні в графі станів інфраструктури хмарного застосунку, це призводить до ускладнення умов правил X_r.

При проведенні реактивного масштабування необхідно враховувати можливість виникнення осциляцій кількості віртуальних машин, для уникнення цього небажаного ефекту слід ретельно підбирати граничні показники правил масштабування. Для того, щоб метод не видавав команд масштабування у відповідь на одиничні піки навантаження в умовах правил X_r використовують усереднені показники за декілька останніх кроків часу, з цією ж метою деякі різновиди методу видають команду масштабування тільки за послідовного виконання умови X_r протягом кількох кроків часу.

В набір правил R можна додати правила з умовами, які виконуються в певний момент часу, без прив’язки до рівня використання системних ресурсів. Такі правила дають можливість провести масштабування напередодні подій, що збільшують завантаженість хмарного застосунку. Для мінімізації видатків можуть бути зроблені модифікації правил з врахуванням політики ціноутворення хостинга. Наприклад, при погодинній тарифікації немає сенсу здійснювати масштабування вниз протягом години після операції масштабування вгору.

Основними перевагами застосування методу реактивного масштабування до хмарного застосунку з періодичними піками навантаження є простота реалізації методу, його висока швидкодія та можливість початку експлуатації без збору додаткової інформації про інфраструктуру хмарного хостингу. До обмежень методу можна віднести реактивну природу, яка відображається в його назві – команда про проведення масштабування може бути віддана тільки тоді, коли виявлена нестача обчислюючих ресурсів, при цьому під час виділення нових ресурсів функціонування хмарного застосунку може відбуватися в нештатному режимі. Також, складною процедурою є встановлення порогових значень в правилах r, а також вибір дії масштабування $\langle S', \Delta N \rangle$ – через велику кількість комбінацій вибір оптимальної є нетривіальним.

Можливість використання методу без додаткових знань про функціонування хмарного застосунку в умовах конкретної інфраструктури хмарного хостингу дозволяють використовувати описаний метод як резерв для більш складних методів масштабування – при неможливості прийняття рішення щодо масштабування за допомогою власних ресурсів вони можуть використовувати результати роботи методу реактивного масштабування.

Метод прийняття рішення на основі прогнозу кількості мережевих запитів

За умови, що хмарний застосунок має стан $\langle S_i, N_i, P_i \rangle$ із графу станів інфраструктури хмарного застосунку, можливі *d(i)* варіанти проведення масштабування, де *d* – функція, що повертає степінь вершини графа. Позначимо множину можливих варіантів масштабування як E_i. Для вибору оптимального варіанту масштабування розрахуємо значення оціночної функції $G(e)$, $e \in E_i$ наступним чином:

$$G(e) = P_{\text{ІНФР}} + k * P_{\text{ВТРАТ}}, \quad (1)$$

Де $P_{\text{інфр}}$ – вартість утримання інфраструктури хмарного застосунку протягом наступних $t_{спок}$ хвилин, $P_{\text{втрат}}$ – вартість втрати одного користувача, k – кількість користувачів, втрачених через надмірну завантаженість хмарного застосунку протягом наступних $t_{спок}$ хвилин, e – дуга графу станів інфраструктури хмарного застосунку $\left(\langle S_i, N_i, P_i \rangle, \langle S_j, N_j, P_j \rangle, t_{ij}\right)$.

Вартість утримання інфраструктури визначається як сума вартості утримання до закінчення масштабування та після його закінчення, кожен з доданків розраховується як значення тарифу за одну хвилину P помножена на кількість хвилин $t_{спок} - t_{ij}$ та t_{ij}:

$$P_{ІНФР} = P_j t_{ij} + P_j\left(t_{СПОК} - t_{ij}\right). \quad (2)$$

Кількість користувачів, втрачених через надмірну завантаженість хмарного застосунку розраховується наступним чином:

$$k = \sum_{t=1}^{t_{спок}} v_t q_{ВТРАТ}(t). \quad (3)$$

де v_t – кількість користувачів хмарного застосунку в період часу $(t;t+1)$, $q_{втрат}(t)$ – імовірність того, що користувач перестане користуватися хмарним застосунком. Підставивши (2) та (3) в (1) отримаємо більш детальний вираз для обчислення оціночної функції:

$$G(e) = P_j t_{ij} + P_j\left(t_{СПОК} - t_{ij}\right) + P_{ВТРАТ} \sum_{t=1}^{t_{спок}} v_t q_{ВТРАТ}(t). \quad (4)$$

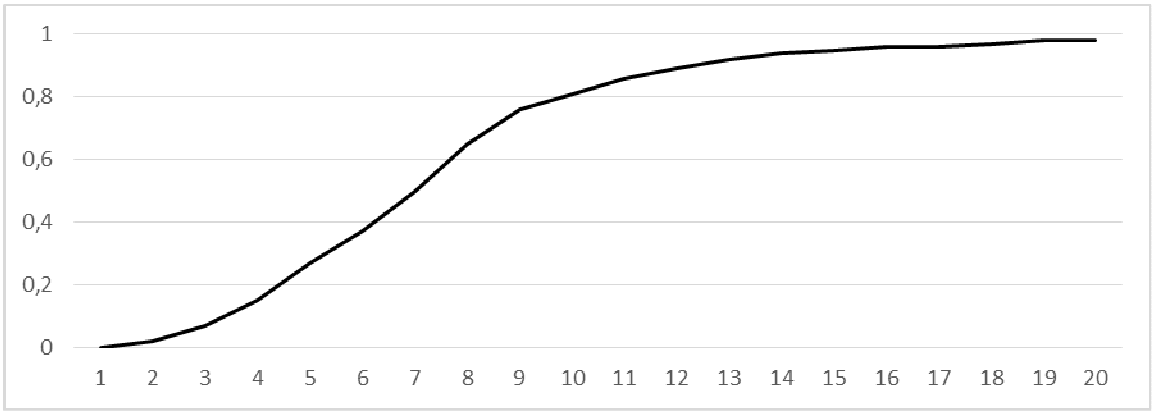


Рисунок 2 – Імовірність втрати користувача в залежності від часу запиту

Для визначення імовірності втрати користувача $q_{втрат}$ використаємо функцію розподілу імовірностей, що представляє залежність імовірності втрати користувача від часу виконання мережевого запиту $F(\tau)$ – імовірність втрати користувача зростає разом з часом очікування відповіді від хмарного застосунку. Дослідження [3-6] показують, що кількість користувачів веб-сторінок починає зменшуватись при $\tau \geq 2c$. Портал Pear1 hosting [3] при цьому наводить функцію розподілу імовірності втрати користувача, що показана на рисунку 2.

Позначимо випадкову величину, що відповідає часу виконання одного мережевого запиту хмарним застосунком за R. Статистичні характеристики R залежать від завантаженості хмарного застосунку в конкретний момент часу t та від стану інфраструктури хмарного застосунку. Визначимо імовірність втрати користувач через математичне сподівання часу виконання мережевого запиту $M(R)$:

$$q_{ВТРАТ}(t) = F\left(M\left(R_{j,t}\right)\right). \quad (5)$$

При умові, що $t_{ij} \leq h$ та з врахуванням (5) вираз (3) можна представити наступним чином:

$$k = \sum_{t=1}^{t_{ij}} v_t F\left(M\left(R_{i,t}\right)\right) + \sum_{t=t_{ij}+1}^{h} v_t F\left(M\left(R_{j,t}\right)\right) + \sum_{t=h+1}^{t_{спок}} v_t F\left(M\left(R_{j,t}\right)\right). \quad (6)$$

Враховуючи те, що прогноз кількість мережевих запитів робиться на h кроків вперед необхідно визначити значення v_t та R для проміжку часу $(h; t_{\text{СПОК}}]$. Використаємо прогноз на момент часу h для цього часового проміжку, тоді останній доданок в (6) може бути представлений добутком:

$$k = \sum_{t=1}^{t_{ij}} v_t F\left(M(R_{i,t})\right) + \sum_{t=t_{ij}+1}^{h} v_t F\left(M(R_{j,t})\right) + (t_{\text{СПОК}} - h) v_h F\left(M(R_{j,h})\right), \quad (7)$$

Таким чином, вираз (4) може бути деталізований за допомогою співвідношення (7):

$$G(e)\,|\,t_{ij} \le h = P_j t_{ij} + P_j\left(t_{\text{СПОК}} - t_{ij}\right) + $$

$$+P_{\text{ВТРАТ}}\left(\sum_{t=1}^{t_{ij}} v_t F\left(M(R_{i,t})\right) + \sum_{t=t_{ij}+1}^{h} v_t F\left(M(R_{j,t})\right) + (t_{\text{СПОК}} - h) v_h F\left(M(R_{j,h})\right)\right). \quad (8)$$

Для випадку коли $t_{ij} > h$ аналогічні перетворення приведуть до виразів (9) та (10):

$$k = \sum_{t=1}^{h} v_t F\left(M(R_{i,t})\right) + (t_{ij} - h) v_h F\left(M(R_{i,h})\right) + (t_{\text{СПОК}} - t_{ij}) v_h F\left(M(R_{j,h})\right) \quad (9)$$

$$G(e)\,|\,t_{ij} > h = P_j t_{ij} + P_j\left(t_{\text{СПОК}} - t_{ij}\right) + $$

$$+P_{\text{ВТРАТ}}\left(\sum_{t=1}^{h} v_t F\left(M(R_{i,t})\right) + (t_{ij} - h) v_h F\left(M(R_{i,h})\right) + (t_{\text{СПОК}} - t_{ij}) v_h F\left(M(R_{j,h})\right)\right). \quad (10)$$

Таким чином, на основі прогнозу кількості користувачів хмарного застосунку та середнього часу виконання мережевого запиту можна обчислити значення оціночної функції $G(e)$ для всіх можливих способів масштабування $e \in E_i$ хмарного застосунку, що має стан інфраструктури $\langle S_i, N_i, P_i \rangle$ в момент часу t. Спосіб масштабування з найменшим значенням оціночної функції використовується в якості обраного варіанта масштабування.

Формальний запис правила масштабування має наступний вигляд:

$$\text{якщо } t > t_{\text{гран}},$$
$$\text{то здійстини масштабування } e_{min} | G(e_{min}) = \min_{e} G(e)$$

$$\text{якщо } e_{min}.S' \neq S_i \text{ або } e_{min}.\Delta N \neq 0,$$
$$\text{то } t_{\text{гран}} \leftarrow t + \Delta t$$

де t – поточний момент часу, $t_{\text{гран}}$ – момент часу в який закінчується період стабілізації після минулої операції масштабування, Δt – проміжок часу спокою після масштабування впродовж якого нові операції масштабування не здійснюються.

Прогнозування часу виконання мережевого запиту

Для визначення значення оціночної функції ефективності масштабування хмарного застосунку (2.10) необхідно мати інформацію про кількість користувачів хмарного застосунку в хвилину v та про середній час виконання мережевого запиту $M(R)$. Побудуємо прогноз цих величин на основі прогнозу кількості мережевих запитів та статистичної інформації про функціонування хмарного застосунку.

Відношення кількості мережевих запитів до кількості унікальних користувачів при великій кількості користувачів є сталою величиною для конкретного хмарного застосунку. Тому прогноз кількості користувачів v_t може бути розрахована наступним чином:

$$v_t = x_t * c_{\text{корист}},$$

де t – проміжок часу для якого здійснюється прогнозування, x_t – прогноз кількості мережевих запитів за проміжок часу *t*, $c_{корист}$ – коефіцієнт, що показує частоту запитів від одного користувача за одиницю часу для конкретного хмарного застосунку. Значення коефіцієнта $c_{корист}$ визначається шляхом періодичного збору статистики про кількість унікальних користувачів та кількість мережевих запитів під часу роботи хмарного застосунку.

Для прогнозування часу виконання мережевого запиту на основі кількості мережевих запитів в одиницю часу необхідно мати інформацію про виділені обчислювальні потужності. Таку інформацію містить опис інфраструктури хмарного застосунку $\langle S, N, P \rangle$, а саме розмір віртуальних машин *S* та їх кількість *N*. Представимо можливі комбінації значень змінних x, S та *N* за допомогою гіперкуба станів хмарного застосунку[7, 8]. Кількість вимірів такого гіперкуба дорівнює трьом, в комірках зберігається масив, кожен елемент якого представляє собою час виконання окремого мережевого запису. В комірці також зберігається середнє значення елементів масиву, що дозволяє пришвидшити розрахунки. Для зменшення кількості комірок гіперкубу та полегшення прогнозування[9] необхідно провести дискретизацію змінної x, розбивши діапазон її значень на 3-10 відрізків.

Перевагою використання гіперкубу для зберігання даних про середній час мережевого запиту є можливість такого впорядкування даних, при якому виникає змога використання інформації з сусідніх комірок у випадку відсутності інформації в комірці, що описує стан хмарного застосунку. Графічне зображення гіперкубу станів хмарного застосунку наведене на рисунку 3, пунктирними стрілками показані напрями зростання середнього часу виконання мережевого запиту.

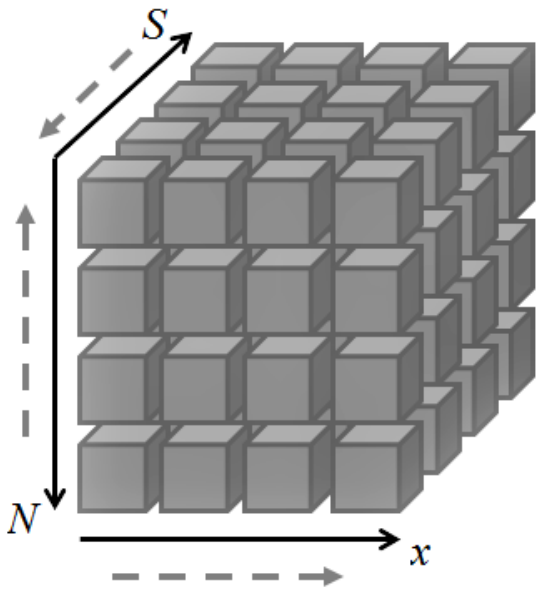


Рисунок 3 - Гіперкуб станів хмарноо застосунку

За умови, що зростання часу виконання мережевого запиту змінюється лінійно при пересуванні вздовж одного з вимірів гіперкубу, можна ввести алгоритм визначення часу виконання мережевого запиту, який дозволяє доповнити інформацію про пусті комірки гіперкуба за допомогою лінійної інтерполяції.

Крок 1. Початок роботи.

Крок 2. Введення інформації про розмірність гіперкубу станів хмарного застосунку k_S, k_N, k_x, множину значень його комірок C та цільовий стан iS, iN, ix, глибину пошуку *z*.

Крок 3. Якщо $C_{iS,iN,ix}$ має значення – виконати крок 4, інакше – крок 5.

Крок 4. Вивести $C_{iS,iN,ix}$, перейти до кроку 13.

Крок 5. Для *j* від 1 до z виконати кроки 6 – 11.

Крок 6. Якщо $C_{iS-j,iN,ix}$ існує та має значення і $C_{iS+j,iN,ix}$ існує та має значення – перейти до кроку 7, інакше – перейти до кроку 8.

Крок 7. Вивести $\left(C_{iS-j,iN,ix} + C_{iS+j,iN,ix}\right)/2$, перейти до кроку 13.

Крок 8. Якщо $C_{iS,iN-j,ix}$ існує та має значення і $C_{iS,iN+j,ix}$ існує та має значення – перейти до кроку 9, інакше – перейти до кроку 10.

Крок 9. Вивести $\left(C_{iS,iN-j,ix} + C_{iS,iN+j,ix}\right)/2$, перейти до кроку 13.

Крок 10. Якщо $C_{iS,iN,ix-j}$ існує та має значення і $C_{iS,iN,ix+j}$ існує та має значення – перейти до кроку 11, інакше – перейти до кроку 5.

Крок 11. Вивести $\left(C_{iS,iN,ix-j} + C_{iS,iN,ix+j}\right)/2$, перейти до кроку 13.

Крок 12. Вивести невизначений результат.

Крок 13. Кінець.

Схема алгоритму визначення часу виконання мережевого запиту зображена на рис. 4. За рахунок введення змінної *z* алгоритм дозволяє проводити прогнозування на основі комірок гіперкубу, значення яких є невизначеним.

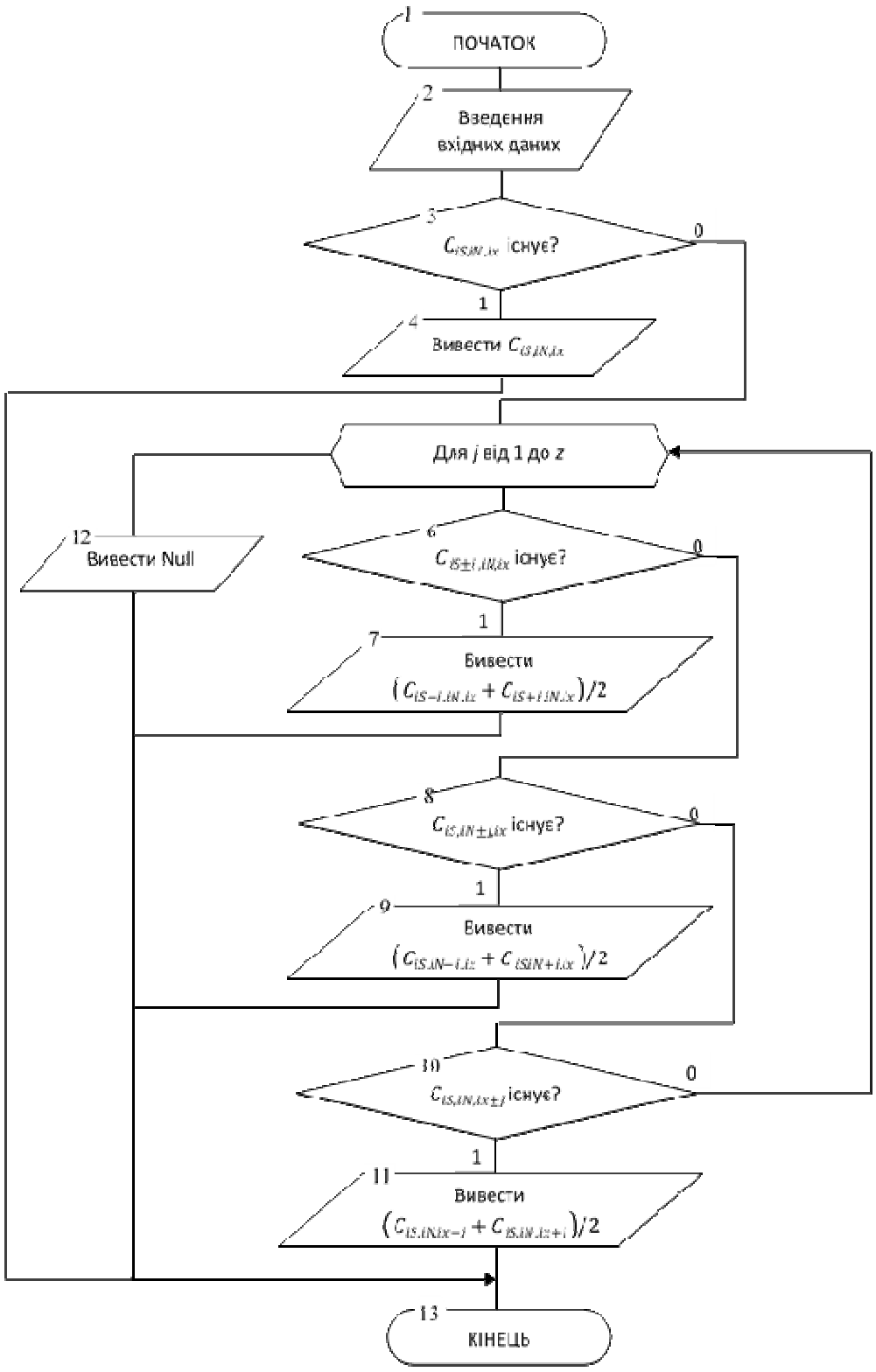


Рисунок 4 – Алгоритм визначення часу виконання мережевого запиту

Висновки

Таким чином, на основі інформації про стан хмарного застосунку було сформовано алгоритм, що на основі набору правил реактивного масштабування та оцінки ефективності здійснює прийняття рішення щодо проведення масштабування хмарного застосунку. Розроблена оціночна функція варіантів масштабування дозволяє отримати економічну оцінку ефективності, що базується на вартості утримання інфраструктури хмарного застосунку та на оцінці кількості користувачів, що перестануть користуватися хмар-

ним застосунком через його перевантаженість. Розроблений алгоритм може бути використаний в якості складової інформаційної технології масштабування хмарного застосунку.

Список літератури

1. Lorido-Botran T. Auto-scaling techniques for elastic applications in cloud environments/ T. Lorido-Botran, J. Miguel-Alonso, J. A. Lozano// Department of Computer Architecture and Technology, University of Basque Country, Tech. Rep. – EHU-KAT-IK-09, 2012 – № 12.
2. RightScale Cloud Management [Електронний ресурс] – режим доступа: http://www.rightscale.com/.
3. How a Slow Website Impacts Your Visitors and Sales [Електронний ресурс] – режим доступа: http://www.peer1.com/knowledgebase/how-slow-website-impacts-your-visitors-and-sales .
4. Nah F.F.H. A study on tolerable waiting time: how long are Web users willing to wait?// Behaviour & Information Technology, 2004 – №23(3) – С. 153-163.
5. How Loading Time Affects Your Bottom Line [Електронний ресурс] – режим доступа: https://blog.kissmetrics.com/loading-time/ .
6. Menasce D. Load testing of web sites.// Internet Computing, 2002 – № 6(4) – С. 70-74.
7. Савчук Т. О. Прогнозування стану техногенної ситуації за допомогою імітаційної моделі, що базується на багатовимірному рівнянні дифузії / Т. О. Савчук, А. В. Козачук// Вісник національного університету "Львівська політехніка". Інформаційні системи та мережі – 2014 – №783, – С. 473-478
8. Савчук Т. О. Визначення доцільності використання багатовимірного підходу до прогнозування стану техногенної ситуації/ Т. О. Савчук, А. В. Козачук// Вимірювальна та обчислювальна техніка в технологічних процесах – Хмельницький, 2014 - № 2(47) – С. 179-182
9. Data Mining with Microsoft SQL Server 2008. Jamie MacLennan, ZhaoHui Tang, Bogdan Crivat. 2008. – 672 pages. – ISBN: 978-0-470-27774-4.

Відомості про авторів

Савчук Тамара Олександрівна – к.т.н. проф. кафедри Комп'ютерних наук ВНТУ. Місто Вінниця, Хмельницьке шосе 95.

Козачук Андрій Валерійович – аспірант кафедри Комп'ютерних наук ВНТУ. Місто Вінниця, Хмельницьке шосе 95.

КОМП`ЮТЕРНІ СИСТЕМИ ТА КОМПОНЕНТИ

УДК 621.375.024

О. Д. АЗАРОВ, Л. В. КРУПЕЛЬНИЦЬКИЙ, В. А. ГАРНАГА, Д. Ю. ПОЗНЯК

Вінницький національний технічний університет, м. Вінниця

ПОЛІГАРМОНІЙНІ МЕТОДИ ВИМІРЮВАННЯ ЧАСТОТНИХ ХАРАКТЕРИСТИК ЗВУКОВИХ КАНАЛІВ І ТРАКТІВ

Анотація. У статті здійснено аналіз методів вимірювання амплітудно-частотної характеристики каналів передавання звукової інформації, визначено переваги і недоліки відомих методів вимірювання. Запропонований новий метод вимірювань, який дозволяє швидко отримати крім АЧХ каналу передавання інформації й інші важливі параметри, зокрема, груповий час затримки, коефіцієнт нелінійних спотворень та інші. Показано прилад АПЗТ, побудований на запропонованому методі вимірювань параметрів звукових трактів, а також наведено його параметри. Представлено графіки АЧХ, які отримано із використаня запропонованого методу.

Ключові слова: АЧХ, полігармонійні сигнали, звукові тракти, вимірювання.

Аннотация. В статье проведен анализ методов измерения амплитудно-частотной характеристики каналов передачи звуковой информации, определены преимущества и недостатки известных методов измерения. Предложен новый метод измерений, который позволяет быстро получить кроме АЧХ канала передачи информации и другие важные параметры, в частности, групповое время задержки, коэффициент нелинейных искажений и другие. Показан прибор АПЗТ, построенный на предложенном методе измерений параметров звуковых трактов, а также приведены его параметры. Представлены графики АЧХ, полученные с использованием предложенного метода.

Ключевые слова: АЧХ, полигармонические сигналы, звуковые тракты, измерения.

Abstract. The article analyzes the methods of measuring the amplitude-frequency responses of the transmit audio information identified advantages and disadvantages of known methods of measurement. A new method of measurement, which allows a quick response in addition to channel information and other important parameters, such as group delay, distortion factor and others. Showed APZT device, built on the proposed method of measuring parameters of audio paths, and given the options. AFC is a graph obtained using the proposed method.

Key worlds: AFC, polyharmonic signals, sound path, measurement.

Вступ

Сприйняття людиною інформації в значній мірі відбувається на слух. Для успішного засвоєння цієї інформації потрібно забезпечити належну якість звукового каналу передавання інформації, оскільки, при недостатній якостіінформація маже бути сприйнята зі спотвореннями і неправильно інтерпретована. Крім того, довгострокове прослуховування неякісного звукового сигналу впливає на самопочуття та здоров’я людини. Саме тому є важливою задача контролю якості каналів передавання звукової інформації. Водночас, варто відзначити, що цією проблемою вже досить давно займаються науковці [1–4] та інженери і розроблено методики оцінювання, а також підготовлено відповідні державні стандарти [5] до визначення якості каналів передавання інформації.

Амплітудно-частотна (АЧХ)є однією з найважливіших характеристик якості каналів і трактів радіомовлення та звукового супроводу телебачення. При вимірюванні АЧХ виникає також задача одночасного оцінювання фазочастотної характеристики (ФЧХ) та характеристики групового часу затримки (ГЧЗ). Коректність виміру частотних характеристик реальних каналів і трактів залежить від рівня вимірювального сигналу, рівня шумів, нелінійних спотворень та наявності специфічних для цифрових систем властивостей, пов’язаних з ущільненням звукового сигналу. Так звану малосигнальну АЧХ, що дозволяє прогнозувати потенційно досяжні характеристики тих чи іншихелектронних схем, можна отримати лише шляхом комп’ютерного моделювання і аналізу схем. У практиці вимірювань використовується інший підхід; в якому АЧХ і ФЧХ каналу або тракту вимірюються в рамках діапазону сигналу на певному рівні, наприклад – 21 дБ [5].

Аналізуючи підходи і методи вимірювання до вказаної задачі можна прийти до висновку, що існуючі підходи є достатньо громіздкими, і те яким чином та з якими засобами вони виконувалися є застарілими. Розвиток технічних засобів дозволяє запропонувати нові підходи, щодо реалізації таких вимірювань.

Актуальність

Відомі декілька методів вимірювання та оперативного контролю АЧХ та її окремих параметрів – з використанням послідовного набору сигналів з фіксованими частотами, з використанням сигналу, частота якого змінюється в часі та з використанням шумоподібного сигналу з рівномірним спектром частот. Вказані методи мають свої недоліки, пов’язані як зі складністю процедури вимірювань, так і з неможливістю адекватного оцінювання параметрів цифрових трактів з ущільненням звукових сигналів. Разом з тим, нові можливості відкривають комп’ютерні методи вимірювань з використанням високоякісних ЦА- і АЦ-перетворювачів та методів спектральної обробки сигналів.

Таким чином, **метою** статті є розгляд відомих методів вимірювання АЧХ звукових трактів та аналіз можливостей пропонованогополігармонійногометоду, який дозволяєз використанням одного вимірювального сигналу отримати усі необхідні частотні характеристики (АЧХ, ФЧХ та ГЧЗ) звуковому тракту та оцінити якість цифрового ущільнення звуку.

Задачідосліджень:

1. Запропонувати полігармонійні методи для визначення АЧХ, ФЧХ та ГЧЗ звукових трактів.
2. Провести порівняльний аналіз запропонованих і відомих методів визначення частотних характеристик звукових трактів.

Розв’язання задач

Проаналізуємо відомі методи визначення АЧХ каналів їх переваги і недоліки.

Найпростіший із методів пов’язаний з використанням генератора низьких частот та вольтметра або вимірювача рівня. Вимірювання виконують почергово для кожної точки із фіксованого ряду частот, подаючи на вхід звукового тракту одно тональний (моногармонічний) сигнал з генератора. Вимірювачем оцінюють середньоквадратичне значення вихідного сигналу. Недоліком такого моногармонічногометоду є досить великі затрати часу на вимірювання та обробку результатів. Як правило, передбачається вимірювання АЧХ при рівні вхідного сигналу мінус 21 дБ відносно номінального, а рівні вихідних сигналів необхідно нормувати відносно рівня на опорній частоті 1 кГц. Результати, звичайно, необхідно представляти у децибелах, у табличній та графічній формах. При вимірюванні АЧХ міжміських каналів передачі звукового мовлення затрати часу зростають також у зв'язку з необхідністю узгоджувати процес вимірювання між двома операторами , а в результати вимірювань вноситься додаткова випадкова похибка через флуктуацію в часі коефіцієнту передачі каналу. Моногармонійний метод погано пристосований до контролю форми АЧХ при оперативних регулюваннях.

У моногармонійному методі додатково, при наявності взірцевого опорного каналу, за допомогою фазометра, можна виміряти також і ФЧХ. Така можливість є тільки в студійних звукових трактах без пристроїв запису-відтворення звуку. Характеристику групового часу затримки, яка математично є похідною від ФЧХ, можна оцінити виконуючи вимірювання різниці фаз між близькими частотами і враховуючи можливі переходи фази через період. Але, через складність, такий метод вимірювання ГЧЗ практично не використовують.

У аналогових спектроаналізаторах для оцінювання форми АЧХ використовується вимірювальний сигнал з частотною модуляцією – змінюється миттєва частота гармонічного сигналу при незмінній амплітуді. Такий метод найбільш придатний для швидкого графічного відтворення форми АЧХ на екранах з запам’ятовуванням. Через плавну зміну частоти сигналу можливе вимірювання ГЧЗ, але у відомих спектроаналізаторах для звукового діапазону частот ця можливість не використана. При вимірюванні АЧХ звукових трактів метод має недоліки , пов'язані з невисокою точністю вимірювань, складністю отримання відліків на фіксованих частотах та неможливістю синхронізувати вимірювання в міжміських каналах звукового мовлення та з'єднувальних лініях.

У цифрових приладах для вимірювання АЧХ отримав поширення метод, що грунтується на шумоподібному вимірювальному сигналі. Формування шумоподібного сигналу з рівномірно розподіленим спектром відбувається в цифровому вигляді за рахунок генератора псевдовипадкової послідовності. За результатами вимірювань оцінюють спектр вихідного сигналу, форма обвідної якого відтворює АЧХ тракту. Такий метод більше пристосований до специфіки вимірювань в звукових каналах і трактах, не вимагає попередньої синхронізації, дозволяє за рахунок цифрового спектрального аналізу або фільтрації отримати відліки АЧХ на окремих частотах, може використовуватись при оперативних регулюваннях. Недоліком шумоподібного сигналу є принципова неможливість вимірювання ФЧХ та ГЧЗ.

Імпульсно-подібний сигнал, наближений до дельта-функції також має рівномірний спектр і використовується, наприклад, для виміру часу акустичної реверберації приміщень. Він може також використовуватись при вимірюванні АЧХ, ФЧХ та ГЧЗ електричних звукових трактів, оскільки окремі спектральні складові дельта-функції мають однакову амплітуду і початкову фазу. Однак, порівняно з шумоподібним сигналом, амплітуда кожної окремої спектральної складової імпульсного сигналу суттєво менша амплітуди сумарного імпульсу. А це призводить до підвищеного впливу шумів і завад тракту на результати розрахунку АЧХ, ФЧХ і ГВЗ.

Для реалізації полігармонійного методу вимірювання використовується програмно-апаратний комплекс структуру якого наведено на рис. 1. Він складається з програмного забезпечення, яке дозволяє згенерувати тестовий сигнал та самого комплексу, що реалізує синтез в аналоговому вигляді тестовий сигнал і потім передає його через канал зв’язку, якість якого потрібно визначити.

Для вимірювання АЧХ і ФЧХ звукових каналів і трактів пропонуєтьсявикористовуватиполігармонійний сигнал (ПГС), який сформовано з суми 10÷20 гармонічних складових, значення частот яких можуть задаватись користувачем із нормованого ряду або довільно.

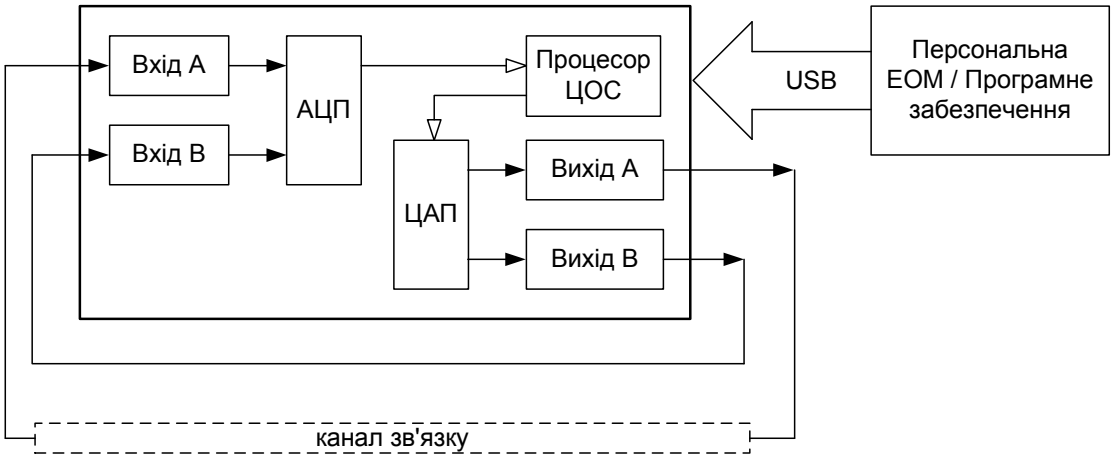


Рисунок 1 – Узагальнена структура пристрою, що реалізує полігармональний метод вимірювання АЧХ

Всі амплітуди і початкові фази гармонічних складових однакові:

$$S(t) = \sum_{i=1}^{N} A_i \sin(2\pi f_i t)$$

де A_i – амплітуда кожної з N гармонійних складових.

Приклад осцилограми такого полігармонічного сигналу наведено на рис. 1, а його спектр – на рис. 2 (задано частоти 32.5, 63, 125, 250, 500, 1000, 2000, 4000, 10000, 16000 Гц). Для отримання АЧХ і ФЧХ необхідно виконати аналіз спектру вихідного сигналу звукового тракту, вибрати опорну частоту (наприклад, 1000 Гц) та розрахувати відхилення рівня сигналу на інших частотах, що дає нерівномірність АЧХ. Для отримання ФЧХ розраховується різниця фаз між кожною спектральною складовою та опорною. Оскільки вимірювання АЧХ і ФЧХ відносні, по одній виборці сигналу, то мінімізуються похибки через флуктації коефіцієнту передачі звукового каналу, а також немає потреби в опорному сигналі для виміру ФЧХ. Останнє суттєво, наприклад, для вимірювання характеристик міжміських каналів звукового мовлення або трактів запису-відтворення магнітофонів.

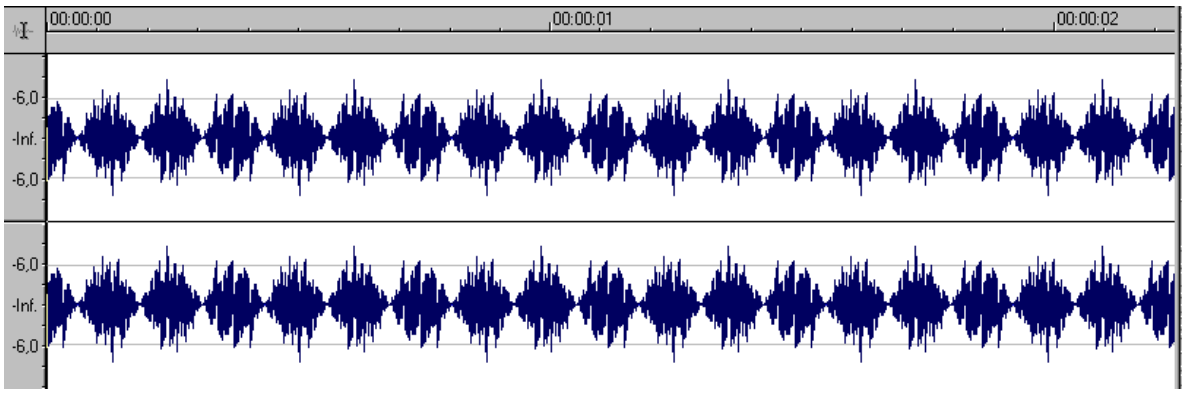


Рисунок 2 – Форма тестового сигналу

Для детальнішого аналізу форми АЧХ та виміру характеристики ГЧЗ пропонується так званий імпульсоподібний сигнал (ІПС), в якому число гармонічних складових порівняно з ПГС збільшено до 50÷150. Назва ІПС пояснюється формою осцилограми сигналу, наведеної на рис. 4. Як видно із спектру ІПС на рис. 5, спектральні складові заповнюють весь діапазон смуги частот 20 Гц – 20 кГц. При цьому АЧХ каналу передавання можна визначити у вигляді відношення отриманого сигналу $S'(t)$ до згенерованого $S(t)$:

$$\frac{S'(t)}{S(t)} = \frac{\sum_{i=1}^{N} A_i' \sin(2\pi f_i t)}{\sum_{i=1}^{N} A_i \sin(2\pi f_i t)},$$

при цьому відношення A_i' / A_i дозволить оцінити коефіцієнт передачі звукового тракту на відповідній частоті.

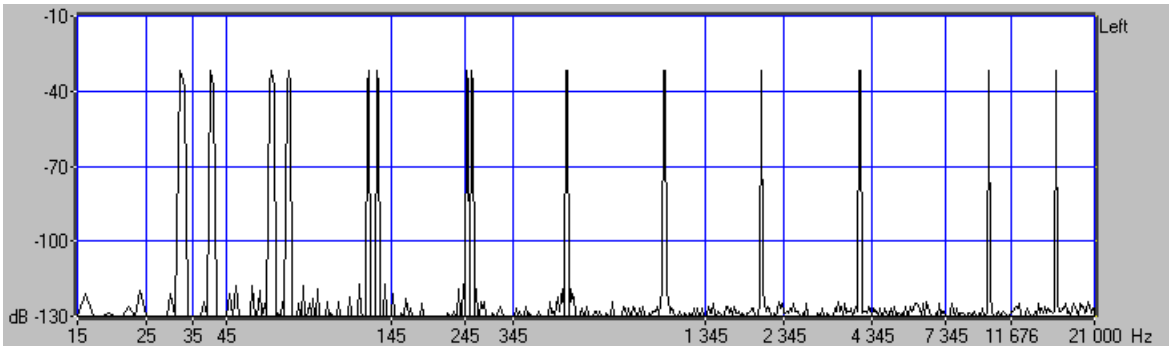


Рисунок 3 – Спектр тестового сигналу

Розрахунок ГЧЗ можливий за рахунок малого кроку по частоті між спектральними складовими.Характеристика групового часу затримки (ГЧЗ) показує різницю в часовій затримці сигналів різних частот, що проходять через звуковий тракт. ГЧЗ визначають як похідну по частоті від фазочастотної характеристики (ФЧХ):

$$G(f) = \frac{d\Psi(f)}{df},$$

де обчислення похідної замінюється обчисленням відношення приросту ФЧХ до приросту частоти на імпульсно-подібному сигналі ІПС.

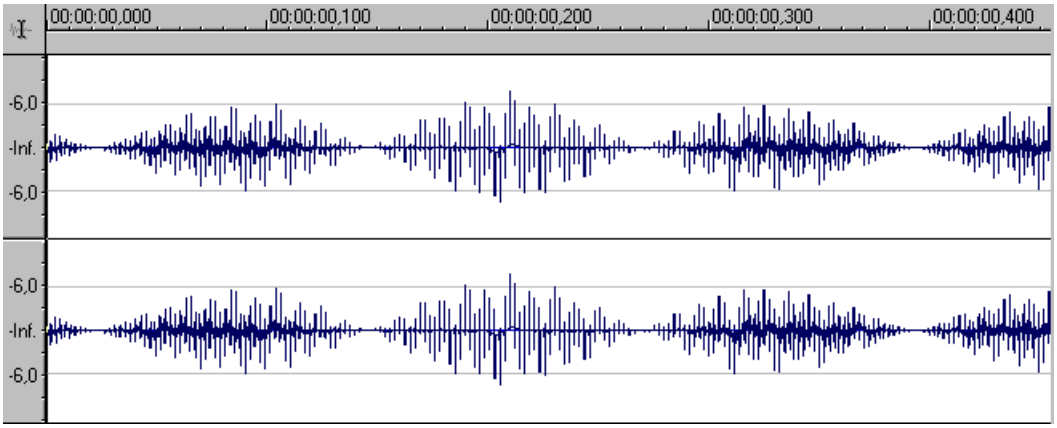


Рисунок 4 – Форма імпульсо-подібного сигналу

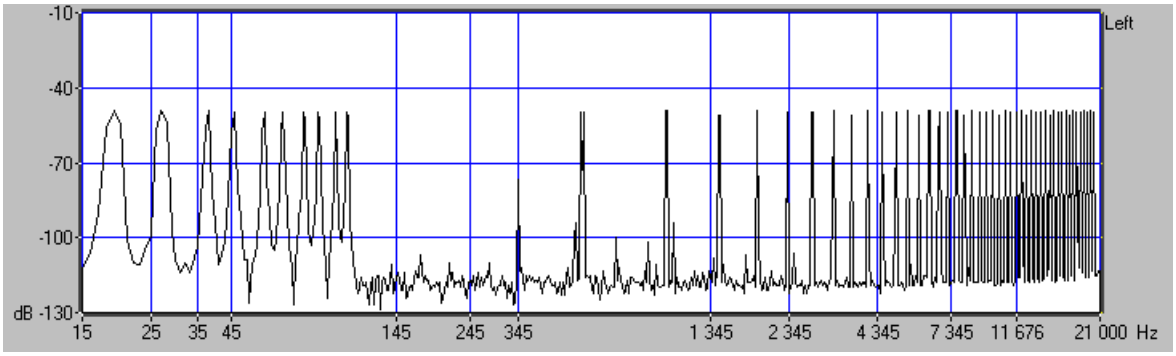


Рисунок 5 – Спектр імпульсо-подібного сигналу

Метод вимірювання АЧХ за допомогою ІПС наочно демонструє свої перваги при контролі параметрів трактів, що мають в своєму складі апаратуру компресії-декомпресії за методами

MPEG.Застосуваннякомпресіїдозволяєсуттєвозменшити полосу пропускання каналу, але зрозуміло, щоцепризводить до зниженняякості сигналу. Вимірювання АЧХ такихканалів методами моно гармонічного абочастотномодульованого тестового сигналу не виявляєвідхиленьвіднорми. В той же час, в деякихвипадкахсуб'єктивна оцінка на реальному звуковому сигналі не дозволяє оцінити АЧХ як таку, що задовольняє відповідним нормам.

Тестовий сигнал, що складається з багатьох спектральних складових є найбільш наближеним до реального звукового сигналу. Саме тому метод вимірювання за допомогою ПГС та ІПС дозволяє найбільш об'єктивно оцінити АЧХ тракту під час передачі звукового сигналу. На рис 5 наведено спектр ІПС після проходження через компресор-декомпресор MPEG-2 з швидкістюпропускання каналу 384 кБіт/с.

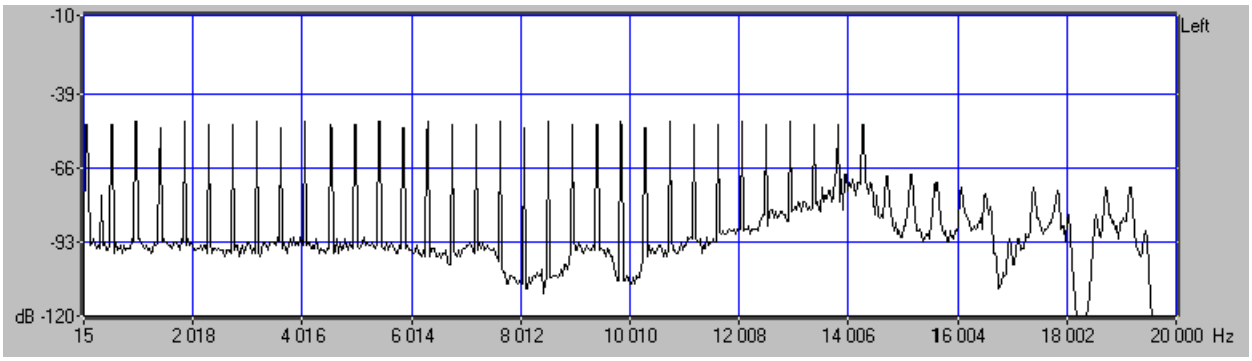


Рисунок 6 – Спектр імпульсо-подібного сигналу після проходження каналу зі стисненням

Авторами було проведено тестування різних методів компресії з різноманітними полосами пропускання, всі результати яких не можуть бути включені до цієї статті. Відзначимо лише, що найкращі результати показав метод компресії MPEGLayer 3. Спектр ІПС після проходження через компресор-декомпресор MPEGLayer 3 з швидкістю пропускання каналу 128 кБіт/с наведено на рис 7.

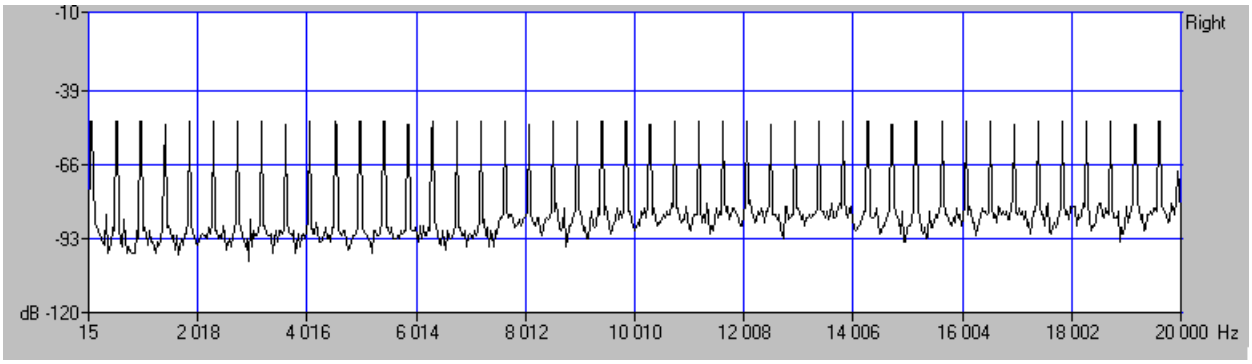


Рисунок 7– Спектр імпульсо-подібного сигналу після проходження каналу зі компресор-декомпресоромMPEGLayer 3

Технічна реалізація методів вимірювань з використанням полігармонійних сигналів можлива в системах з цифроаналоговим і аналого-цифровим перетворювачами (ЦАП і АЦП) та процесором цифрової обробки сигналів. Прикладом такої вимірювально-інформаційної системи є, розроблений за участю авторів, комп’ютерний аналізатор параметрів звукових трактів АПЗТ-02, зовнішній вигляд якого наведено на рис.8.

АПЗТ-02 призначений для автоматизованого виміру електричних характеристик трактів формування та передачі звукових програм радіомовлення та звукового супроводу телебачення, низькочастотних трактів радіоелектронної апаратури, магнітофонів, вимірювачів рівня, тощо. При вимірюванні параметрів звукових трактів АПЗТ-02 замінює собою до 14 окремих приладів: середньоквадратичний вольтметр, селективний вольтметр, вимірювач АЧХ, ГЧЗ, фазометр, генератор низьких частот, вимірювач нелінійних спотворень, спектроаналізатор, частотомір, вимірювач інтегральних та псофометрічних шумів, детонометр, генератор для вимірювання параметрів квазіпікових індикаторів, вимірювач опору навантаження. АПЗТ-02 пройшов державну метрологічну атестацію і забезпечує всі вимірювання, що передбачені ГОСТ 11515 "Канали і тракти звукового мовлення. Основні параметри якості. Методи вимірювань".

Рисунок 8 – Зовнішній вигляд АПЗТ-02

При вимірюванні АЧХ з застосуванням АПЗТ-02, полігармонічний сигнал розраховується як сума гармонічних складових, відтворюється за допомогою ЦАП та подається на вхід звукового тракту, щовимірюється. Вихідний сигнал тракту в АЦП аналізатораперетворюється в цифрову форму, процесорвиконуєшвидкеперетворення Фур'є і розраховує спектр сигналу. Визначення параметрів АЧХ, ФЧХ та ГЧЗ здійснюється шляхом аналізу спектру. Для зменшення систематичних похибок в АПЗТ-02 здійснюється автокалібрування, під час якого визначається, а в подальших розрахунках враховується АЧХ і ФЧХ вимірювальних каналів.

Характеристики входів-виходів АПЗТ-02:

- вхід і вихід - симетричні, стереофонічні;
- діапазон робочих частот - 20 Гц - 20 кГц;
- вхідні діапазони вимірювань - від +18 дБ до мінус 30 дБ з кроком 6 дБ (0 дБ = 0.775 В еф.);
- вихідний рівень сигналу генератора - від +12 дБ до мінус 76 дБ з кроком 0.1 дБ;
- вхідний опір - 150 Ом, 200 Ом, 600 Ом, 100 кОм;
- вихідний опір - 0 Ом, 600 Ом.

Таблиця 1 – Основніхарактеристики, щоможуть бути визначені за допомогою АПЗТ-02 та їх діапазони

Характеристики	**Діапазони значень**
середньоквадратична напруга	від +20 до мінус 90 дБ
селективна напруга	від +20 до мінус 90 дБ
шуми інтегральні і псофометричні	від мінус 20 до мінус 90 дБ
рівень селективної перешкоди	від мінус 20 до мінус 90 дБ
частота вхідного сигналу	від 20 до 20000 Гц
амплітудно-частотна характеристика	від 20 до 20000 Гц
фазочастотна характеристика і різниця фаз	від мінус 90 до +90 градусів
груповий час затримки	від 0.5 до 50 мс
коефіцієнт гармонік	від 0.05 до 5%
коефіцієнт різницевого тону	від 0.05 до 5%
коефіцієнт інтермодуляционнихспотворень	від 0.05 до 5%
рівень селективної перешкоди	від мінус 20 до мінус 90 дБ
захищеність від перехідних перешкод	від 30 до 80 дБ
захищеність від паразитної амплітудної модуляції	від 30 до 80 дБ
коефіцієнт детонації магнітофона	від 0.01 до 5%;
тривалість імпульсного сигналу для випробувань індикаторів	від 1 до 220 мс

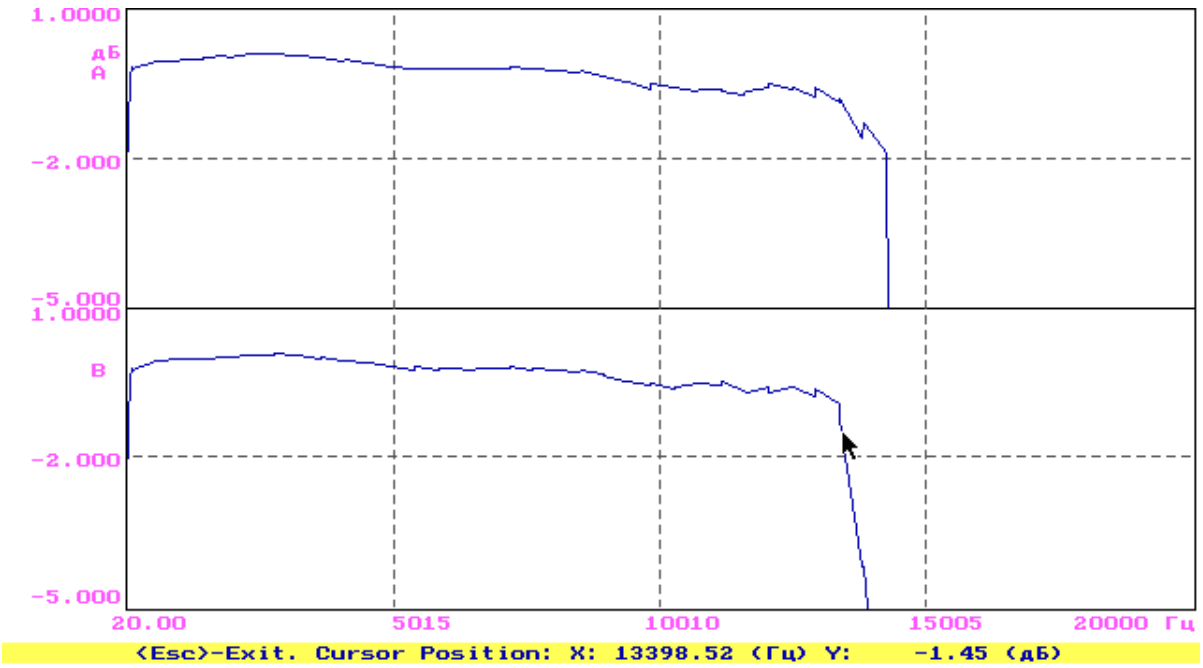


Рисунок 9 – АЧХ стереоканалів міжміського тракту звукового мовлення Київ–Одеса

Висновки:

1. Запропоновано методи полігармонійного вимірювання параметрів каналів передавання. Вони за рахунок наявності у складі тестового сигналу великої кількості гармонійних складових (20÷150) у одному сигналі складної форми,дозволяють швидко оцінити АЧХ, ФЧХ і ГЧЗ звукового тракту у всьому діапазоні робочих частот, а також отримати і інші характеристики, які наведено у табл. за один прохід пересилання тестового сигналу.

2. Запропонований пристрій має два канали, які дозволяють використовувати процедури самокоригування і за рахунок цього забезпечити високі метрологічні характеристики пристрою.

3. Запропонований полігармонійний метод, дозволяє адекватно оцінити реальну смугу пропускання каналів і трактів з цифровим ущільненням сигналів, оскільки на такому сигналі відслідковується обмеження психоакустичних методів ущільнення звуку.

Література

1. Азаров О. Д. Основитеорії аналого-цифрового перетворення на основінадлишковихпозиційних систем числення :монографія. – УНІВЕРСУМ-Вінниця, 2010. – 232 с.

2. Азаров О. Д. Двотактні підсилювачі постійного струму для багато розрядних перетворювачів форми інформації, щосамокалібруються : монографія / О. Д. Азаров, В. А. Гарнага. – УНІВЕРСУМ-Вінниця, 2011. – 156 с.

3. Азаров О. Д. Методи та засобивисокоточногослідкувального аналого-цифрового перетворення з ваговоюнадлишковістю. :монографія / О.Д. Азаров, О.В. Дудник. - УНІВЕРСУМ-Вінниця, 2014.- 120 с.

4. Крупельницький Л.В., Азаров О.Д. Аналогові пристрої самокоригуючих АЦП для систем обробляння низькочастотних сигналів / Під редакцією Азарова О.Д. – Монографія – 2005.

5. ГОСТ 11515-91 Каналы и тракты звукового вещания. Основныепараметры качества. Методы измерений.

Інформація про авторів

Азаров Олексій Дмитрович – д. т. н., професор, декан факультету інформаційних технологій та комп’ютерної інженерії Вінницького національного технічного університету.

Гарнага Володимир Анатолійович – к. т. н., доцент, кафдра обчилювальної техніки Вінницького національного технічного університету.

Крупельницький Леонід Віталійович – к. т. н., доцент, кафдра обчилювальної техніки Вінницького національного технічного університету.

Позняк Дмитро Юрійович – провідний програміст НТЦ "Аналого-цифрові системи" Вінницького національного технічного університету.

УДК 621.396

Г. Г. БОРТНИК, В. В. КИЧАК, Н. О. ПУНЧЕНКО

Вінницький національний технічний університет. Вінниця

ОЦІНЮВАННЯ ЗАВАДОСТІЙКОСТІ ЧАСТОТНО-ІМПУЛЬСНОГО МЕТОДУ КОДУВАННЯ ІНФОРМАЦІЇ ПРИ ПОБУДОВІ НАДВИСОКОЧАСТОТНИХ ЦИФРОВИХ ПРИСТРОЇВ

Анотація. Розроблено метод оцінювання виграшу по завадостійкості при застосуванні частотно-імпульсного методу кодування інформації.Показано, що застосування частотно-імпульсного методу кодування інформації дає можливість суттєво підвищити завадостійкість пристроїв передавання та оброблення інформації, а виграш по завадостійкості може перевищувати декілька одиниць і залежить від індексу модуляції.

Ключові слова:частотно-імпульсний метод кодування інформації, частотна модуляція, завадостійкість.

Аннотация.Разработан метод оценки выигрышапо помехоустойчивости при использовании частотно-импульсного метода кодирования информации. Показано, что применение частотно-импульсного метода кодирования информации дает возможность значительно повысить помехоустойчивость устройств передачи и обработки информации, а выигрыш по помехоустойчивости может превышать несколько единиц и зависит от индекса модуляции.

Ключевые слова:частотно-импульсний метод кодированнияинформации, частотнаямодуляция, помехоустойчивость.

Abstract. The noise-stability gain estimation method at information pulse-frequency coding application is developed.The information pulse-frequency coding application is shown to allow information transmission and processing devices' noise-stability significant increasing, also the noise-stability gain is shown to exceed several unities and to depend on a modulation index.

Keywords:nformation pulse-frequency coding application, pulsesmodulation,significant.

Вступ

При розробці цифрових пристроїв і систем передавання та оброблення інформації важливим є їх здатність протистояти дії шкідливих завад, які приводять до зниження достовірності переданого повідомлення. Одним із важливих показників якості та достовірності оброблення та передавання інформації є завадостійкість. Оскільки в діапазоні надвисоких частот доцільно використовувати частотно-імпульсний метод кодування інформації [1], тому доцільно оцінити його завадостійкість у порівнянні з традиційним імпульсно-потенціальним методом кодування інформації.

Постановка задачі

Завданням цієї праці є розробка методу оцінювання завадостійкості при кодуванні інформації частотно-імпульсними сигналами та дослідження її залежності від параметрів модульованих сигналів і частотної модуляції.

Розв'язання задачі

Частинним критерієм якості передавання сигналів при використанні модуляції є відношення сигнал/шум на виході демодулятора (приймача)

$$\left(S/N\right)_{\text{вих}} = \frac{P_{C\text{вих}}}{P_{Ш\text{вих}}}.$$

Якщо $K = \frac{U_m}{U_g}$ – пік-фактор сигналу і $U_m = 1$, то $U_g = \frac{1}{K}$ і потужність сигналу визначається за виразом

$$P_{C\text{вих}} = \frac{1}{K^2}.$$

Тоді відношення сигнал/шум на виході демодулятора (приймача) визначається за співвідношенням

$$\left(S/N\right)_{\text{вих}} = \frac{1}{K^2 \int_0^{F_C} N_m(t)dt}, \quad (1)$$

де $N_m(t)$ - спектральна густина потужності шуму.

Другою важливою характеристикою радіотехнічних пристроїв приймання та оброблення сигналів є відношення сигнал/шум на вході демодулятора

$$\left(\frac{S}{N}\right)_{\text{вх}} = \frac{P_{C\text{вх}}}{P_{Ш\text{вх}}} - \frac{P_{C\text{вх}}}{N_0 F}, \tag{2}$$

де N_0 - спектральна густина потужності білого шуму.

Узагальнений критерій якості (завадостійкості) пристрою модуляції сигналів, або виграш при застосуванні відповідного методу модуляції можна розрахувати за виразом [2]

$$Q = \frac{\left(S/N\right)_{\text{вих}}}{\left(S/N\right)_{\text{вх}}}. \tag{3}$$

Цей критерій слугує мірою достовірності приймання сигналів для заданого виду модуляції.

При $Q > 1$ - відношення сигнал/шум у процесі модуляції зростає, при $Q < 1$ - відношення сигнал/шум у процесі модуляції зменшується, і замість виграшу система модуляції дає програш, тобто вплив шуму на виході демодулятора (приймача) підвищується.

Проведемо аналіз завадостійкості при застосуванні частотної модуляції.

Частотна модуляція відноситься до інтегральних видів модуляції, тобто модульований сигнал входить у вираз для ЧМ-сигналу під знаком інтегралу. Окрім того, ЧМ – нелінійна модуляція, тому що ЧМ-сигнал можна отримати з модульованого сигналу S(t) шляхом виконання нелінійних операцій.

$$U_{\text{чм}}(S,t) = U_0 cos\left[\omega_0 t + \Delta\omega\psi(t)\right], \tag{4}$$

де $\Delta\omega$ - девіація частоти; $\psi(t) = \int_0^T S(\tau)d\tau$ - повна фаза сигналу.

Потенційна завадостійкість ЧМ визначається на базі виразу для вихідного шуму приймача [3]

$$G_{Ш}(f) = \frac{\left(2\pi f\right)^2 N_0}{\overline{\left(\frac{dU}{d\psi}\right)^2}}, \tag{5}$$

$\sim\sim$ - усереднення у часі;

——— - усереднення за реалізацією (ансамблем).

Тобто, енергетичний спектр завади на виході ЧМ-приймача є параболічним (пропорційним квадрату частоти) або «рожевий шум».

Для ЧМ-сигналу (4) знайдемо частинну похідну

$$\frac{\partial U}{\partial\psi} = -\Delta\omega U_0 sin\left[\omega_0 t + \Delta\omega\psi(t)\right]. \tag{6}$$

Виконаємо усереднення квадрату частинної похідної ЧМ-сигналу у часі за реалізацією.

$$\begin{aligned}\overline{\left(\frac{\partial U}{\partial\psi}\right)^2}^2 &= \frac{1}{T}\int_0^T dt = \frac{\Delta\omega^2 U_0^2}{T}\int_0^T sin^2\left[\omega_0 t + \Delta\omega\psi(t)\right]dt = \\ &= \frac{\Delta\omega^2 U_0^2}{2T}\int_0^T\left\{1 + cos\left[2\omega_0 t + \Delta\omega\psi(t)\right]\right\}dt = \frac{1}{2}\Delta\omega^2 U_0^2\end{aligned}. \tag{7}$$

Оскільки $\omega_0 T >> 1$, тому значення $\int_0^T cos 2\omega_0 t \cdot dt = \frac{sin 2\omega_0 T}{2\omega_0} < \frac{1}{2\omega_0}$ логічно не враховувати.Підставивши (7) в (5) отримаємо:

$$G_{Ш}(t) = 2\left(\frac{2\pi f}{\Delta\omega}\right)^2 \frac{N_0}{U_0^2} \tag{8}$$

Модульований ЧМ-сигнал – нестаціонарний процес, тому для знаходження його потужності необхідно виконати усереднення квадрату сигналу у часі та за реалізацією

$$P_{Свих} = \overline{U^2_{\text{ЧМ}}(S,t)} = \frac{1}{T}\int_0^T U^2_{\text{ЧМ}}(S,t)\,dt = \frac{1}{2T}U_0^2\left\{1+cos\left[2\omega_0 t + \Delta\omega\overline{\psi(t)}\right]\right\}dt = \frac{1}{2}U_0^2 \quad (9)$$

Модульований сигнал $S(t)$ - центрований сигнал, тому $\overline{\psi(t)} = 0$.

Знаходимо виграш ЧМ-системи модуляції, підставивши у вираз (3) вирази (1), (2) і (8), (9)

$$Q_{\text{чм}} = \frac{N_0\Delta F_{\text{сфм}}}{K^2\frac{1}{2}U_0^2\int_0^{F_C} 2\left(\frac{2\pi f}{2\pi\Delta f}\right)^2\frac{N_0}{U_0^2}df} = \frac{3\Delta f^2\Delta F_{\text{сфм}}}{k^2F_{\text{С}}^3} = \frac{3m_{\text{чм}}^2\Delta F_{\text{сфм}}}{k^2F_{\text{С}}}, \quad (10)$$

де $\Delta F_{\text{сфм}}$ - смуга частот ЧМ – сигналу; $F_{\text{С}}$ - максимальна частота модульованого сигналу; $m_{\text{чм}}$ - індекс ЧМ; $m_{\text{чм}} = \frac{\Delta f}{F_{\text{С}}}$.

Введемо частинний критерій частотної надлишковості $\beta_f = \frac{\Delta f_{\text{см}}}{F_C}$.

Тоді вираз (10) буде виглядати

$$Q = \frac{3m^2\beta_{f\text{чм}}}{k^2}. \quad (11)$$

При ЧМ виграш може бути значно вищим 1, тобто $Q_{\text{чм}} >> 1$. Це $Q_{\text{чм}}$ досягається за рахунок розширення смуги частот ЧМ сигналу при збільшенні m. Таким чином, ЧМ – завадостійкість може бути підвищеною не тільки за рахунок збільшення потужності сигналу $P_{\text{С}}$, але і за рахунок розширення смуги робочих частот

$\Delta f_{\text{чм}} = 2f_B\left(1 + m_{\text{чм}} + \sqrt{m_{\text{чм}}}\right)$ – формула Манаєва для m=0,1,2,…,20,

де $f_{\text{В}}$ - верхня частота модульованого сигналу.

Розрахунки проводились для синусоїдального моделювального сигналу – $k = \sqrt{2}$, телевізійного сигналу – $k = \sqrt{3}$ та багатотонального сигналу– $k = 4$.

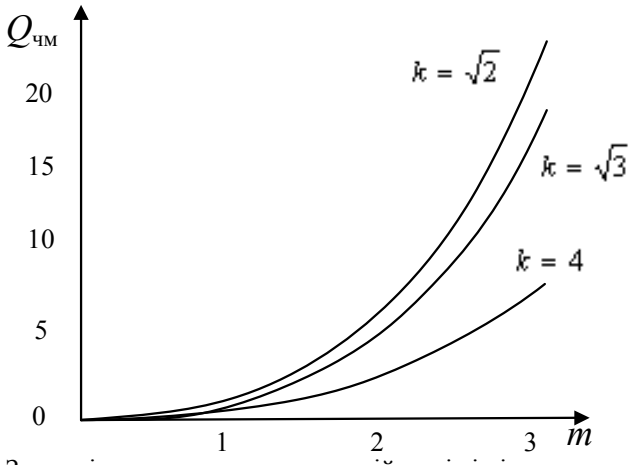


Рисунок 1 – Залежність виграшу по завадостійкості від індексу модуляції для частотно-модульованих сигналів

Як видно із рис. 3.1, при індексі модуляції більше1, виграш по завадостійкості перевищує 1 і досягає більше 8 при m більше 3.

Висновки

Розроблено метод оцінювання завадостійкості при застосуванні частотно-модульованих сигналів для кодування інформації. Показано, що виграш по завадостійкості залежить від індексу модуляції та може досягати більше 8 при індексі модуляції більше 3.

Список літератури

1. Кичак В.М. Синтез частотно-імпульсних елементів цифрової техніки. Монографія. – Вінниця: УНІВЕРСУМ-Вінниця. – 2005. – 291 с.

2. Бондарєв А.П. Критерій граничної завадостійкості системи фазової синхронізації. /А.П.Бондарєв, О.В.Капшій. //Вісник ДУ «ЛП» Радіоелектроніка та телекомунікації. – 1998. – №352. –С. 98-100.

3. Помехоустойчивость и эффективность систем передачи информации. / Под ред. А.Г. Зюко. – М.: Радио и связь. – 1985. –272 с.

Відомості про авторів

Бортник Геннадій Григорович, доцент, к.т.н, проф. кафедри ТКСТБ, ВНТУ.
Кичак Володимир Васильович, к.т.н., асистент кафедри ТКСТБ, ВНТУ.
Пунченко Наталя Олегівна, к.т.н доцент кафедри метрології Одеської національної академії зв'язку.

УДК 004.03; 004.3; 004.4; 004.5; 004.7

В. В. ЕРЕСЬКО, Н. В. НЕСТЕРЕНКО

Институт кибернетики им. В.М.Глушкова НАН Украины, г. Киев

ТИПОВЫЕ СИСТЕМЫ ИНФОРМАЦИОННОЙ ПОДДЕРЖКИ ДЛЯ ПРИМЕНЕНИЯ В РАЗЛИЧНЫХ ОБЛАСТЯХ

Анотація. Розглянуто сучасні типові системи інформаційної підтримки для автоматизації різноманітних технологічних процесів у промисловості, управлінні, науці, освіті та побуті. Запропоновано типові системи, що відрізняються від існуючих підходами до побудови та застосування. Запропоновано нову концепцію комплексної віртуалізації окремих елементів системи, підсистем і типових систем як складову частину загальної методології побудови систем інформаційної підтримки і складних комп'ютерних систем.
Ключові слова: системи інформаційної підтримки, віртуальні компоненти.
Аннотация. Рассмотрены современные типовые системы информационной поддержки для автоматизации различных технологических процессов в промышленности, управлении, науке, образовании и быту. Предложены типовые системы, отличающиеся от существующих подходами к построению и применению. Предложена новая концепция комплексной виртуализации отдельных элементов системы, подсистем и типовых систем как составная часть общей методологии построения систем информационной поддержки и сложных компьютерных систем.
Ключевые слова: системы информационной поддержки, виртуальные компоненты.
Abstract. The approach to development of information support system for automation of various processes in industry, administration, science, education and everyday life is described. The typical systems different from existing and based on new approaches to their construction and application are proposed. The new concept of comprehensive virtualization of the system components, subsystems and typical systems model proposed in article. This concept is a part of the methodology of development of information systems and complex computer systems.
Key words: system of information support, virtual devices.

Введение

При разработке современных систем информационной поддержки (СИП) и систем оперативного управления обычно преследуется цель повысить оперативность и эффективность принимаемых решений с максимальным использованием современных вычислительных средств, разнообразных информационных ресурсов, современных средств связи, обработки и передачи информации. При построении таких систем в целом, возникают и решаются дополнительные важные задачи – снять с конкретного пользователя системы большой объем рутиной работы по поиску, организации сбора, структуризации и хранению информации, организации взаимодействия конкретных аппаратно-программных устройств и другие.

В Институте кибернетики им. В. М. Глушкова НАН Украины постоянно проводятся теоретические и методологические исследования по созданию компьютерных систем различного назначения, разработке требований к их архитектурно-структурной организации, развитию концепций, принципов и методологии построения различных компьютерных систем и оптимизации информационных процессов. В ходе практических разработок, создаются варианты типовых СИП (например, [1]) и аппаратно-программных средств для построения разнообразных компьютерных систем и комплексов.

Особенностями современного этапа развития компьютерных систем являются повсеместное формирование единой информационной среды, внедрение процессов информатизации и виртуализации на всех уровнях во всех сферах управления, экономики, науки, образования и в быту. Информационные и коммуникационные технологии используются также при разработке эффективных технологических процессов производства услуг и товаров, построении систем государственного управления и организации оптимальной работы бизнес-структур, в том числе и организации виртуального взаимодействия субъектов хозяйствования различных экономических структур (например, [2]). Создаваемые системы позволяют обеспечить эффективность процесса поиска, накопления, передачи, отображения и дальнейшего использования информации.

В то же время современный этап характеризуется необходимостью обеспечения совместной работы и совместимости по различным критериям вновь создаваемых систем с уже работающими. Особенно важно это в условиях достаточно большого разнообразия современных аппаратных, аппаратно-программных и программных средств и отсутствия единых комплексных подходов и сложившейся методологии построения перспективных компьютерных систем, учитывающих все эти особенности.

Актуальность

Актуальность создания типовых конфигураций СИП различного назначения, устранения возникающих при этом трудностей, а также решение в целом проблем системной интеграции при построении таких систем определяется рядом важных факторов.

В настоящее время СИП в различном виде применяются практически во всех сферах и областях человеческой деятельности, среди которых можно выделить, например, такие как:

– поддержка принятия обоснованных решений в государственной сфере и бизнесе;

– построение корпоративных систем различного назначения;

– автоматизация управления технологическими процессами в производстве;

– предоставление разнообразных информационных услуг широкому кругу пользователей (в т.ч. в сфере государственного документооборота);

– информационное обеспечения нужд отдельных граждан (прогнозы погоды для места нахождения абонента, получение карты местности, разнообразной дорожной и туристической информации, путеводителей и другое);

– экологический мониторинг;

– комплексная автоматизация научных исследований и экспериментов;

– информационное обеспечение дистанционного обучения (в том числе – в образовании) и прочие.

Проблемы и задачи, которые сегодня приходится решать государственным органам управления, компаниям, разработчикам, научным работникам и другим лицам при принятии решений, имеют как интенсивную (качественную), так и экстенсивную (количественную) природу. Эти проблемы и задачи, отличаются для каждого конкретного потребителя по своему конечному результату, форме и характеру необходимой информации, выработанному на основе этой информации решению. Тем не менее, схемы получения информации, процесс выработки решений, исходные данные и другое имеют высокую степень пересечения, как в части сбора информации и ее источников, так и в части алгоритмов принятия решений и дальнейшей работы с информацией.

В связи с развитием технологий в последнее время сложилась и продолжает развиваться ситуация так называемого «информационного хаоса», требующая решения проблемы информационного упорядочивания и оптимизации процессов поиска необходимой информации, относящейся к различным аспектам деятельности государства, компаний разработчиков, научных учреждений, сферы образования.

Решение всего многообразия задач получения, хранения, обработки и предоставления информации в различных сферах привело к тому, что на современном этапе развития автоматизации обработки информации и создания автоматизированных систем управления появились разные концепции, а на их основе новые и разнообразные системы, у которых архитектура, организационная структура и решаемые задачи весьма многообразны. Данные системы для решения новых (возникающих) типов задач обычно реконфигурируются, являясь эволюционным развитием самих себя, или дополняются новыми компонентами в рамках интеграции общих корпоративных систем.

В последние годы происходят серьезные эволюционные изменения в части подходов к автоматизации работы с данными, информацией, документообороту, что связано со многими факторами, среди которых, как развитие новых технологий передачи, накопления, обработки и отображения информации (данных, знаний), так и лавинообразный рост ее объемов. Если новые технологии предоставляют возможности реализовать ранее недоступные алгоритмы и механизмы обработки информации (например, глобальное взаимодействие с партнерами, потребителями и другое), доступ к практически неограниченному ее объему, то возрастание объемов информации и знаний приводит к определенным трудностям при работе и принятии решений.

Одним из характерных признаков существующей и используемой ныне информации (данных, знаний) является форма ее представления – в настоящее время накопилась критически важная для бизнеса масса документов, существующих только в электронном виде (например, в виде электронных документов). При этом информация хранится, как в структурированном, так и в неструктурированном видах, что выдвигает новые требования к системам, да и к самим принципам, обработки информации. Подвергается определенной ревизии и сам понятийный аппарат или дифференциация между такими понятиями, как данные, информация, знания и так далее.

Появление класса систем информационной поддержки (СИП), реализующих различные интеллектуальные технологии обработки и предоставления информации пользователям, создание конкретных типовых вариантов реализации таких систем стало особенно актуальным на современном этапе, с появлением и развитием концепций сервис-ориентированных систем и ряда других технологий. Речь идет, в первую очередь, о массовом внедрении во всех сферах человеческой деятельности разнообразных систем принятия решений, информационно-поисковых систем, систем комплексной автоматизации различных технологических процессов в производстве, бизнесе, науке, образовании и в быту.

Цель

Основной целью работы являлась разработка элементов концепции и методов системной интеграции для эффективного проектирования и обеспечения работы типовых систем информационной поддержки, в том числе распределенных, с применением современных информационных технологий.

Кроме того, на основе анализа архитектурно-структурной организации существующих информационно-аналитических систем, систем поддержки принятия решений, систем автоматизации

различных технологических процессов, использования современных информационных технологий и технологий передачи данных предполагалось создать варианты типовых автоматизированных систем информационной поддержки, которые позволят (при использовании их самостоятельно или в качестве подсистем более сложных систем) отработать методику построения типовых СИП.

Решение задач

В Институте кибернетики им. В. М. Глушкова НАН Украины разработан ряд вариантов типовых систем информационной поддержки [1]. Созданные варианты СИП использовались как для проверки и подтверждения теоретических положений методологии создания СИП, так и для создания самостоятельных аппаратно-программных продуктов различного практического применения. Варианты типовых СИП постоянно дорабатывались в связи с дальнейшим развитием средств вычислительной техники, появлением высокопроизводительных мобильных персональных вычислительных средств, новых методов и стандартов передачи информации и мобильной связи [3], увеличением требуемых объемов информации, усилением требований к интерфейсу пользователя (таких, например, как «дружественность», «гибкость», «интуитивность», «универсальность»). Разработаны варианты типовых СИП с расширенными функциональными возможностями, с использованием новых информационных технологий, с изменяемой в зависимости от областей применения архитектурой и организационной структурой, созданы опытные образцы аппаратно-программных средств типовых СИП для некоторых областей применения. Доработана также методика создания самих СИП. Каждая из созданных типовых систем может работать как самостоятельно (локально или в сети) так и в составе более масштабной системы, когда можно рассматривать типовую систему как подсистему СИП более высокого уровня. При этом используются типовые интерфейсы пользователя, справочные подсистемы, базы данных и другие элементы подсистем, которые быстро конфигурируются под конкретные нужды пользователей и дополняются новыми элементами в случае необходимости.

Общая структура организации взаимодействия пользователей с типовыми СИП как подсистемами распределенных СИП более высокого уровня приведена на рис. 1. Методика и алгоритмы выбора оптимальной конфигурации аппаратных и программных средств (и аппаратно-программных платформ как основного базиса) подсистем СИП и СИП в целом зависят от конкретной сферы применения, задач, количества пользователей и могут гибко подбираться и настраиваться во время создания конкретного экземпляра системы.

Рисунок 1 – Пример организации взаимодействия пользователей с типовыми СИП

Разработаны и разрабатываются типовые автоматизированные системы информационной поддержки для принятия решений в сфере государственного управления, разных отраслях производства, для автоматизации различных технологических и бизнес процессов, экологического мониторинга, для решения разнообразных задач в сфере науки и образования и другие. Варианты типовых СИП разрабатываются как с использованием методов системной интеграции типовых программных и аппаратных средств, так и с применением программных и аппаратно-программных средств собственной разработки.

При разработке вариантов СИП и ее подсистем учитывались следующие основные требования:

– необходимость обеспечения мобильности (как процесса непрерывного получения необходимой информации из СИП в любой точке, где находится пользователь, даже при быстром перемещении относительно точек доступа);

– «дружественный» интерфейс пользователя, возможность (при желании пользователя) контролировать различные параметры (например, состояние подключения, уровень сигнала, информацию о параметрах поиска, статистику текущего сеанса обмена информацией с системой, скорость соединения, загруженность процессора);

– возможность формирования пользователем отдельных виртуальных информационных пространств и баз данных для каждой отдельной задачи или сеанса поиска информации;

– интуитивный интерфейс пользователя для удобного настраивания виртуального «информационного пространства» и для работы с большими объемами информации;

– возможность хранения и удобной для пользователя подачи разнородной информации;

– возможность формирования отчетов о работе в нужной форме;

– возможность экспорта полученных результатов во внешние базы данных или различные пакеты ПО для дальнейшей обработки;

– отсутствие необходимости ручной отладки пользователем «технических» параметров связи, выбора протоколов, каналов связи и так далее;

– возможность интеграции с различным технологическим оборудованием разных производителей и взаимодействия с программными и аппаратными компонентами систем учета и контроля, технологическим оборудованием (например, электронными весами, сканерами, датчиками) и др.

Еще одна особенность разработанных типовых систем – гибкая информационная «многоуровневость». Например, чаще всего используются основные три уровня. Для них можно распределить информацию и основное назначение каждого уровня таким образом:

– уровень «глоссария», когда на введенное поисковое слово или аббревиатуру пользователю предоставляется основное трактование, краткое объяснение, перевод или расшифровка;

– уровень более широкого объяснения, полной информации, углубление в «основные» детали и др.;

– уровень дополнительного «разъяснения», с использованием наглядного материала, ссылок на первоисточники, ссылок на различные ресурсы, связанные с проблемой или темой, дополнительный «внешний» поиск и прочее.

При необходимости, для каждого запроса отдельно можно во время работы увеличить количество уровней справочника или, при определенных требованиях к информации, наоборот – перераспределить и уменьшить. При этом важно, что всегда можно получить информацию о первоисточнике, библиографических и историографических ссылках, авторстве материалов и другое.

В ходе исследований и разработок предложена концепция виртуализации отдельных элементов системы, подсистем и типовых систем как составная часть методологии построения СИП и сложных компьютерных систем в целом, использующая единые подходы и методы практически на всех уровнях системы. Термин «виртуальное» используется в этом случае как понятие «созданное или расширенное за счет использования общих (разделенных, распределенных) компонентов и ресурсов». Таким образом, виртуальные элементы являются, по сути, временными наборами («квазипостоянными» множествами) ресурсов различного типа и назначения, что позволяет под конкретные условия или новую задачу быстро, эффективно и гибко создавать или перенастраивать их.

Принципы виртуализации отдельных компонентов систем, сервисов, создания виртуальных систем и лабораторий достаточно широко применяются в различных сферах, как информационной поддержки пользователей, так и в целом в сфере управления и создания бизнес-проектов и предприятий [4, 5]. В предложенной концепции используется широкое комплексное внедрение принципов виртуализации элементов компьютерных систем, в первую очередь – СИП, на всех уровнях и для всех элементов, что позволит использовать единые подходы к созданию новых систем, реконфигурации уже существующих, эффективному применению как систем в целом, так и их отдельных компонентов конкретными пользователями.

Информационные системы, как совокупность аппаратных и программных средств, не зависимо от характера возникающих новых задач, обычно эволюционируют не автономно, а в сложной взаимосвязи со всем комплексом информационного окружения в корпоративной среде. Предлагаемые подходы поз-

волят минимизировать затраты на реконфигурацию или переналадку отдельных компонентов общей системы за счет применения единых, внешних для отдельного компонента, структур данных и метаданных, протоколов обмена данными, принципов организации единого информационного пространства и, таким образом, уменьшить зависимость отдельных аппаратно-программных компонентов системы и ее подсистем от других.

Основа концепции построения СИП – это реализация многоуровневой виртуализации подсистем и апаратно-программных средств системы на нескольких основных уровнях, таких как, например:

– уровень виртуальных датчиков, приборов, органов управления, средств отображения информации, обработки и т.д., построенных в виде специализированных модулей на микроконтроллерах или на ПЛИС, самостоятельных или встраиваемых, которые обеспечивают роботу непосредственно с объектами мониторинга (управления), а также взаимодействие с пользователями определенных профессий, занятых обслуживанием данных объектов и оборудования;

– уровень виртуальных подсистем, в том числе распределенных, предназначенных для решения конкретных типовых задач пользователя для заданной предметной области с комплексным использованием разнообразных виртуальных средств разных уровней и элементов других подсистем;

– уровень типовых виртуальных рабочих мест пользователей (и операторов, которые учатся) на базе совместно используемых виртуальных средств и подсистем разных уровней;

– уровень типовых виртуальных информационных (рабочих) пространств пользователей (индивидуальных или для рабочей группы), которые используют общие данные подсистем разных уровней.

Количество уровней, их характеристики, аппаратно-программные средства и подсистемы, которые применяются, должны определяться при разработке конкретной системы или при реконфигурации существующей типовой системы под определенные задачи.

При принятии оперативных решений отдельными лицами, научными работниками, руководителями различного ранга, компаниями (как производителями технологий, так и поставщиками-интеграторами готовых продуктов) требуется учет множества факторов и аспектов, гибкое структурирование получаемой информации, конфигурирование формы ее представления и объемов в зависимости от возникшей задачи, часто кратковременной. В предложенной концепции именно виртуализация информационного пространства пользователя, – как в целом, так и для отдельных задач, – рассматривается как перспектива организации и построения подсистем в компьютерных системах. Разработка вариантов динамичных структур типовых виртуальных подсистем и организация в них взаимодействия аппаратно-программных средств для создания виртуальной информационной среды, где главными являются информационные потребности пользователя рассматриваются как перспективные направления создания систем информационного обслуживания пользователя.

Соединив в типовых подсистемах информационные объекты между собой и с «объектами реального мира» в модели, можно построить локализованные виртуальные информационные пространства (ВИП) для конкретных пользователей, задач, приложений, в том числе с геопозиционированным доступом к информации определенных узлов распределенной системы. При этом для пользователя или приложения создается эффект «прозрачности» распределенности: местонахождение функций, данных, их фрагментация и репликация явно не заметны пользователю. Средства позиционирования позволяют осуществлять доступ в требуемое «виртуальное информационное пространство», в зависимости от местности, для создания связи между ВИП и «фрагментом реального мира». Такой подход позволяет оптимизировать использование аппаратно-программных средств СИП при создании, например, систем экологического мониторинга, управления сложными технологическими процессами и так далее.

Для конкретного узла или участка СИП, как нижнего, так и более высокого уровней, такие описания и модели могут содержать абстрактные и конкретные наборы параметров и данных для построения виртуальных средств, рабочих мест, подсистем СИП. Это, в свою очередь позволяет обеспечить виртуализацию и самоорганизацию информационных структур в зависимости, например, от моделей близлежащего участка СИП, используемой для мониторинга сенсорной сети, схемы маршрутизации данных. Таким образом, возможно гибкое построение виртуального объекта и модели внешнего объекта с учетом его и собственного позиционирования (например, местоположения).

Главной особенностью разрабатываемых подходов к построению СИП и их компонентов является ориентация всего процесса построения и использования систем на потребности пользователя, необходимость обеспечения эффективности, надежности, прозрачности или незаметности работы системы при выполнении пользователем его непосредственных задач.

В конечном итоге разработанные подходы к построению СИП и элементы методологии создания систем, где главными являются информационные потребности пользователя, позволят создавать виртуальные информационные пространства множества пользователей, необходимые и достаточные для решения их круга задач. Различные комбинации таких информационных пространств будут составлять, в свою очередь, фрагменты общего (глобального) информационного пространства общества.

При этом на всех уровнях, от виртуальных элементов до систем и предприятий, можно выделить общие черты, такие как, например:

– высокая интеграция средств, ресурсов, опыта и другого в рамках необходимой целесообразности;

– организация виртуальных элементов вокруг ключевых задач, проектов, производственных и других процессов на время жизненного цикла задачи, процесса или проекта;

– создание автономных или совместно работающих элементов, рабочих мест, рабочих групп, территориально удаленных друг от друга;

– возможность быстрого создания, развития, гибкого переструктурирования и расформирования в нужное время любого элемента любого уровня;

– распределенное управление, возможность централизации и децентрализации управления компонентами СИП по необходимости;

– возможность обмена информацией между различными виртуальными компонентами, совместное использование одних и тех же источников данных разными пользователями на всех уровнях, с широким распределением возможностей, прав, полномочий и другое;

− использование неоднородных аппартно-программных, компьютерных (вычислительных) средств, различных сетей и технологий передачи данных;

– возможность гибкой организации группового взаимодействия специалистов различных областей знаний или учащихся, в том числе и в сетевых чатах, теле- и видеоконференциях [6, с. 29];

– полная ориентация на конкретного пользователя для обеспечения максимально быстрого выполнения его задач, наиболее полного удовлетворения его потребностей на основе интеграции всех необходимых средств и ресурсов при помощи новейших информационных и телекоммуникационных технологий и мобильных персональных вычислительных средств;

– формирование из виртуальных пространств различных пользователей и подсистем разных уровней единого информационного пространства с исчерпывающей информацией, необходимой для анализа задачи, подготовки управленческих решений, решения разнообразных задач пользователей по необходимым направлениям и др. И наоборот – возможность из существующего единого информационного пространства (набора ресурсов) сформировать виртуальное пространство как срез или сочетание необходимых и достаточных данных по запросу в различных видах представления информации [4].

Примеры типовых вариантов структурной организации виртуализированных элементов систем типа СИП для узлов системы на нижнем уровне, организации виртуальных рабочих мест и виртуальных рабочих групп из виртуальных компонентов на более высоких уровнях в общем информационном пространстве СИП приведены на рис. 2. В качестве примера использования принципа виртуализации для построения общедоступных компонентов на самом нижнем уровне системы показан комплексный виртуальный прибор (ВП) для сбора данных (для работы используются виртуальные датчики веса, температуры, считыватели штрих-кодов). При построении виртуальных автоматизированных рабочих мест (ВАРМ) пользователей, решающих различные производственные, управленческие или другие задачи, используются виртуальные приборы или разделяемые ресурсы других ВАРМ. При таком подходе, на базе персонального мобильного вычислительного средства одного пользователя могут быть сконфигурированы несколько виртуальных рабочих мест для решения как ряда смежных (похожих), так и нескольких в принципе различных самостоятельных задач. На основе нескольких ВАРМ возможно создание виртуальных рабочих групп (ВРГ), которые могут формироваться для решения конкретных задач не только из специалистов разных подразделений одной организации, но и вообще из независимых людей разных профессий, находящихся в различных географических точках. Вопросы организации ВРГ, прав доступа, договоренностей, доверия и другие аспекты создания виртуальных рабочих групп и организаций являются темой отдельных исследований. Показанные элементы, образующие различные виртуальные СИП и ее подсистемы, в частности мобильные СИП, могут работать как самостоятельно, так и являться подсистемой или фрагментом более масштабной системы. Алгоритмы работы коммутаторов-маршрутизаторов (КМ) и протоколы передачи информации в данной работе не рассматривались. Другие средства, обеспечивающие связь элементов систем и их информационное взаимодействие, на рис. 2 не показаны. В общем случае они вообще могут принадлежать, например, операторам Интернета или общедоступных сетей связи, телекоммуникационным каналам независимых провайдеров.

Системы создаются или реконфигурируются из типовых СИП под конкретную задачу, набор или круг задач или решаемых проблем интеграцией необходимых аппаратно-программных, информационных, организационно-технологических, человеческих, телекоммуникационных и других ресурсов. Виртуальность структуры такой системы можно представить, как временно организованное для работы одного или нескольких пользователей взаимодействие реально существующих общедоступных компонентов и ресурсов (общего или потенциально общего пользования). Взаимодействие специалистов организуется через средства телекоммуникации, позволяя создавать виртуальные подразделения или организации для выполнения общей задачи или решения определенной проблемы.

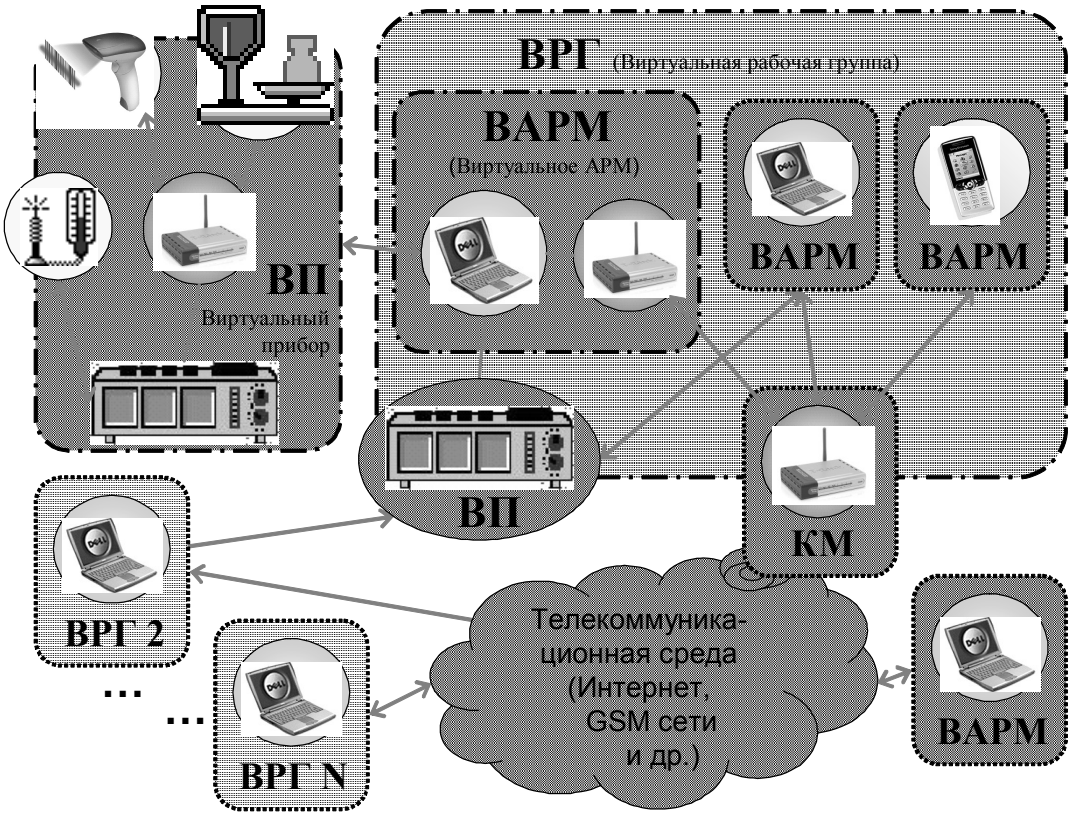


Рисунок 2 – Примеры организации взаимодействия виртуализированных элементов системы нижнего уровня, виртуальных рабочих мест и виртуальных рабочих групп в информационном пространстве СИП

Такое подразделение не имеет отраслевых или ведомственных ограничений и позволяет организовывать эффективную работу и гибкую оперативную реакцию на различные факторы. В качестве примера уже применяемых подходов виртуализации можно привести «виртуальные лаборатории» в науке и промышленности или «виртуальные классы» в дистанционном образовании.

Предложенные принципы построения виртуальных динамичных структур аппаратно-программных средств и подсистем с нефиксированной архитектурой и территориальной организацией можно рассматривать как основу методологических подходов к построению перспективных СИП и других компьютерных систем различного назначения, как современную концепцию построения виртуальных групп, организаций и предприятий для решения различных конкретных задач.

Выводы

1. Предложены универсальные компьютеризированные системы информационной поддержки принятия решений и автоматизации технологических процессов и задач управления, отличающиеся от существующих внедрением единых подходов к созданию их компонентов. Эти подходы характеризуются комплексным применением современных методов и методик виртуализации отдельных компонентов и подсистем всех уровней, которые позволяют эффективно, в относительно короткие сроки, с минимальными финансовыми затратами создавать новые или гибко перенастраивать существующие системы под новые, в том числе кратковременные, задачи сфер производства, автоматизации различных технологических процессов, процессов управления, образования, научных исследований и в быту. Возможность гибко создавать в разработанных системах виртуальные компоненты, виртуальные рабочие места, виртуальные рабочие группы, виртуальные информационные пространства пользователя или задачи позволяют эффективно решать широкий круг разнообразных задач.

2. Разработанные варианты типовых СИП отличаются от известных систем подобного назначения комплексной виртуализацией компонентов системы на всех ее уровнях, возможностью создавать и гибко реконфигурировать для решаемых задач временные виртуальные информационные пространства пользователя, возможностью изменять форму представления и объемы получаемой информации, комплексными подходами к обеспечению работоспособности системы и ее аппаратно-программных компонентов, возможностью оперативно получать информацию на портативные и мобильные устройства

(в том числе на планшеты, КПК, смартфоны и мобильные телефоны), универсальностью типовых приложений, использованием для работы пользователя стандартных общедоступных браузеров, офисного ПО и типовых средств разработки.

3. Варианты разработанных типовых СИП и их компоненты построены так, что они в короткое время могут быть легко настроены, доработаны и адаптированы для решения типовых задач: информационной поддержки и автоматизации в управлении при принятии решений; на производстве для автоматизации разнообразных технологических процессов; в науке для автоматизации научного эксперимента; создания разнообразных справочных системы, систем дистанционного обучения, дистанционных виртуальных лабораторий для образования; в быту (формирование справочников, баз данных, баз знаний, обеспечение различной геозависимой информацией и другое).

4. При построении разнообразных АРМ, адаптации и реконфигурации типовых СИП относительно легко учитывать условия и требования заказчика на уровне формирования метаданных для конфигурации системы, ее отдельных компонентов и интеллектуальных интерфейсов пользователя. Также упрощается решение проблемы обеспечения совместимости как на уровне протоколов, системного и прикладного ПО, форматов данных и др., так и на уровне электромагнитной совместимости (каналы связи, частоты, помехи и др.) нескольких устройств, работающих рядом или удаленных друг от друга. Большое внимание уделяется обеспечению надежности, непрерывности и бесперебойности их работы, независимо от характера и объема выполняемых задач, как всей системой, так и отдельными узлами, подсистемами и элементами.

5. В ходе работы были исследованы существующие типы информационных систем различного назначения, их архитектурно-структурная организация и концепции построения. Системы для принятия решений в сфере бизнеса, государственного управления, проектирования, науки, образования выделены как отдельные самостоятельные классы. Определены тенденции и перспективы развития таких систем, варианты архитектурно-структурных решений для реализации СИП и их подсистем.

Список литературы

1. Интегрированная аналитическая система информационной поддержки (АСИП) решения задач в области СВТ: концепция, реализация, перспективы / Коваль В. Н., Палагин А. В., Яковлев Ю. С., Валькман Ю. Р., Ересько В. В., Баринова И. А., Нестеренко Н. В., Курзанцева Л. И., Карпман Л. Я., Якуба А. А. // Проблемы программирования. – 2000. – № 1–2. – С. 398–408.

2. Веремеєнко Н.В. Стимули розвитку віртуальних підприємств в умовах інформатизації України // Вісник соціально-економічних досліджень. – № 34. – С. 32–36.

3. Шахнович И. Мир связи. Современные технологии беспроводной связи / И. В. Шахнович. – 2-е изд., испр. и доп. – М. : Техносфера, 2006. – 288 с. – ISBN 5-94836-070-9.

4. Вютрих Х.А. Виртуализация как возможный путь развития управления / Х. А. Вютрих, А. Ф. Филипп // Проблемы теории и практики управления. – 1999. – № 5. – С. 10.

5. Организация виртуальных предприятий : монография / Л. А. Тимашова, С. К. Рамазанов, Л. А. Бондар, В. А. Лещенко. – Луганск : Вид-во СНУ ім. В. Даля, 2004. – 368 с.

6. Макарова М. Віртуальні підприємства і телеробота як нові соціально-економічні явища // Економіст. – 2003. – № 12. – С. 27–32.

Сведения об авторах

Ересько Виталий Владимирович – главный инженер-электроник отдела Института кибернетики им. В.М.Глушкова НАН Украины, (+38044)-526-32-07, г. Киев, пр. Академика Глушкова, 40.

Нестеренко Николай Васильевич – научный сотрудник Института кибернетики им. В.М.Глушкова НАН Украины, (+38044)-526-32-07, г. Киев, пр. Академика Глушкова, 40.

УДК004.032.26

О. К. КОЛЕСНИЦЬКИЙ, Ю. В. ЛЕВЧЕНКО

Вінницький національний технічний університет, м. Вінниця

ІМПУЛЬСНИЙ НЕЙРОННИЙ ЕЛЕМЕНТ НА СТАБІЛІТРОНІ

Анотація. У даній роботі проведено дослідження та запропоновано нову спрощену схему моделінейронного елемента, а також проведено перевірку на відповідність пропонованої схеми моделі нейрону до математичної моделі LIF-нейрону. Для створення схеми моделі нейронного елементу було використано середовище моделювання NI Multisim, а для відображення математичної моделі, відповідно, середовище CSIM. Проведене дослідження комп'ютерної моделі запропонованої схеми нейронного елемента на стабілітроні показує, що його передатна характеристика адекватна передатній характеристиці LIF-нейрону з великою точністю. Запропонована схема нейронного елемента може використовуватися для побудови апаратних реалізацій імпульсних нейронних мереж з великою кількістю елементів.

Ключові слова: нейросистеми;нейроннімережі;LIF-нейрон;CSIM;Multisim; оптичні входи та виходи.

Аннотация.В данной работе проведено исследование и предложено новую упрощенную схему модели нейронного элемента, а также проведена проверка на соответствие предлагаемой схемы модели нейрона к математической модели LIF-нейрона. Для создания схемы модели нейронного элемента было использовано среду моделирования NI Multisim, а для отображения математической модели, соответственно, среду CSIM. Проведенное исследование компьютерной модели предложенной схемы нейронного элемента на стабилитроне показывает, что его передаточная характеристика адекватна передаточной характеристике LIF-нейрона с большой точностью. Предложенная схема нейронного элемента может использоваться для построения аппаратных реализаций импульсных нейронных сетей с большим количеством элементов.

Ключевые слова:нейросистемы; нейронные сети; LIF-нейрон; CSIM; Multisim; оптические входы и выходы.

Abstract. In current worka research was held anda new simplified circuit model of neural element was proposed as well as verification for correspondence of the proposed neuron model scheme with the mathematical model of LIF-neuron. A simulation environment NI Multisimwas used for the creation of neural element scheme model and the environment CSIM was usedfor the mathematical modeldisplaying.The performed research of computer model for the proposed scheme of neural element created usingzener diodeshows that itstransmitting characteristic is adequate to according characteristic of the LIF-neuronwith great accuracy. The proposed neural element scheme can be used to construct hardware implementations of spiking neural networks with a large number of elements.

Key words: neural system; neural networks; LIF-neuron; CSIM; Multisim; optical inputs and outputs.

Вступ

Зацікавленість штучними нейронними мережами швидко виросла за останні декілька років. Фахівці з багатьох областей, таких як технічне конструювання, філософія, фізіологія і психологія, заінтриговані можливостями, що надає ця технологія, і шукають застосування їм всередині своїх дисциплін. Максимум переваг від нейронних мереж можна отримати при їх апаратній реалізації [1]. Чим більше буде нейронів у мережі, тим складніші завдання можна буде вирішувати [2]. Але більшість відомих схем нейронних елементів на сьогодні побудовані на основі мікроелектронної технології [3,4], а тому використовують електричні зв'язки при формуванні із них мереж нейронів. Планарна природа напівпровідникових чипів не дозволяє створювати велику кількість зв'язків між нейронами в межах кристалу. Вирішити цю проблему можна при використанні оптичних зв'язків між нейронами в мережі. Тому існує задача створення схем нейронних елементів, в яких введення нових вузлів та зв'язків надасть змогу працювати з оптичними сигналами. Завдяки наявності оптичних входів та виходів схема буде зручна у використанні при побудові нейронних мереж з великою кількістю елементів та зв'язків між ними. Це розширює функціональні можливості схеми, оскільки організація великої кількості оптичних зв'язків реалізується набагато простіше, ніж електричних, за допомогою оптичних та голографічних засобів. Але мало створити схему нейронного елемента, треба ще довести, що її функціонування відповідає математичним моделям нейронів [5], на яких будуються потужні теоретичні засади функціонування імпульсних нейронних мереж [6]. Тому виникає необхідність проведення моделювання розроблених схем нейронних елементів та перевірка відповідності створених схем математичним моделям нейронів.

Постановка задачі

На сучасному етапі, на жаль, переважає підхід до створення нейросистемна основі формальних нейронів, які оперують статичними бінарними чи аналоговими сигналами на відміну від біологічних нейронів, які оперують імпульсними сигналами [7]. Тому найперспективнішими для технічної реалізації є саме імпульсні нейрони [1,2].

Нейронні елементи і мережі можна реалізовувати:

- програмно;
- програмно-апаратно;
- апаратно.

Максимальна перевага – створення реальної нейронної мережі. Чим більше нейронних елементів буде в мережі – тим краще. На даний час існує велика кількість різних варіантів нейронних елементів: механічних, магнітоелектричних, термоелектричних, квантових, на основі надпровідних матеріалів, на транзисторних схемах, на приладах з від'ємним опором – R-негатронах, на біспін-приладах, на операційних підсилювачах, на цифрових мікросхемах та мікроконтролерах. Складні фізичні моделі нейронних елементів дозволяють найбільш точно відобразити всі процеси та функціонування

біологічних нейронів, проте містять велику кількість елементів і створення великих масивів таких нейронів є занадто складною технічною задачею. Альтернативний підхід полягає у створенні якомога простіших апаратних реалізацій нейронних елементів зі збереженням головних функції нейрону. Найперспективнішою елементною базою для створення нейронних елементів є функціональні електронні прилади, використання яких забезпечить схемотехнічну простоту, високу надійність, економічність, технологічність, малі габарити та вагу [1].

Необхідно використовувати прості схеми апаратної реалізації нейронів. З іншого боку виникає необхідність, щоб створені схеми відповідали математичним моделям нейронних елементів, для яких уже доведено теореми повноти та представимості[6,7]. Також роздільна та апроксимаційна властивості імпульсних нейронних мереж доведені саме для LIF-нейронів [5]. Роздільна властивість передбачає, що будь-які дві вхідні функції мережі, які значно різняться, спричиняють неперервні стани мережі, які також значно різняться. Апроксимаційна властивість передбачає, що елемент зчитування має здатність апроксимувати будь-яку дану неперервну функцію f, що відображає поточні неперервні стани мережі у поточні виходи [6,7].

Мета

Мета статті – дослідження схемотехнічної реалізації імпульсного нейрону на стабілітроні та доведення її адекватності математичній моделі LIF-нейрону.

Аналіз існуючих моделей нейронних елементів

Є відомі схеми нейронних елементів, але кожна із них має ряд недоліків.

Пристрій для моделювання нейрона (патент України №55921[8]), який містить джерело живлення, два фотодіода,три резистора, три конденсатори, тиристор, світлодіод. Недоліком даного пристрою є складність схемо технічної реалізації, низька швидкодія та великі струми перемикання.

Інший відомий пристрій для моделювання нейрона (патент України № 22956[9]) містить два фотоприймачі, електронний ключ, біспін-прилад, резистор навантаження, модулятор лазерного випромінювання, лазер та голограму, на якій записано зовнішні зв'язки моделі нейрону. Недоліком даного пристрою також є його апаратна складність. Детальний аналіз відомих схем нейронних елементів наведено в [10].

Зазначені вище недоліки вимагають розробки як найпростішої схеми нейронного елемента, що зможе легко використовуватися при побудові імпульсних нейронних мереж з великою кількістю елементів та зв’язків між ними. Це й розширить функціональні можливості нейронного елемента, оскільки організація великої кількості оптичних зв'язків реалізується набагато простіше, ніж електричних, за допомогою оптичних та голографічних засобів.

Схема імпульсного нейронного елемента на стабілітроні

Інтелектуальні системи на основі штучних нейронних мереж дозволяють ефективно вирішувати задачі розпізнавання образів, прогнозування, оптимізації, діагностики, кластеризації, асоціативної пам’яті та керування. Штучний нейрон є елементарним функціональним модулем, з множини яких будуються штучні нейронні мережі. Він являє собою модель біологічного нейрона, але не в сенсі способу функціонування, а лише здійснення відповідних перетворень над вхідними сигналами. Імпульсний нейрон виконує зважене підсумовування (інтегрування) вхідних сигналів до досягнення цією сумою певного заданого порогу, після чого на виході нейрона формується вихідний імпульс, а зважена сума обнулюється.

Авторами була запропонована модель нейрона, яка містить порогів елемент, виконаний у вигляді стабілітрона, резистор та конденсатор, який під'єднаний паралельно до стабілітрона та резистора, вивід стабілітрона є виходом моделі, два фотодіоди та світловипромінювач, джерело живлення [11]. На рис. 1 зображено схему запропонованої моделі нейрона. Вона містить конденсатор 1, стабілітрон 2, резистор 3, перший 4 та другий 5 фотодіоди, світловипромінювач 6, джерело живлення 11. Активним елементом моделі є стабілітрон 2, оптична апертура першого фотодіода 4 є входом 7 для всіх збуджувальних сигналів моделі нейрона, оптична апертура другого фотодіода 5 є входом 8 для всіх гальмівних сигналів моделі нейрона, оптична апертура світловипромінювача 6 є оптичним виходом 9 моделі нейрона, анод стабілітрона 2з'єднаноз електричним виходом 10 моделі нейрона [5].

Модель нейрона працює таким чином. Перший фотодіод 4 приймає збуджувальні вхідні оптичні сигнали, а другий фотодіод 5 приймає гальмівні вхідні оптичні сигнали. Збуджувальні вхідні оптичні сигнали перетворюються першим фотодіодом 4 в струм, що заряджає конденсатор 1, а гальмівні вхідні оптичні сигнали перетворюються другим фотодіодом 5 в струм, що розряджає конденсатор 1. Вихідний імпульс формується на резисторі 3 (електричний імпульс) та світловипромінювачі 6 (оптичний імпульс) в момент, коли напруга заряду конденсатора 1 під впливом вхідних сигналів досягає порогового значення, достатнього для збудження в стабілітроні 2 процесу лавинного множення. В момент досягнення порогу провідність стабілітрона 2 різко збільшується, в результаті чого на електричному

виході 10 виникає імпульс напруги, який затухає по мірі розряду конденсатора 1 (світловипромінювач 6 формує в цей час вихідний оптичний імпульс). Розрядження конденсатора 1 викликає запирання стабілітрона 2 і процес повторюється знов. Частота вихідних імпульсів моделі прямо пропорційна різниці між фотострумом першого фотодіода 4 та другого фотодіода 5 [11].

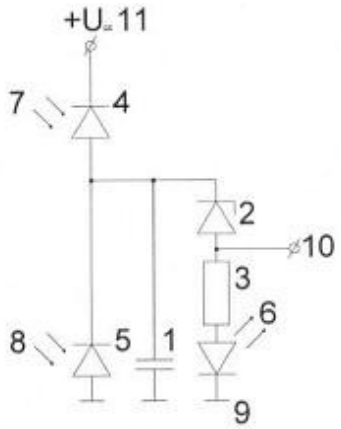


Рисунок 1 – Модель нейрона

Як збуджувальні, так і гальмівні сигнали моделі нейрону є оптичними, тобто мають однакову додатну полярність (як у біологічних нейронів) на відміну від відомих схем [3,4], де збуджувальні електричні сигнали повинні мати додатну полярність, а гальмівні електричні сигнали повинні мати від'ємну полярність. Наявність однакової полярності збуджувальних та гальмівних сигналів підвищує адекватність цієї моделі своєму біологічному прототипу.

Запропонована схема нейронного елемента дає змогу працювати з оптичними сигналами, що розширює функціональні можливості нейронного елемента та дозволяє мати збуджувальні та гальмівні сигнали однакової полярності, а це підвищує адекватність моделі до свого біологічного прототипу. Завдяки наявності оптичних входів та виходів дана модель може легко використовуватись при побудові імпульсних нейронних мереж з великою кількістю елементів та зв'язків між ними, оскільки організація великої кількості оптичних зв'язків реалізується набагато простіше, ніж електричних, за допомогою оптичних та голографічних засобів.

Моделювання схеми імпульсного нейронного елементу в середовищі NI Multisim

Для того, щоб зробити висновок про можливість застосування запропонованої схеми нейронного елемента на стабілітроні у нейронних мережах на її основі, потрібно перевірити відповідність функціонування цієї схеми та математичної LIF-моделі нейрону [5]. Для цього треба мати передатні характеристики обох моделей нейрона – фізичної моделі на стабілітроні[11] та математичної LIF-моделі нейрону [5]. Для отримання передатної характеристики нейронного елементу на стабілітроні було створено її комп'ютерну модель у програмному середовищі NIMultisim[12], яку представлена на рис. 2.

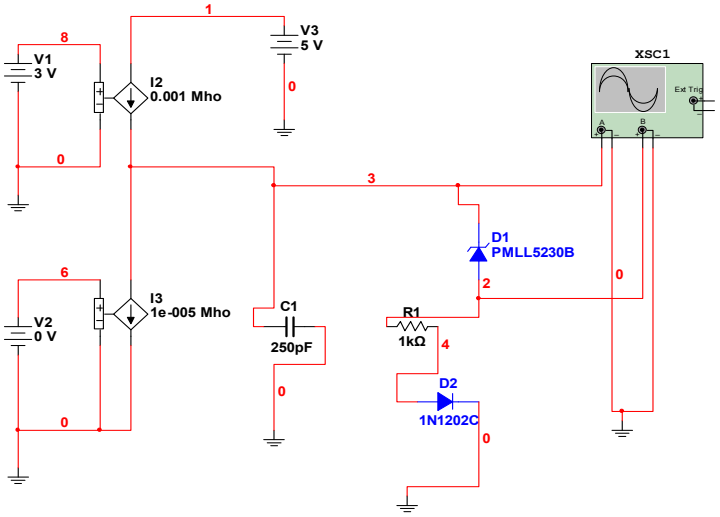


Рисунок 2 – Комп'ютерна модель схеми нейронного елементу на стабілітроні

Схема складається з джерела напруги збуджувального сигналу V1 та джерела напруги гальмівного сигналу V2, керованих напругою джерел струму I2 та I3, джерела живлення V3, конденсатора C1, стабілітрона D1, резистора R1 та світловипромінювача D2. Для контролю форми імпульсів використовується осцилограф. Внаслідок проведеного моделювання отримано залежність частоти вихідних імпульсів нейронного елемента від струму збудження. Комп’ютерне моделювання схеми нейронного елемента проводилось при таких параметрах схеми: V1=0…3V, V2=0V, V3=5V; коефіцієнт передачі керованого напругою джерела струму 0,001 Mho; конденсатор ємністю 250pF; резистор – 1 kΩ.

Порівняння передатних характеристик нейронів

Передатна характеристика нейронного елемента на стабілітроні отримана в попередньому пункті була порівняна із передатними характеристика LIF-нейрону [5], яку було отримано в середовищі моделювання CSIM [13]. Рисунок 3 відображає відповідно дві передатні характеристики: схеми на стабілітроні та LIF-нейрону. Ці графіки дозволяють надалі визначити ступінь адекватності запропонованої схеми нейронного елемента на стабілітроні до LIF-нейрона. Як індикатор адекватності візьмемо середнє відносне відхилення графіка передатньої характеристики запропонованої схеми нейрона на стабілітроні, отриманого в середовищі MULTISIM, від графіка передатної характеристики LIF-нейрона, отриманого у середовищі CSIM [13].

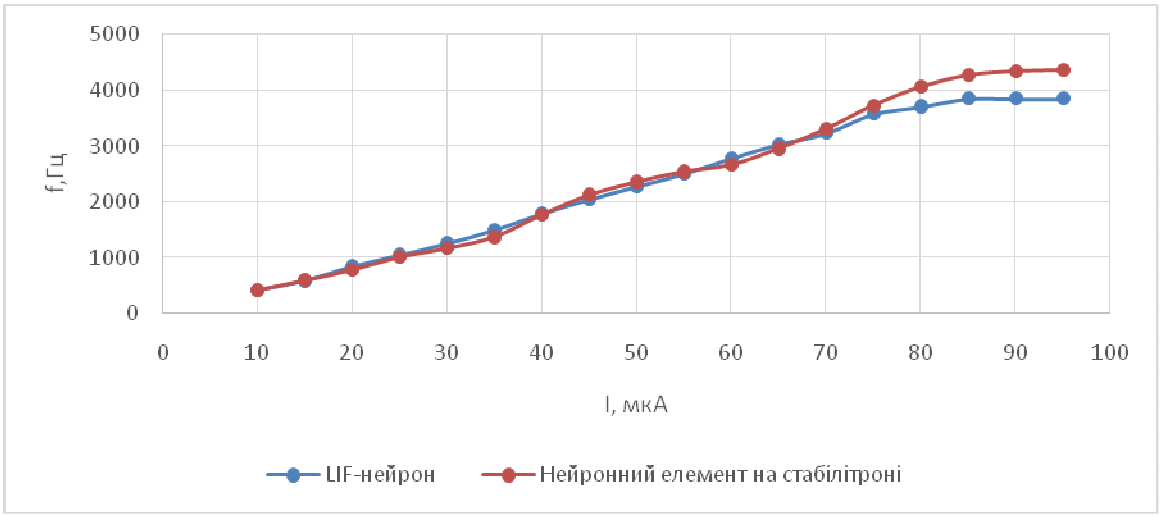


Рисунок 3 – Графік залежності передатних характеристик нейрону із середовищ CSIM та MULTISIM

Середнє відносне відхилення цих двох графіків розрахуємо за формулою (1):

$$b = \sum \frac{|(y_{ms} - y_{cs})|}{y_{cs}} \cdot \frac{1}{N} \cdot 100\%, \tag{1}$$

де b – середнє відносне відхилення, y_{ms} – значення передатньої характеристики запропонованої схеми нейрона на стабілітроні, отриманої в середовищі MULTISIM, y_{cs} – значення передатньої характеристики LIF-нейрона, отриманого у середовищі CSIM, N – кількість значень вимірювань.

Отже, за допомогою формули (1), підставивши усі необхідні параметри, можна остаточно визначити середнє відносне відхилення:

$$b = 0{,}925539 \cdot \frac{1}{18} \cdot 100 = 5{,}14\% \cdot$$

Цей результат свідчить про те, що передатна характеристика запропонованої схеми відрізняється від математичної моделі LIF-нейрона лише на 5,14%, що свідчить про високий ступінь їх адекватності.

Висновки

Проведене дослідження комп’ютерної моделі запропонованої схеми нейронного елемента на стабілітроні показало, що його передатна характеристика адекватна передатній характеристиці LIF-нейрону з точністю в межах 5,14%. Цедозволяє зробити висновок, що запропонована схема нейронного елемента може використовуватися для побудови апаратних реалізацій імпульсних нейронних мереж з великою кількістю елементів.

Перелік літературних джерел

1.Бардаченко В.Ф. Перспективи застосування імпульсних нейронних мереж з таймерним представленням інформації для розпізнавання динамічних образів / В.Ф. Бардаченко, О.К. Колесницький, С.А. Василецький - УСіМ.-2003-№6.- С. 73-82.

2. Kolesnytskyj O. K. Optoelectronic Implementation of Pulsed Neurons and Neural Networks Using Bispin-Devices / O. K. Kolesnytskyj, I. V. Bokotsey, S. S. Yaremchuk // Optical Memory & Neural Networks (Information Optics). – 2010. – Vol.19. – №2. – P.154–165. – ISSN 1060-992X.

3. DARPA SyNAPSEProgram [Електронний ресурс] / Дхармендра Мода. — Режим доступу: http://www.artificialbrains.com/darpa-synapse-program#256processor

4. NeuroGrid Project. Grid technology for neuroscience [Електроннийресурс] / Andrew Simpson. — Режимдоступу: http://www.neurogrid.ac.uk/

5.Gerstner W. Spiking Neuron Models. Single Neurons, Populations, Plasticity / Wulfram Gerstner, Werner M. Kistler. – Cambridge University Press, 2002. – 494p. – ISBN 0 521 89079 9 (paperback). – ISBN 0 521 81384 0 (hardcover). [Електронний ресурс]. — Режим доступу: http://diwww.epfl.ch/~gerstner/BUCH.html. – ISSN 0-521-89079-9.

6.W. Maass, T. Natschlaeger, and H. Markram. Computational models for generic cortical microcircuits. In Computational Neuroscience: A Comprehensive Approach, J. Feng, editor, chapter 18, pages 575-605. Chapman&Hall/CRC, BocaRaton, 2004.

7. Maass W. Networks of spiking neurons: the third generation of neural network models / W. Maass // *Neural Networks*, 1997, №10. — P. 1659-1671. – ISSN 0893-6080.

8. Пат. №55921 України. Пристрійдлямоделюваннянейрона / КолесницькийО.К., БокоцейІ.В., Павлов С.М., КолесницькаГ.О. Опубл. Бюл. №24, 27.12.2010.

9.Пат. №22956України. Модельнейрона/ КожемякоВ.П., КолесницькийО.К. НізельськийМ.Б., ВасилецькийС.А. Опубл. Бюл. №3, 5.05.98.

10. Бардаченко В. Ф. Таймернінейронніелементитаструктури. Монографія / В. Ф. Бардаченко, О. К. Колесницький, С. А. Василецький. – Вінниця : УНІВЕРСУМ-Вінниця, 2005, 126 с. – ISBN 966-641-109-1.

11.Пат. 76924 Україна, МПК G 06 G 7/60. Модель нейрона / О. К. Колесницький, Ю. В. Левченко, Г. О. Колесницька, заявник і власник патенту Вінницький нац. технічн. ун-т – № 201207181; заявлено 13.06.12; опубл. 25.01.13, Бюл.№2.

12. Хернитер М.Е. Multisim. Современная система компьютерного моделирования и анализа схем электронных устройств / М. Е. Хернитер – 2011 – С. 501.

13. CSIM: A Neural Circuit SIMulator Version 1.1 User Manual / The IGI LSM group. – Institute for Theoretical Computer Science, 2006. – 63p.

Інформація про авторів

Колесницький Олег Костянтинович, к.т.н., доцент кафедри комп'ютерних наук ВНТУ.
Левченко Юрій Вікторович, студент групи 1КН-14мі, ВНТУ.

УДК 519.876.5

С. Я. КРЕПИЧ, М. П. ДИВАК

Тернопільський національний економічний університет

ПОРІВНЯННЯ ЧАСОВОЇ СКЛАДНОСТІ РЕАЛІЗАЦІЇ ПРОЦЕДУР ВИПАДКОВОГО ПОШУКУ В ЗАДАЧІ СИНТЕЗУ ФІЛЬТРА ТА ДОПУСКІВ НА ПАРАМЕТРИ ЙОГО ЕЛЕМЕНТІВ

Анотація. В статті розглянуто приклад задачі синтезу аналогового фільтра при заданих допустимих значеннях вихідних характеристик. Обгрунтовано метод побудови цільової функції задачі та наведено приклад перетворення задачі синтезу до оптимізаційної задачі для пошуку хоча б одного розв’язку інтервальної системи нелінійних алгебричних рівнянь. Однак суттєвою різницею застосування зазначеного методу в праці є використання при обчисленні інтервалів значень вихідних характеристик не просто оцінки вектора номінальних значень параметрів (як це в класичних схемах), а певні обчислені інтервальні оцінки цього вектора, які обумовлені допусками на параметри елементів.

Ключові слова: методи випадкового пошуку, вихідна характеристика, допуски на параметри

Аннотация. В статье рассмотрен пример задачи синтеза аналогового фильтра при заданных допустимых значениях исходных характеристик. Обоснован метод построения целевой функции задачи и приведен пример преобразования задачи синтеза к оптимизационной задаче для поиска хотя бы одного решения интервальной системы нелинейных алгебраических уравнений. Однако существенной разницей применения данного метода в работе является использование при исчислении интервалов значений исходных характеристик не просто оценки вектора номинальных значений параметров (как это в классических схемах), а определенные вычисленные интервальные оценки этого вектора, которые обусловлены допусками на параметры элементов.

Ключевые слова: методы случайного поиска, исходная характеристика, допуска на параметры

Abstract. At the article considers the sample of the task of analysis analog filter for specified admissible values of output characteristics and tolerances on the parameters of its elements. Was grounded method for constructing of the objective function of task of the analog filter synthesis and presented example of converting synthesis problem to an optimization problem to find at least one solution ISNAE. However, a essential difference in the use of this method in work is the use for the calculation of interval values of output characteristics not only evaluation vector of nominal parameter values (as in the classical schemes), but some calculated interval valuations of this vector, which are caused by the specified tolerances for the parameters of element.

Keywords: procedures of random search, output characteristics, tolerance of the parameters.

Вступ

У традиційній постановці задачу синтезу радіоелектронних пристроїв математично формулюють як оптимізаційну задачу, в якої цільова функція забезпечує вимоги до характеристик синтезованого пристрою [1]. При розв’язуванні задачі синтезу з одночасним забезпеченням функціональної придатності радіоелектронних пристроїв необхідно обчислити параметри у такий спосіб, щоб з врахуванням їх відхилень, які задаються допусками, забезпечити функціональну придатність радіоелектронного пристрою. Зазначені задачі мають велику практичну значущість [2].

На сьогоднішній день існуючі методи розв’язку задач синтезу радіоелектронних пристроїв досить часто використовують процедури випадкового пошуку. Серед найбільш відомих методів синтезу радіоелектронних пристроїв при одночасному задані допустимих значень вихідних характеристик відомими є такі процедури випадкового пошуку: з лінійною тактикою; по найкращій спробі; із використанням направляючого конуса; із адаптацією розподілу випадкового кроку [3]. Разом з тим часову складність реалізації цих процедур саме для задач синтезу радіоелектронних пристроїв за умови заданих допустимих значень вихідних характеристик є малодослідженою.

Методи синтезу радіоелектронних пристроїв описані в працях[4, 6, 7]. Методи синтезу радіоелектронних пристроїв при заданих допустимих значеннях вихідних характеристик в яких використовуються також процедури випадкового пошуку наведені у працях [8, 9]. Проте, на сьогодні, відсутні публікації щодо обґрунтування вибору процедури випадкового пошуку за критерієм мінімальної часової складності при реалізації методів синтезу радіоелектронних пристроїв при заданих допустимих значеннях вихідних характеристик і одночасно заданих допусків на параметри його елементів.

Метою роботи є проведення порівняльного аналізу часової складності методів синтезу радіоелектронних пристроїв при одночасному задані допустимих значень вихідних характеристик та допусків на параметри його елементів з застосування різних процедур випадкового пошуку і на цій основі обґрунтування вибору методу з мінімальною часовою складністю. Для реалізації мети пропонується використати приклад синтезу аналогового фільтра.

Постановка задачі

Розглянемо приклад синтезу аналогового фільтра представленого схемою на рис. 1 при заданих допустимих значеннях його модуля коефіцієнта передачі на різних частотах.

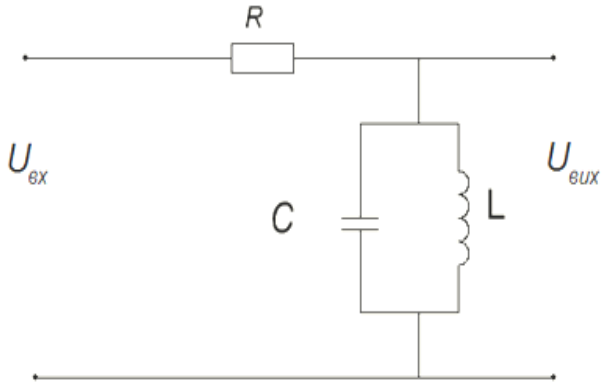


Рисунок 1 – Схема аланологового фільтра

Вираз для обчислення модуля коефіцієнта передачі фільтра, зображеного на рисунку 1, представлений формулою [4]:

$$\left|K(f)\right| = \frac{1}{\sqrt{R_0^2 \cdot (\frac{1}{2\pi \cdot f \cdot L_0} - 2\pi \cdot f \cdot C_0)^2 + 1}}, \quad (1)$$

де f - частота; R_0, C_0, L_0 - невідомі (в задачі синтезу) номінальні значення параметрів елементів фільтра, відповідно: резистора, конденсатора та котушки індуктивності.

За вихідні характеристики фільтра, які необхідно забезпечити в процесі синтезу приймемо модуль коефіцієнта передачі на різних частотах [5]. Задамо необхідні номінальні значення модуля коефіцієнта передачі $\left|K(f_i)\right| = K_{0i}, i = 1,..., N$ на різних частотах і допустимі відхилення від цих значень

$$[K_{0i}^-; K_{0i}^+] = [\left|K(f)\right| - \delta K_i \cdot \left|K(f)\right|; \left|K(f)\right| + \delta K_i \cdot \left|K(f)\right|], \quad (2)$$

де $\delta K_i > 0$ - відносне відхилення модуля коефіцієнта передачі від номінальних значень; $[K_{0i}^-; K_{0i}^+]$ - інтервали допустимих значень модуля коефіцієнта передачі. Тоді вважатимемо, що зазначене коло є функціонально придатним.

Позначимо допуски на усі параметри елементів фільтра та представимо вектор параметрів елементів фільтра в інтервальному вигляді:

$$[\vec{b}_0 + \vec{b}_0 \cdot \delta\vec{b}^-; \vec{b}_0 + \vec{b}_0 \cdot \delta\vec{b}^+] = ([R_0^-; R_0^+], [C_0^-; C_0^+], [L_0^-; L_0^+]) \quad (3)$$

де

$$\begin{aligned} [R_0^-; R_0^+] &= [R_0 - R_0 \cdot \delta R; R_0 + R_0 \cdot \delta R], \\ [C_0^-; C_0^+] &= [C_0 - C_0 \cdot \delta C; C_0 + C_0 \cdot \delta C], \\ [L_0^-; L_0^+] &= [L_0 - L_0 \cdot \delta L; L_0 + L_0 \cdot \delta L] \end{aligned} \quad (4)$$

$\delta R > 0,\ \delta C > 0,\ \delta L > 0$ - відносні відхилення параметрів елементів фільтра.

Зауважимо, що номінальні значення параметрів R_0, C_0, L_0 елементів фільтра є невідомими, також невідомими є їх допуски $\delta R, \delta C, \delta L$. Задамо обмеження на модуль коефіцієнта передачі для різних частот у вигляді:

$$\begin{cases} K_{01}^{-} \le \left|K(f_1)\right| \le K_{01}^{+} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ K_{0i}^{-} \le \left|K(f_i)\right| \le K_{0i}^{+} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ K_{0N}^{-} \le \left|K(f_N)\right| \le K_{0N}^{+} \end{cases} \quad (5)$$

Тепер підставимо в систему (5) інтервали $[R_0^-; R_0^+]$, $[C_0^-; C_0^+]$, $[L_0^-; L_0^+]$, які представляють допуски на параметри елементів разом з номінальними значеннями параметрів R_0, C_0, L_0 у формулі (1). Тоді отримаємо інтервальну систему нелінійних алгебричних рівнянь (ІСНАР):

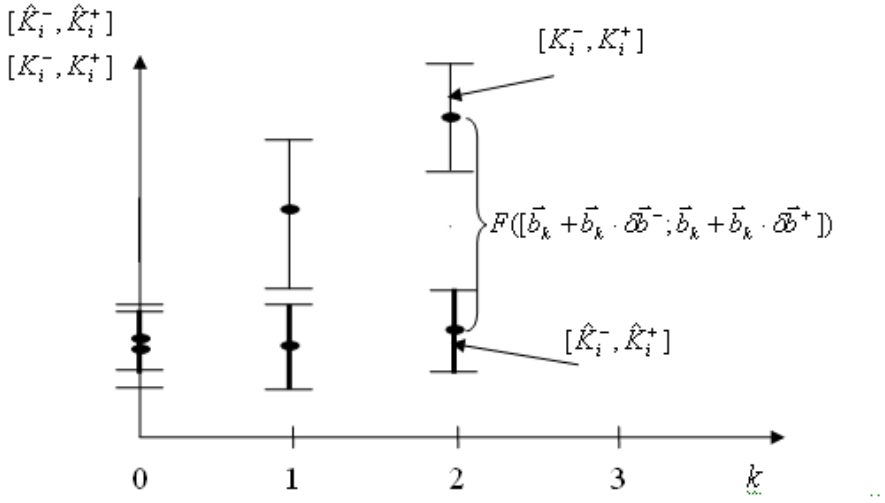
$$\begin{cases} K_{01}^{-} \le \dfrac{1}{\sqrt{[R_0^-; R_0^+]^2 \cdot (\dfrac{1}{2\pi \cdot f_1 \cdot [L_0^-; L_0^+]} - 2\pi \cdot f_1 \cdot [C_0^-; C_0^+])^2 + 1}} \le K_{01}^{+} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ K_{0i}^{-} \le \dfrac{1}{\sqrt{[R_0^-; R_0^+]^2 \cdot (\dfrac{1}{2\pi \cdot f_i \cdot [L_0^-; L_0^+]} - 2\pi \cdot f_i \cdot [C_0^-; C_0^+])^2 + 1}} \le K_{0i}^{+} \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ K_{0N}^{-} \le \dfrac{1}{\sqrt{[R_0^-; R_0^+]^2 \cdot (\dfrac{1}{2\pi \cdot f_n \cdot [L_0^-; L_0^+]} - 2\pi \cdot f_n \cdot [C_0^-; C_0^+])^2 + 1}} \le K_{0N}^{+} \end{cases} \quad (6)$$

Розв’язок цієї системи є не опуклою областю в просторі параметрів елементів радіоелектронних кіл. Тому будемо шукати деяку апроксимацію цієї області у вигляді вписаного прямокутного паралелепіпеда. Такий підхід забезпечує одночасне знаходження параметрів РЕК і їх допусків. Для знаходження зазначеного розв’язку використовуємо методи інтервальної арифметики. При цьому, враховуючи множинність розв’язків інтервальної системи нелінійних алгебричних рівнянь(6), для знаходження хоча б одного розв’язку із множини скористаємось, за аналогією, результатами праці [6], перетворивши зазначену задачу до дискретної оптимізаційної задачі:

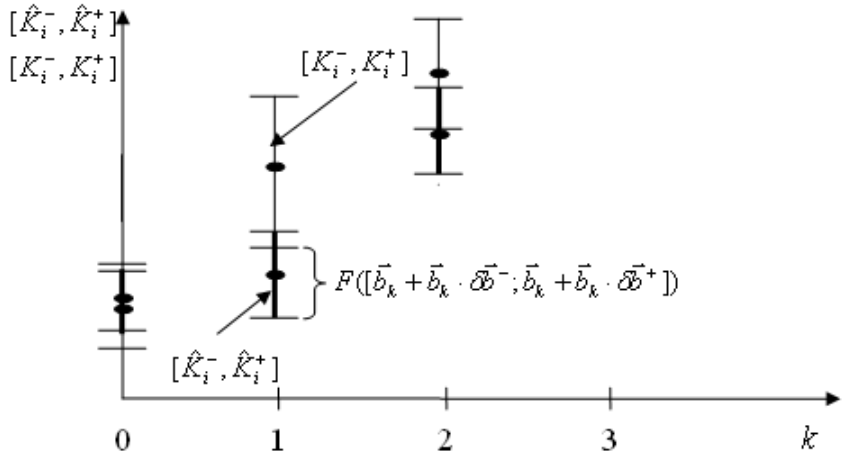
$$F([\vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^{-}; \vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^{+}]) \xrightarrow{\vec{b}_k, \delta\vec{b}_k} \min,\ \vec{b}_j > 0,\ \delta b_j > \delta b_j^z,\ j = 1..m \quad (7)$$

де δb_j^z - задані мінімально допустимі для даної схеми допуски на параметри елементів, $F([\vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^{-}; \vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^{+}])$ - значення функції мети, яка побудована на основі ІСНАР (6) і на кожній ітерації визначає досягнуту «якість» наближення оцінки вектора параметрів фільтра до шуканих номінальних значень. При цьому ітераційну процедуру оцінювання вектора параметрів фільтра необхідно організувати у такий спосіб, щоб забезпечити зменшення значень функції мети $F([\vec{b}_1]) > \ldots > F([\vec{b}_k]) > \ldots > F([\vec{b}_{k=K}] = [\vec{b}_0] \in \Omega)$ за скінчену та якомога меншу кількість ітерацій $k = K$.

Розглянемо детально процедуру перетворення задачі знаходження хоча б одного розв’язку ІСНАР (6) до оптимізаційної задачі (7). Скористаємось графічною ілюстрацією для пояснення процедури отримання функції мети.



а)



б)

Рисунок 2 – Ілюстрація для визначення функції мети в задачі синтезу: а) випадок $[\hat{K}_i^-; \hat{K}_i^+] \cap [K_i^-; K_i^+] = \varnothing$, б) випадок $[\hat{K}_i^-; \hat{K}_i^+] \cap [K_i^-; K_i^+] \neq \varnothing$

На рисунку 2 наведені задані інтервали допустимих значень на вихідні характеристики РЕП $[K_i^-; K_i^+]$ та інтервали допустимих значень на вихідні характеристики РЕП, отримані з одного розв'язку ІСНАР(6):

$$[\hat{K}_i^-; \hat{K}_i^+] = \frac{1}{\sqrt{[R_0^-; R_0^+]^2 \cdot (\frac{1}{2\pi \cdot f_i \cdot [L_0^-; L_0^+]} - 2\pi \cdot f_i \cdot [C_0^-; C_0^+])^2 + 1}} \quad (8)$$

Умовою забезпечення функціональної придатності, як витікає із системи (6), є $[\hat{K}_i^-; \hat{K}_i^+] \subset [K_i^-; K_i^+] \forall i = 1..N$.

Тоді, спираючись на результати праці [7], функцію мети $F([\vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^-; \vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^+])$ визначаємо, як різницю центрів найбільш віддалених між собою інтервалів $[\hat{K}_i^-; \hat{K}_i^+], i = 1,..., N$, отриманих для поточного наближення вектора оцінок параметрів $[\vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^-; \vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^+]$, та відповідних заданих інтервалів $[K_{0i}^-; K_{0i}^+]$ - у випадку, коли найбільш віддалені між собою інтервали

для *i*- ї вихідної характеристики не перетинаються. Формально зазначену умову запишемо у такому вигляді:

$$F([\vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^-; \vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^+]) = \max_{i=1,..,N} \left\{ \left| mid([\hat{K}_i^-; \hat{K}_i^+]) - mid([K_{0i}^-; K_{0i}^+]) \right| \right\}, \quad (9)$$

якщо $[\hat{K}_i^-; \hat{K}_i^+] \cap [K_{0i}^-; K_{0i}^+] = \varnothing \quad \exists i = 1,..., N$, де $mid(\cdot)$ - означає оператор виділення центру інтервалу.

Для випадків, коли усі інтервали $[\hat{K}_i^-; \hat{K}_i^+], i = 1,..., N$ й відповідні задані інтервали допустимих значень вихідних характеристик фільтра перетинаються, то спираючись на результати праці [6] функцію якості наближення $F([b_k + \vec{b}_k \cdot \delta\vec{b}^-; \vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^+])$ в задачі (7) визначатимемо у такому вигляді :

$$F([b_k + \vec{b}_k \cdot \delta\vec{b}^-; \vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^+]) = \max_{i=1,..,N} \left\{ \left| wid([\hat{K}_i^-; \hat{K}_i^+]) - wid([\hat{K}_i^-; \hat{K}_i^+] \cap [K_{0i}^-; K_{0i}^+]) \right| \right\}, \quad (10)$$

якщо $[\hat{K}_i^-; \hat{K}_i^+] \cap [[K_{0i}^-; K_{0i}^+] \neq \varnothing \quad \forall i = 1,..., N$, де $wid(\cdot)$ - означає оператор визначення ширини інтервалу.

Зазначені випадки зображені на рисунку 2 а) та б).

Перейдемо до розгляду методів розв’язування задачі (7). Одним з методів пошуку невідомого вектора параметрів аналогового фільтра є метод випадкового пошуку. Метод випадкового пошуку застосовано для подібних задач розглядався в багатьох працях, зокрема в [8]. Проте в нашому випадку задача суттєво ускладнюється через те, що при обчислені інтервалів значень $[\hat{K}_i^-; \hat{K}_i^+]$ вихідних характеристик фільтра використовуємо не просто оцінки $\vec{b}_k$ вектора номінальних значень параметрів $\vec{b}_0$ (як це в класичних обчислювальних схемах), а певні інтервальні оцінки $[\vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^-; \vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^+]$ цього вектора, обумовлені заданими допусками на параметри елементів фільтра.

Виходячи з вище викладеного, для розв’язування задачі синтезу аналогового фільтра у вигляді задачі оптимізації (4), пропонується розглянути та дослідити процедури випадкового пошуку, які описано у [9]:

- випадковий пошук із лінійною тактикою;
- випадковий пошук по найкращій спробі;
- випадковий пошук із використанням направляючого конуса;
- випадковий пошук із адаптацією розподілу випадкового кроку.

В основі обчислювальних схем розглянутих у [9] є правила урахування історії пошуку у послідовності $[\vec{b}_1] \to [\vec{b}_2] \to ... \to [\vec{b}_i] \to ... \to [\vec{b}_k]$. Зокрема, основним правилом, на яких ґрунтуються процедури випадкового пошуку та обчислювальні схеми оптимізації, є правило, що імовірність вдалого кроку у напрямку який був успішним на попередніх ітераціях, тобто коли функція мети зменшувалась, буде більшою ніж ймовірність вдалого кроку вибраного випадковим чином.

Процедура випадкового пошуку з лінійною тактикою (ВПзЛП). Випадковий пошук такого роду побудуємо за допомогою двох операторів: генерування випадкового вектора параметрів та їх допусків $\vec{B} = (\vec{b}_k, \delta\vec{b}_k)$, або повторення кроку за аналогією з попередньою ітерацією. Дія кожного із них може привести до одного з двох результатів: значення функції $F([\vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^-; \vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^+])$ або зменшиться, або залишиться без змін. Процедура випадкового пошуку з лінійною тактикою спирається на наступне очевидне припущення відносно об’єкта оптимізації: ймовірність вдалого кроку у напрямку який був успішним на попередній ітерації, тобто коли функція мети зменшувалась, буде більшою, ніж ймовірність вдалого кроку вибраного випадковим чином. При пошуку параметрів елементів фільтра це означатиме доцільність повторити успішний крок, а у випадку невдачі – обирати напрям пошуку випадковим чином. Лінійність тактики вказаної процедури полягає в імітації лінійної поведінки, тобто в прямому повторенні вдалого кроку[10].

Формально правило обчислення поточного значення параметрів фільтра у цій процедурі випадкового пошуку запишемо у вигляді:

$$\vec{B}_k = \vec{B}_{k-1} + \Delta\vec{B}_k \quad (11)$$

де

$$\Delta\vec{B}_k = \begin{cases} r\cdot\vec{\xi}, якщо F([\vec{b}_k+\vec{b}_k\cdot\delta\vec{b}^-;\vec{b}_k+\vec{b}_k\cdot\delta\vec{b}^+])-F([\vec{b}_{k-1}+\vec{b}_{k-1}\cdot\delta\vec{b}^-;\vec{b}_{k-1}+\vec{b}_{k-1}\cdot\delta\vec{b}^+])\geq 0 \\ \Delta\vec{B}_{k-1}, якщо F([\vec{b}_k+\vec{b}_k\cdot\delta\vec{b}^-;\vec{b}_k+\vec{b}_k\cdot\delta\vec{b}^+])-F([\vec{b}_{k-1}+\vec{b}_{k-1}\cdot\delta\vec{b}^-;\vec{b}_{k-1}+\vec{b}_{k-1}\cdot\delta\vec{b}^+])<0, \end{cases} \quad (12)$$

r – поточний радіус; $\vec{\xi}$ – одиничний випадковий вектор, рівномірно розподілений за усіма напрямами простору параметрів, що оптимізуються, та їх допусків; $\Delta\vec{B}_k, \Delta\vec{B}_{k-1}$ – величина кроку на k–тій та на k-1 ітераціях.

Вище зазначена схема має просту геометричну інтерпретацію. Фактично, це спуск із довжиною кроку r або у випадковому напрямку $\vec{\xi}$ у просторі оцінок вектора параметрів та їх допусків $\vec{B} = (\vec{b}_k, \delta\vec{b}_k)$, або із послідовним застосуванням оператора повторення по поверхні, яка задана функцією $F([\vec{b}_k+\vec{b}_k\cdot\delta\vec{b}^-;\vec{b}_k+\vec{b}_k\cdot\delta\vec{b}^+])$ у напрямку такому ж як на попередніх кроках, до тих пір, поки значення функції $F([\vec{b}+\vec{b}\cdot\delta\vec{b}^-;\vec{b}+\vec{b}\cdot\delta\vec{b}^+])$ не почне зростати. Це є сигналом того, що обраний напрям більше не перспективний, і потрібно обирати новий, - випадковий.

Розглянемо специфіку і можливості описаної обчислювальної схеми. Нехай p – ймовірність того, що величина випадкового кроку $\Delta\vec{B}_k = r\cdot\vec{\xi}$ успішна, а q – ймовірність повторення попереднього успішного кроку. Описана обчислювальна схема буде ефективною, коли часто «спрацьовує» оператор повторення кроку, так як в ці моменти гарантовано зменшується значення функції мети. Звідси випливає: для ефективної роботи даного алгоритму необхідно, щоб ймовірність q була більшою ніж p. Спираючись на властивості функції мети $F([\vec{b}+\vec{b}\cdot\delta\vec{b}^-;\vec{b}+\vec{b}\cdot\delta\vec{b}^+])$, можемо припустити, що ефективність розглянутої лінійної стратегії у процедурі випадкового пошуку буде висока для початкових ітерацій, коли оцінки параметрів є «грубими» і функція мети на цій ділянці є монотонно спадною.

Процедура випадкового пошуку по найкращій спробі (ВПпНС). Випадковий пошук такого роду побудований на обчислені значень функції мети в L точках, згенерованих випадковим чином на деякій локальній області пошуку. Набір цих точок представимо таким виразом:

$$\vec{B}_l = \vec{B}_{k-1} + r\cdot\vec{\xi}_l, l = 1,...,L \quad (13)$$

де $\vec{B}_l$ – одна інтервальна точка із набору, що претендують на «покращену» оцінку вектора $\vec{B}_k$ параметрів фільтра на k –тій ітерації; $\vec{\xi}_l$ – одиничний випадковий вектор, рівномірно розподілений за усіма напрямками простору параметрів, які оптимізуються.

Вибір напрямку $\vec{\xi}^*$ найкращої спроби визначаємо наступним чином:

$$F([\vec{b}_{k-1}+\vec{b}_{k-1}\cdot\delta\vec{b}^-;\vec{b}_{k-1}+\vec{b}_{k-1}\cdot\delta\vec{b}^+]+r\cdot\vec{\xi}^*) = \min_{l=1,...,L}(F([\vec{b}_{k-1}+\vec{b}_{k-1}\cdot\delta\vec{b}^-;\vec{b}_{k-1}+\vec{b}_{k-1}\cdot\delta\vec{b}^+]+r\cdot\vec{\xi}_l)) \quad (14)$$

Основна складність у використанні розглянутої обчислювальної схеми полягає у виборі числа L. Чим більше значення L, тим вищою буде обчислювальна складність реалізації методу із цією процедурою, оскільки збільшується кількість поточних оцінок параметрів фільтра, для яких необхідно обчислювати значення функції мети за формулами (9) чи (10). В той же час зменшення значення L може призвести до збільшення кількості ітерацій обчислювального методу і тим самим підвищити його часову складність.

Процедура випадкового пошуку із використанням направляючого конуса (ВПізВНК). Розглянемо обчислювальну схему у процедурі випадкового пошуку мінімуму функції мети $F([\vec{b}+\vec{b}\cdot\delta\vec{b}^-;\vec{b}+\vec{b}\cdot\delta\vec{b}^+])$ із використанням направляючого конуса [11].

На початковій ітерації випадкового пошуку (k=0) задаємо початкове наближення вектора параметрів $\vec{B}_0$. В околі цього наближення на поверхні уявної гіперсфери, радіусом r, тобто на відстані r від точки $\vec{B}_0$ в просторі параметрів на основі рівномірного закону розподілу генеруємо n випадкових точок:

$$\vec{B}_n = \vec{B}_0 + r \cdot \vec{\xi}_n, n = 1,..., N \tag{15}$$

Серед згенерованих точок вибираємо точку, яка забезпечує найменше значення функції мети:

$$\vec{B}_1 = \underset{n=1,...,N}{\arg\min}(F([\vec{b}_0 + \vec{b}_0 \cdot \delta\vec{b}^-; \vec{b}_0 + \vec{b}_0 \cdot \delta\vec{b}^+] + r \cdot \vec{\xi})). \tag{16}$$

Отримана оцінка вектора параметрів фільтра є наближенням для наступної ітерації. Додатково у цій процедурі обчислюємо вектор пам’яті, який визначає успішний напрям пошуку:

$$\vec{w} = (\vec{B}_1 - \vec{B}_0)/r. \tag{17}$$

На наступних ітераціях в просторі параметрів будуємо уявний гіперконус із вершинами в точках $\vec{B} = (\vec{b}_k, \delta\vec{b}_k)$, яка є поточною оцінкою вектора параметрів, з кутом розкриття ψ і віссю $\vec{w}_K$. Цей гіперконус «відсікає» від гіперсфери з центрами в точках $\vec{B} = (\vec{b}_k, \delta\vec{b}_k)$ і радіусом *r* деякі поверхні. На отриманих поверхнях в просторі параметрів та їх допусків генеруємо на основі рівномірного закону розподілу *n* випадкових точок за формулою (15), де вектор $\vec{\xi}_n$ в даному випадку обчислюємо виходячи із обмежень на параметри конусів. Серед згенерованих точок вибираємо точку, яка забезпечує найменше значення функції мети:

$$\vec{B}_{k+1} = \underset{n=1,...,N}{\arg\min}(F([\vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^-; \vec{b}_k + \vec{b}_k \cdot \delta\vec{b}^+] + r \cdot \vec{\xi}_n)) \tag{18}$$

Отримана оцінка вектора параметрів є наближенням для наступної *k+1* ітерації пошукової процедури.. Додатково у цій процедурі перевизначаємо вектор пам'яті:

$$\vec{w}_{k+1} = \alpha \cdot \vec{w}_k + \beta \cdot \frac{\vec{B}_{k+1} - \vec{B}_k}{r}, \tag{19}$$

де α - ($0 \le \alpha \le 1$) – коефіцієнт забування, а β – коефіцієнт інтенсивності врахування нової інформації.

Пошук продовжується до тих пір, поки зменшується значення функції мети. Якщо ж значення функції мети не зменшується на певній ітерації, то замість конуса використовуємо гіперсферу, як на початковій ітерації для заданого вектора оцінок параметрів. Якщо ж і далі серед згенерованих точок не можливо знайти точку, яка забезпечує зменшення функції мети, то у цьому випадку налаштовуємо довжину кроку *r*, як правило зменшуємо її.

Процедура випадкового пошуку із адаптацією розподілу випадкового кроку (ВПізАРВК). Основною характеристикою цього розподілу є математичне сподівання $M(\vec{\xi})$. Суть випадкового пошуку із адаптацією розподілу випадкового кроку полягає у зміні $M(\vec{\xi})$ таким чином, щоб якомога більше враховувати «історію» пошуку, а саме «успішні» кроки. Отже математичне сподівання випадкового напрямку $\vec{\xi}$ змінюватимемо із урахуванням вектора пам’яті $\vec{w}_{K+1}$. Тоді вектор, що задає випадковий напрямок у задачі пошуку матиме такий вигляд:

$$\vec{\xi}_{k+1} = \vec{\xi}_0 + \vec{w}_{k+1}, \tag{20}$$

де $\vec{\xi}_0$ - деякий «нульовий вектор», що задається математичним сподіванням початкового розподілу.

Вектор пам’яті $\vec{w}_{k+1}$ визначає зміщення «випадкового руху» в процесі пошуку. Передбачається, що це зміщення має бути направлене в сторону, де функція мети набуває нульового значення. У цьому випадку вектор пам’яті відображає історію пошуку і корегує перспективний напрям руху до мети. При цьому напрямок $\vec{w}_{K+1}$ на кожній ітерації корегується.

Очевидно, що у такому підході закладено певну евристику: напрям потрібно формувати як зважену суму випадкових кроків, причому «успішні» кроки потрібно брати із позитивними вагами, а «неуспішні» – з від’ємними. При цьому перевага повинна надаватися даним із «короткою історією».

II. Дослідження обчислювальної складності методу синтезу фільтра та допусків на параметри його елементів при заданих допустимих значеннях модуля коефіцієнта передачі

Розглянемо обчислювальну складність вище описаних методів випадкового пошуку на прикладі аналогового фільтра представленого на рисунку 1. Задамо необхідні номінальні значення модуля коефіцієнта передачі $|K(f_i)| = K_{0i}, i = 1,..., N$ на різних частотах і допустимі відхилення від цих значень в межах 20%, тобто $\delta K_i = 0{,}2$. (табл.1.)

Допустимі значення для цієї характеристики на різних частотах представлено у таблиці 1 та на рис. 3.

Таблиця 1 – Обмеження на вихідні характеристики РЕК на різних частотах

i	$f_i, Гц$	K_{0i}	$K_i^- = K_{0i} - K_{0i} \cdot \delta K_i$	$K_i^+ = K_{0i} + K_{0i} \cdot \delta K_i$
1	70700	0,295810306	0,236648245	0,354972
2	70750	0,324681261	0,259745009	0,389618
3	70800	0,359323772	0,287459018	0,431189
4	70850	0,401452582	0,321162066	0,481743
5	70900	0,453384085	0,362707268	0,544061
6	70950	0,518154327	0,414523461	0,621785
7	71000	0,599350631	0,479480504	0,719221
8	71050	0,699805336	0,559844269	0,839766
9	71100	0,816863688	0,65349095	0,980236
10	71150	0,931096493	0,744877195	1,117316
11	71200	0,997002478	0,797601983	1,196403
12	71250	0,973189315	0,778551452	1,167827
13	71300	0,87619141	0,700953128	1,05143
14	71350	0,756847583	0,605478066	0,908217
15	71400	0,647525248	0,518020198	0,77703
16	71450	0,557289514	0,445831611	0,668747
17	71500	0,48507222	0,388057776	0,582087
18	71550	0,427376843	0,341901475	0,512852
19	71600	0,380851292	0,304681033	0,457022
20	71650	0,342842745	0,274274196	0,411411
21	71700	0,311367179	0,249093743	0,373641

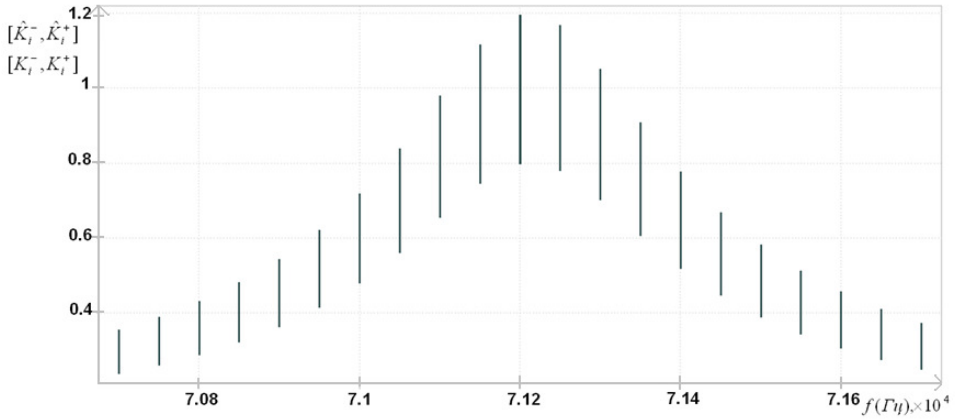


Рисунок 3 – Допустимі (інтервальні) значення для модуля коефіцієнта передачі аналогового фільтра

Програмна реалізація проводилась на ПК з характеристиками: тип процесора - Mobile DualCore Intel Core i5-2410M, 2655 MHz (27 x 98); системна пам'ять - 8106 MB (DDR3-1333 DDR3 SDRAM).

Форма введення початкових даних для усіх методів наведена на рис.4. Кількість випадковим чином згенерованих, згідно з рівномірним законом розподілу, точок в просторі параметрів на відстані радіусу від початково заданого вектора параметрів рівна 100. Загальна кількість ітерацій 1.000.000, яку можна збільшити, якщо «якість» оцінювання не буде досягнуто.

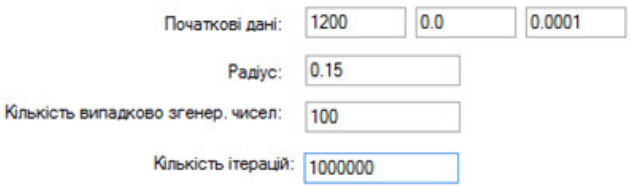


Рисунок 4 – Форма введення початкових даних для методів випадкового пошуку

Рис. 5 ілюструє повідомлення про успішний пошук параметрів елементів фільтра методом випадкового пошуку із лінійною тактикою. Інтервали значень для отриманих параметрів: $[R^-;R^+]=[945{,}25;1044{,}75]Ом$ $[C^-;C^+]=[0{,}46;0{,}5]мкФ$ $[L^-;L^+]=[8{,}5;9{,}5]мкГн$.

Рис. 6 ілюструє повідомлення про успішний пошук параметрів елементів фільтра методом випадкового пошуку по найкращій спробі. Інтервали значень для отриманих параметрів: $[R^-;R^+]=[940{,}5;1039{,}5]Ом$ $[C^-;C^+]=[0{,}48;0{,}53]мкФ$ $[L^-;L^+]=[8{,}7;9{,}7]мкГн$.

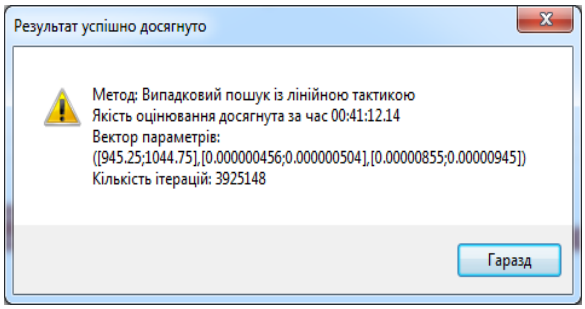


Рисунок 5 – Результат виконання методу випадкового пошуку із лінійною тактикою

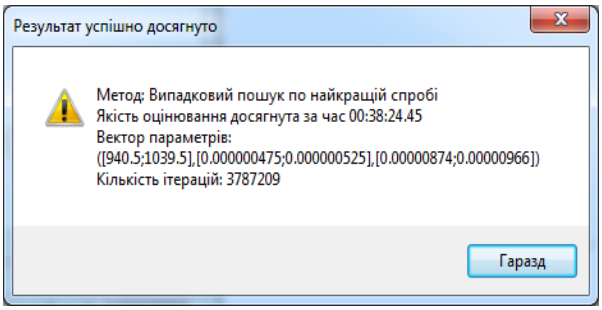


Рисунок 6 – Результат виконання методу випадкового пошуку по найкращій спробі

Рис. 7 ілюструє повідомлення про успішний пошук параметрів елементів фільтра методом випадкового пошуку із використанням направляючого конуса. Інтервали значень для отриманих параметрів: $[R^-;R^+]=[948{,}1;1047{,}9]Ом$ $[C^-;C^+]=[0{,}48;0{,}53]мкФ$ $[L^-;L^+]=[9{,}5;10{,}5]мкГн$.

Рис. 8 ілюструє повідомлення про успішний пошук параметрів елементів фільтра методом випадкового пошуку із адаптацією розподілу випадкового кроку. Інтервали значень для отриманих параметрів: $[R^-;R^+]=[950{,}0;1050{,}0]Ом$ $[C^-;C^+]=[0{,}38;0{,}42]мкФ$ $[L^-;L^+]=[8{,}5;9{,}5]мкГн$.

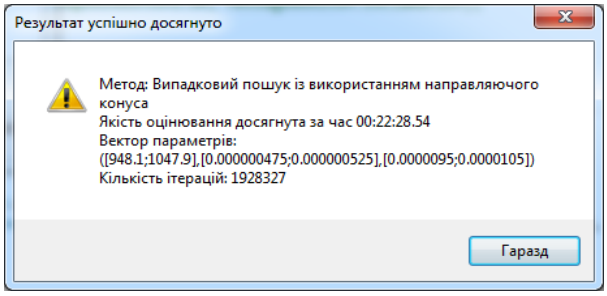


Рисунок 7 – Результат виконання методу випадкового пошуку з використанням направляючого конуса

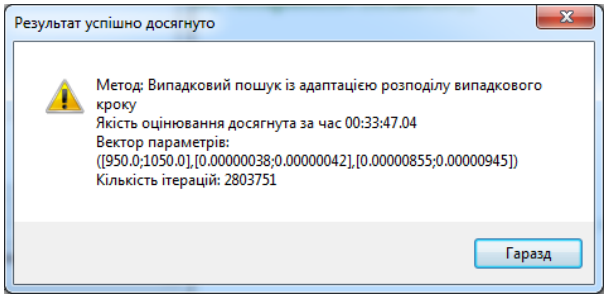


Рисунок 8 – Результат виконання методу випадкового пошуку із адаптацією випадкового кроку

На рис. 9 наведено характеристику модуля коефіцієнта передачі аналогового фільтра з урахуванням заданих допусків на параметри фільтра для кожного з вище представлених методів. Пунктирною лінією наведена характеристика модуля коефіцієнта передачі розрахована методом випадкового пошуку із лінійною тактикою, одинарною лінією – характеристика розрахована методом випадкового пошуку по найкращій спробі, штрих-пунктирною лінією – характеристика розрахована методом випадкового пошуку із використанням направляючого конуса та потовщеною лінією – характеристика розрахована методом випадкового пошуку із адаптацією розподілу випадкового кроку. Вертикальними відрізками наведені інтервали допустимих значень, які визначають функціональну придатність фільтра. Як бачимо з рис. 9, характеристика синтезованого фільтра на заданих частотах знаходиться в допустимих межах.

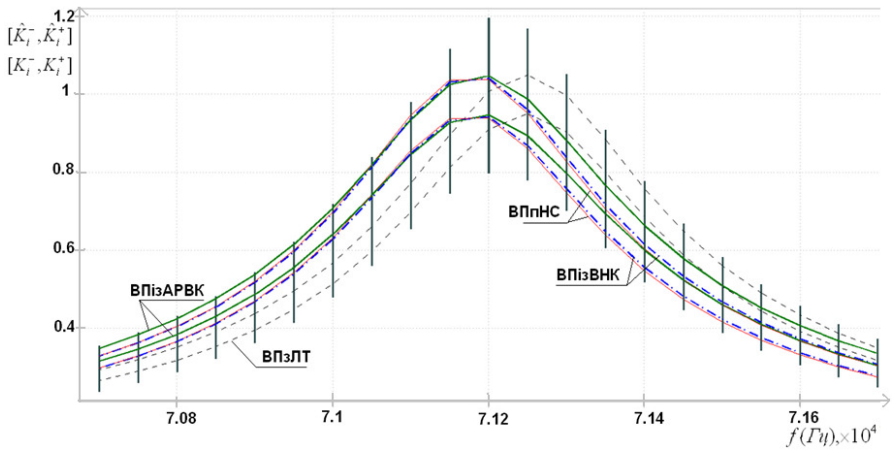


Рисунок 9 – Інтервальні оцінки модуля коефіцієнта передачі аналогового фільтра при застосуванні різних процедур випадкового пошуку

На рис. 10 у вигляді гістограми представлено порівняння часової складності застосування для задачі синтезу аналогового фільтра методів випадкового пошуку. З порівняльної діаграми видно, що найкращим по часовій складності до вказаної задачі є метод випадкового пошуку із використанням направляючого конуса.

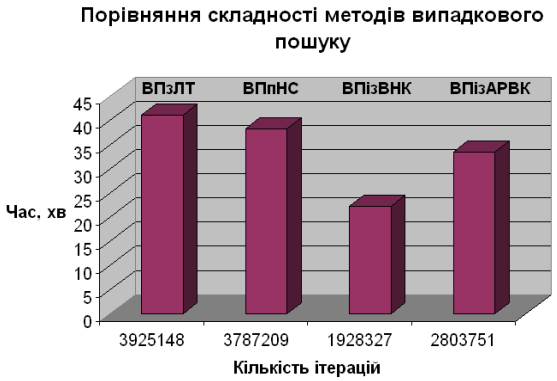


Рисунок 10 – Порівняння часової складності методів випадкового пошуку

III. Висновки

Проведено порівняльний аналіз задачі синтезу радіоелектронних пристроїв з заданими обмеженнями на вихідні характеристики та при заданих допусках на параметри елементів. Порівняльний аналіз проведено на прикладі синтезу параметрів елементів аналогового фільтра.

Показано, що задача синтезу є задачею розв’язку інтервальної системи нелінійних алгебраїчних рівнянь і зводиться до деякої оптимізаційної задачі. Обґрунтовано метод побудови цільової функції задачі синтезу РЕП із заданими обмеженнями на вихідні характеристики та при заданих допусках на параметри елементів.

В процесі дослідження методів синтезу РЕП розглянуто різні процедури випадкового пошуку при реалізації методів синтезу зокрема: метод випадкового пошуку із лінійною тактикою; метод випадкового пошуку по найкращій спробі; метод випадкового пошуку із використанням направляючого конуса; метод випадкового пошуку із адаптацією випадкового кроку.

Встановлено, що найбільш ефективний метод синтезу РЕП, який ґрунтується на процедурі випадкового пошуку із використанням направляючого конуса. Зокрема, часова складність реалізації методу на 91% вища ніж при використанні методу випадкового пошуку із лінійною тактикою, на 65% - методу випадкового пошуку по найкращій спробі та на 47% - метод випадкового пошуку із адаптацією випадкового кроку.

Список використаних джерел

1. Крепич С.Я. Метод синтезу смугового фільтра для заданих обмежень на його модуль коефіцієнта передачі/ С.Я.Крепич/ Сучасні комп’ютерні інформаційні технології: Матеріали IV Всеукраїнської школи-семінару молодих вчених і студентів АСІТ’2014. – Тернопіль: Економічна думка, 2014. – С.26-29

2. Дивак М.П. Вирішення задач синтезу допусків на параметри РЕК методом допускового еліпсоїдного оцінювання з використанням паралельних обчислень/ М.П.Дивак, С.Я.Максимова/ Сучасні комп’ютерні інформаційні технології: Матеріали Всеукраїнської школи-семінару молодих вчених і студентів АСІТ’2011. – Тернопіль: Економічна думка, 2011. – С.97-101

3. Дивак М.П. Ітераційний метод пошуку допустимого розв’язку ІСЛАР в задачах ідентифікації параметрів динамічних моделей “вхід-вихід” / М.П. Дивак, П.Г. Стахів, І.Я. Каліщук // Відбір та обробка інформації. - 2005. – Випуск 23 (99). – С. 40-48.

4. Yurij Bobalo. Estimation of functional usability of radio electronic circuits by applying method of confidence ellipsoids/ Yurij Bobalo, Petro Stakhiv, Svitlana Krepych/ Computational Problems of Electrical Engineering, Lviv Polytechnic National University, Volume 2,No.2, 2012. – P.1-7

5.Кривошейкин А.В. Точность параметров и настройка аналоговых радиоэлектронных цепей/ А.В. Кривошейкин. – М.:Радио и связь, 1983. -136с.

6. Дивак М.П. Ідентифікація параметрів різницевого оператора в задачах моделювання процесів поширення забруднень методами аналізу інтервальних даних / М.П.Дивак, А.В.Пукас, Т.М.Дивак// Зб.Наук.Праць ДонНТУ. Серія інформатика, кібернетика та обчислювальна техніка. – 2009. – Вип.10(153) – С.224-229

7. Дивак М.П. Кількісні характеристики оцінки якості структури моделі у вигляді інтервального різницевого оператора / М.П. Дивак, Т.М. Дивак, І. Ф. Войтюк // Відбір і обробка інформації. Міжвідомчий збірник наукових праць. – Вип. 34 (110). – 2011

8. Стахів П.Г. Синтез радіо-електронних кіл при заданих обмеженнях на вихідні характеристики та за умова заданих допусків на параметри елементів /П.Г.Стахів, М.П.Дивак, С.Я.Крепич// Вимірювальна та обчислювальна техніка в технологічних процесах. Міжнародний науково-технічний журнал, Хмельницьк. – Вип.№3(48), 2014. – С.39-47.

9. Растригин Л.А. Адаптация сложных систем / Л.А. Растригин - Рига: Зинатне, 1981,- 359 с.

10. Алексеева Е.В. Численные методы оптимизации/Е.В.Алексеева, О.А.Кутненко, А.В.Плясунов//Учебное пособие, Новосибирск, 2008. – 126с.

11. Дивак Т.М. Параметрична ідентифікація інтервального різницевого оператора на прикладі макромоделі розподілу вологості у листі гіпсокартону в процесі його сушіння / Т.М. Дивак // Міжнародний науково-технічний журнал “Інформаційні технології та комп`ютерна інженерія”. Вип 3. - 2012.- С.79-85.

Відомості про авторів

Крепич Світлана Ярославівна – аспірант кафедри комп’ютерних наук факультету комп’ютерних інформаційних технологій ТНЕУ.

Дивак Микола Петрович - д.т.н., професор кафедри комп’ютерних наук , декан факультету комп’ютерних інформаційних технологій ТНЕУ.

УДК 004.5; 004.8; 004.89

Л. И. КУРЗАНЦЕВА

Інститут кібернетики ім. В.М. Глушкова НАН України, г.Київ

ОНТОЛОГИЯ – ОСНОВА ДЛЯ ПОСТРОЕНИЯ ИНТЕЛЛЕКТУАЛЬНОГО ЧЕЛОВЕКО-МАШИННОГО ИНТЕРФЕЙСА ДЛЯ ЗНАНИЕОРИЕНТИРОВАННЫХ ОБУЧАЮЩИХ СИСТЕМ

Анотація. У статті розглядається побудова інтелектуального людино-машинного інтерфейсу для знанняорієнтованих навчальних систем, оскільки існуючі інтерфейси користувача не задовольняють вимогам таких систем. Запропоновано вимоги до інтерфейсу, його укрупнена структура та узагальнений алгоритм взаємодії користувачів з системою. Особливістю інтерфейсу є підвищена гнучкість при взаємодії користувача і системи за рахунок використання в його складі онтології процедур функціонування інтерфейсу та засобів для підтримки її роботи. Пропонований інтерфейс надає як викладачеві, так і особі, що навчається, широкий набір функцій адаптації і допомоги при роботі з системою.

Ключові слова: інтелектуальний людино-машинний інтерфейс, знанняорієнтована навчальна система, онтологічний підхід, процедури функціонування інтерфейсу користувача.

Аннотация. В статье рассматривается построение интеллектуального человеко-машинного интерфейса для знаниеориентированых обучающих систем, поскольку существующие пользовательские интерфейсы не удовлетворяют требованиям таких систем. Предложены требования к интерфейсу, его укрупненная структура и обобщенный алгоритм взаимодействия пользователей с системой. Особенностью интерфейса является повышенная гибкость при взаимодействии пользователя и системы за счет использования в его составе онтологии процедур функционирования интерфейса и средств для поддержки ее работы. Предлагаемый интерфейс предоставляет как обучаемому, так и преподавателю широкий набор функций адаптации и помощи при работе с системой.

Ключевые слова: интеллектуальный человеко-машинный интерфейс, знаниеориентированная обучающая система, онтологический подход, процедуры функционирования пользовательского интерфейса.

Abstract. This article discusses the construction of intelligent man-machine interface for knowledge oriented training systems.The existing user interfaces do not meet the requirements of such systems. Requirements for the interface, it enlarged the structure and the generalized algorithm of user interaction with the system are proposed. A feature of the interface is increased flexibility in the interaction of the user and the system by using the ontology in its composition and modalities of the interface means to support its operation. The proposed interface provides both the student and the teacher to adapt a wide range of functions and support when working wih the system.

Keywords: intelligent man-machine interface, a knowledge oriented training system, an ontological approach, procedures for the functioning of the user interface.

Введение.

Современная парадигма непрерывного образования, являющаяся необходимым условием для развития экономики и социальной сферы общества, предполагает развитие у обучаемого способности самостоятельно создавать новые знания.

Широкое внедрение информационных технологий в сферу образования позволяет организовать учебный процесс с применением автоматизированных средств обучения, активизирующих познавательную деятельность обучаемых путем использования современных методов и технологий адаптации обучаемого и системы, что повышает эффективность обучения как группового с преподавателем, так и индивидуального, а также самообучения. При этом оказывается значительная помощь преподавателю.

Одним из значимых направлений создания информационных систем и пользовательских интерфейсов является онтологический подход, который, несмотря на сложности формирования онтологии предметной области, получил достаточно широкое применение.

Онтология – технология, осуществляющая эффективную обработку информации на основе знаний. Она позволяет классифицировать предметную область, гибко перестраивать структуру и взаимосвязи. Также обладает такими свойствами, как наличие иерархии понятий (объектов) в предметной области, возможностью описания важных свойств каждого объекта, связей между ними, сопоставление одного или нескольких понятий и пр. [1].

Как показал анализ публикаций обучающих систем с применением онтологии, такие системы обеспечивают наглядность представления знаний, объединение материала из различных источников и сред, восстановление недостающих логических связей, а также решение проблемы несовместимости и противоречивости понятий и формируют у обучающегося целостное представление предметной области [2-4]. Использование онтологий при построении тестов для проверки уровня подготовки развивает у обучаемых способности к обобщению и заданию проблемных вопросов, а также способно в достаточной мере охватить все темы курса и правильно оценить обучаемого [5]. Кроме этого, онтология может быть средством коммуникации между обучаемым и преподавателем.

Однако, несмотря на преимущества таких систем, они имеют ряд существенных недостатков, в частности, создание онтологии предметной области в таких системах осуществляется вручную, что приводит к увеличению объема работ преподавателя при реализации онтологии и значительному субъективизму при заполнении баз данных, поэтому повторное использование онтологии достаточно проблематично.

Также имеются и другие недостатки, в частности, трудности при контроле и отслеживании процесса обучения и успеваемости большого количества обучаемых, недостаточно развита обратная

связь между обучаемыми и преподавателем, адаптация учебного материала и оценивание знаний осуществляется в зависимости от уровня знаний обучаемых без учета других характеристик, что снижает эффективность обучения.

Такие недостатки преодолимы в знаниеориентированных информационных системах, представляющих одно из наиболее перспективных направлений Computer Science, которое занимается дальнейшей разработкой прикладных аспектов онтологии, связанных с оптимизацией информационного поиска и повышением эффективности процесса приобретения знаний (формализация, структуризация и классификация материала) [6].

Эти системы предполагают автоматизированное наполнение онтологии с помощью возможностей глобальной сети. В настоящее время они находятся на начальной стадии процесса разработки, поскольку для их создания необходимо решение широкого круга вопросов, касающихся использования баз формализованных знаний, а также приобретения знаний (извлечение, представление, обработка и интеграция), что требует системного подхода к решению вопросов с использованием логики, компьютерной и психологической лингвистики, нейрокибернетики, теории семантических сетей и др. научных дисциплин [6].

Архитектурно-структурные решения знаниеориентированных систем соответствуют стандартам обучающих систем для непрерывного образования, и поэтому могут обеспечить следующие возможности: взаимодействие различных систем, включение развивающихся новых информационных технологий без перепроектирования систем; наличие встроенных методов для обеспечения индивидуализированного обучения и т.д.· [7].

Применение таких систем в сфере образования значительно расширит как возможности, так и области применения обучающих систем с использованием онтологии. Например, формирование различных электронных курсов путем многократного использования онтологии предметной области, поскольку наполнение онтологии автоматизированным способом снижает субъективизм и трудозатраты на разработку и повышает уровень качества. Такие преимущества онтологии, как интеграция разнородных знаний и способность извлечения новых знаний, а также наглядность представления знаний, позволит значительно расширить электронный курс и персонализировать его, настроив на характеристики обучаемого [6]. Также возможна модификация курсов с учетом развития предметной области, поддержка динамических курсов.

Актуальность

Эффективность применения знаниеориентированной обучающей системы (ЗОС) напрямую зависит от возможностей пользовательского интерфейса, обеспечивающего пользователя интеллектуальной поддержкой при взаимодействии с системой и реализующего при этом комплекс адаптивных технологий и методов.

Однако на данный момент, исходя из отечественных и зарубежных публикаций, данное направление также как и сами ЗОС, находится на стадии разработки. Существующие пользовательские интерфейсы обучающих систем не могут быть использованы как интерфейсы для ЗОС, поскольку не учитывают ориентацию системы на технологию реконфигурируемого процессинга, не способны в полной мере обеспечить требуемую адаптивность системы к пользователю на основе опыта пользователя [6]. Кроме этого в их составе отсутствуют средства, позволяющие производить как автоматизированное построение онтологии предметной области, так и автоматизированное наполнение ее информацией, получаемой через Internet и из других источников, а также эффективно использовать эту информацию для получения новых знаний [6].

Вышеприведенное свидетельствует об актуальности разработки предложений по построению интеллектуального интерфейса для знаниеориентированных обучающих систем (ИИ_ЗОС).

Требования к интеллектуальному интерфейсу для ЗОС

Функциональные возможности ИИ_ЗОС и особенности их реализации определяются, в основном, требованиями, предъявляемыми к нему как со стороны знаниеориентированной системы, так и со стороны пользователей системы, которыми, в том числе, являются преподаватели и обучаемые.

Таким образом, требования со стороны ЗОС заключаются в необходимости встраивания в интерфейс дополнительного набора средств, включая редактор базы знаний с инструментами разработки онтологий с возможностью автоматизированного построения онтологии и автоматизированного ее наполнения и средствами предоставления доступа к знаниям, обеспечивающими работу пользователя с онтологией, базами знаний и данных, извлечением знаний, а также различных функциональных возможностей, позволяющих анализировать психофизиологические особенности обучаемого для предоставления ему индивидуальных условий обучения.

Данная система по своему назначению повышает эффективность труда преподавателя, а значит, интерфейс должен удовлетворить требования преподавателей [8], предоставив различные технологии,

помогающие в подготовительной работе, которая состоит в анализе результатов предыдущей и планирования будущей (формирование целей, выбор методики обучения), разработке обучающего материала (лекций, заданий, тестов и т.д.) и пересмотра в свете последних достижений науки и техники с учетом подготовленности и индивидуальных особенностей обучаемых. Также ИИ_ЗОС осуществляет контроль действий обучаемых, оценку результатов тестирования и коррекцию обучающего процесса.

Требования со стороны обучаемых заключаются в предоставлении учебного материала, тестовых и практических заданий, помощи и подсказок в форме и темпе подачи материала, тестов, которые определяются в соответствии с учетом особенностей обучаемого. Материал должен подаваться, начиная с начального уровня знакомства с базовыми понятиями дисциплины до сложного, на котором развивается изобретательность и интуиция обучаемого. При этом должны быть определены особенности мышления обучаемого и наличие у него пробелов в знаниях, в соответствии с которыми проведена коррекция процесса обучения. Также должны поддерживаться различные разновидности форм обучения (лекция, практические, лабораторные, контрольные занятия, семинары, консультации, самостоятельное обучение), связь с техническими средствами обучения, в том числе и с использованием телекоммуникационных средств.

Таким образом, исходя из вышеперечисленного, ИИ_ЗОС, задачей которого является интеллектуальная поддержка взаимодействия пользователя и системы, должен обеспечить выполнение следующих требований:

- динамический контроль процессов умственной (интеллектуальной), психологической и физиологической нагрузки;
- возможность прогнозирования поведения обучаемого;
- определение адаптации к различным видам деятельности;
- определение профессионального соответствия к определенным видам деятельности;
- создание психологического портрета обучаемого;
- оценку эффективности выбранной стратегии обучения для обучаемого;
- работу пользователей разного образовательного уровня, разных физических возможностей (в т.ч. инвалидов), разных культур;
- формирование и корректировку курса для каждого обучаемого в соответствии с целями и задачами обучения и индивидуальными психофизиологическими возможностями обучаемого;
- хранение информации об обучаемом на протяжении всего периода обучения для обеспечения возможности адаптации учебных материалов к конкретным его потребностям или возможностям.

Однако реализация требований в полной мере зависит от решения ряда задач, поставленных перед педагогической, социальной, когнитивной отраслями психологии и психолингвистикой.

К ним относятся [9]:

- определение психологических закономерностей между усвоением знаний и умений обучаемого и его психофизиологическими особенностями при работе с использованием обучающей системы;
- поиск психологическо-педагогических параметров обучаемых, определяющих их способности к обучению в целом, а так же к определенной дисциплине;
- создание методик, методов и правил определения количественных оценок этих параметров для регистрации изменения результатов обучения;
- определение критериев, характеризующих эффективность стратегии обучения;
- определение эффективности методов обучения для обучаемых с различными индивидуальными особенностями;
- создание методов и алгоритмов управления обучением.

Особенности построения интеллектуального пользовательского интерфейса для ЗОС

Реализация данных требований, предполагает создание для каждого пользователя ЗОС пользовательского интерфейса со своим набором функций, что обеспечивается созданием в составе интерфейса модели пользователя, отражающей его индивидуальные характеристики.

Поскольку при разработке ИИ_ЗОС, исходя из специфики применения системы, наибольшее внимание уделяется характеристикам обучаемых, которые являются самыми многочисленными пользователями системы, рассмотрим модель обучаемого. Она отображает уровень знаний обучаемого, его когнитивные процессы (поиск, обработка информации и решение задач), психологические характеристики (уровень развития, интересы, стиль и темп обучающей деятельности).

Исходя из задач обучения, будь-то оценка уровня знаний по конкретному предмету, определение шага, на котором была допущена ошибка, или прослеживание всевозможных путей решения задачи, создаются различные типы моделей обучаемых: стереотипные, векторные, сетевые и пр. [10]. Однако, ни одна из них в явном виде не подходит для отображения совокупности получаемых знаний, умений и навыков обучаемого в их динамической взаимосвязи.

Для этого необходимо в составе интерфейса предусмотреть набор методов и средств адаптации возможностей системы к изменяемым в процессе взаимодействия особенностям пользователя, позволяющих кардинально изменить внутренние структуры модели пользователя в процессе функционирования, что в обычных интерфейсах не предусмотрено. Это достигается встраиванием в управляющей структуре интерфейса механизмов, обеспечивающие как модификацию структуры модели пользователя, так и настройку и перенастройку функций интерфейса под разные категории пользователей с учетом их изменяющихся характеристик и типов решаемых задач в процессе работы. Подходом, способным реализовать такое решение при построению ИИ_ЗОС, может служить онтологический подход.

Применение данного подхода при создании пользовательских интерфейсов, исходя из анализа публикаций на эту тему [11-15], показало высокую эффективность его использования. Она достигается как за счет автоматизации разработки и сопровождения пользовательского интерфейса, что снижает трудоемкость создания интерфейса и сокращает время на его разработку, так и за счет расширения функциональных возможностей интерфейса, в т.ч. предоставления автоматического поиска необходимой пользователю информации; управления взаимодействием пользователей как с информационной системой, так и друг с другом на основе опыта, интересов; работы со знаниями с целью контроля за реализацией управляющего воздействия и выявления критической ситуации, а также прогнозирования решений; обеспечения высокого уровня "дружественности" взаимодействия пользователя с системой путем применения в интерфейсе методов и средств поддержки интеллектуальности.

Данные выводы говорят об эффективности применения онтологического подхода и при создании пользовательского интерфейса для ЗОС.

Для реализации вышеуказанных требований при создании ИИ_ЗОС, в его состав включена онтология процедур функционирования интерфейса. При этом под процедурой интерфейса понимается набор операций, реализующих отдельную подпрограмму функции интерфейса, а совокупность процедур реализующих функцию интерфейса называется макропроцедурой. На рис. 1 приведен примерный состав макропроцедур ИИ_ЗОС, а на рис. 2 – фрагмент онтологии.

На фрагменте представлены макропроцедуры из рис. 1 в логической взаимосвязи, отображающие выделенные классы понятий верхнего уровня онтологии и объединенные между собой отношениями.

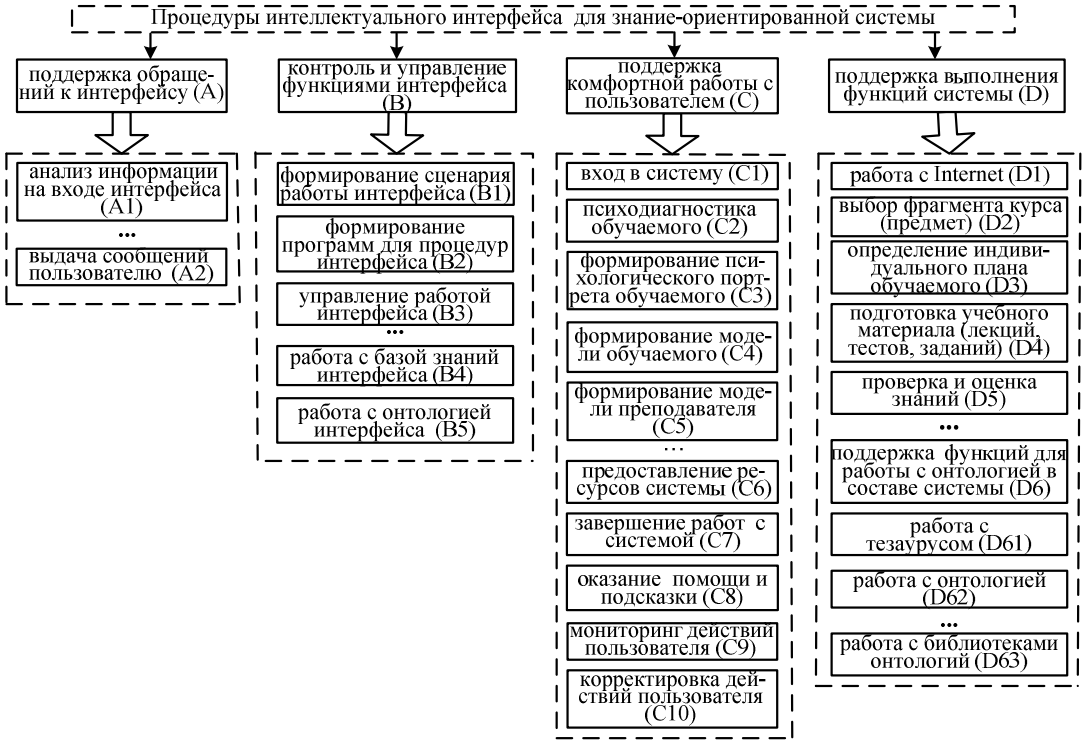
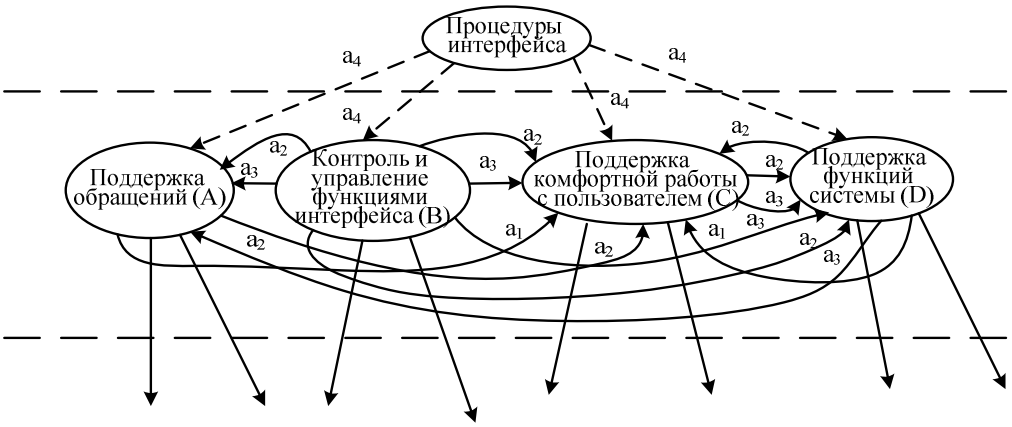


Рисунок 1 – Состав макропроцедур ИИ_ЗОС

Применения процедур для описания работы ИИ_ЗОС дает возможность реализации его функций с учетом как собственной архитектурно-структурной организации, так и требований со стороны системы и пользователя, а применение онтологического подхода позволяет в полной мере учесть все взаимосвязи между процедурами интерфейса. В результате использования онтологии процедур можно создавать

различные варианты пользовательских интерфейсов для ЗОС, а также обеспечить усовершенствование интерфейса, осуществляемое за счет модификации, добавления и удаления компонент, реализующих его функциональные возможности.



Обозначения: a1:= [содержат, включают]; a2:= [используют]; a3:= [определяют]; a4:= [представлены].

Рисунок 2 – Фрагмент онтологии «Процедуры функционирования ИИ_ЗОС»

Структура и функции интеллектуального человеко-машинного интерфейса для ЗОС.

На рис. 3 представлена укрупненная структурная схема ИИ_ЗОС. Все средства поддержки функций ИИ_ ЗОС, реализуемые процедурами интерфейса скомпонованы в четыре блока по назначению. Кратко опишем назначение блоков.

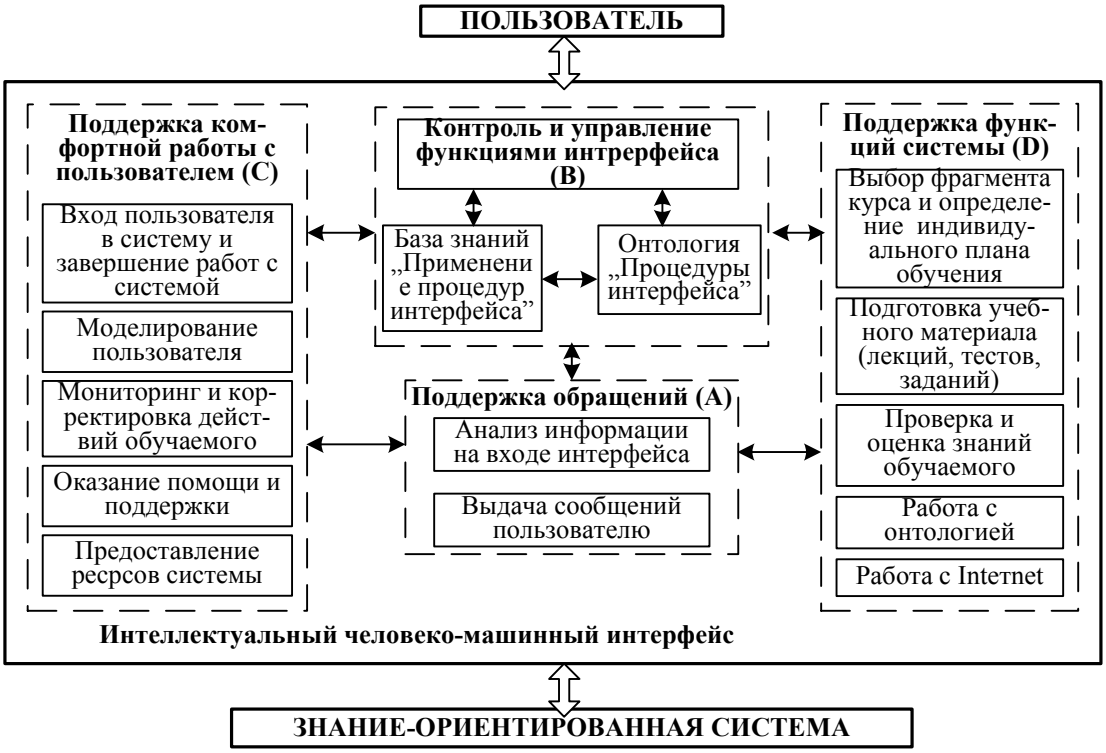


Рисунок 3 – Укрупненная структурная схема ИИ_ЗОС.

Блок «Средства поддержки обращений к интерфейсу» (A) занимается обслуживанием обращений к интерфейсу. В частности, в его функции входит анализ информации, поступающей на вход интерфейса, выдача сообщений пользователю и т.д.

Блок «Средства поддержки контроля и управление функциями интеллектуального человеко-машинного интерфейса» (B) отвечает за организацию работы интерфейса. К его функциям относятся: управление работой самого интерфейса, работа с его базой знаний и онтологией процедур, формирование сценария работы интерфейса и подпрограмм сценария.

Блок «Средства поддержки комфортной работы с пользователем» (C) отвечает за организацию комфортного взаимодействия с пользователем. В его функции входят: оказание поддержки пользователю при входе в систему и завершении работ с системой; формирование модели пользователя и ее модификация (предусматривается создание как модели обучаемого, так и модели преподавателя); предоставление необходимых ресурсов системы пользователю и оказание помощи в соответствии с моделью пользователя; наблюдение за действиями обучаемого, корректировка неверных действий и пр.

Модель обучаемого представляет собой набор характеристик, измеряемых во время работы с системой. В общем случае, она включает персональные данные, степень усвоения знаний изучаемой дисциплины и информационной компетентности, а также психофизиологическое состояния (в т.ч. работоспособность, тревожность). Ее основные функции: оценка уровня обучаемого, диагностика ошибок обучаемого, выявление недостающих и ошибочных знаний, приведших к возникновению ошибок, отображнение индивидуальных особенностей обучаемого, влияющих на учебный процесс [19].

Блок «Средства поддержки функций системы» (D) обеспечивает поддержку выполнения учебных задач, в том числе: выбор преподавателем педагогической стратегии, формирование индивидуальных планов обучения, подготовку учебного материала, проверку и оценку знаний обучаемого, а также поддерживает работу с онтологией предметной области в составе системы и с Internet.

Под работой с онтологией понимается создание, редактирование, отображение, интеграция онтологий, таксономий и отношений и пр. При этом к функциям, поддерживаемых блоком, относятся также работа с тезаурусом и с библиотеками онтологий.

База знаний «Применение процедур интерфейса» содержит набор положений, критериев, данных, предназначенных для построения правил вывода процедур работы интерфейса.

Онтология «Процедуры интерфейса» описывает реализацию функций интеллектуального человеко-машинного интерфейса знаниеориентированной системы с помощью процедур интерфейса с учетом совокупности процессов и взаимосвязей, происходящих при взаимодействии пользователя с системой в процессе ее функционирования.

Обобщенный алгоритм взаимодействия пользователей с знаниеориентированной системой

Знаниеориентированная система имеет в наличие онтологию предметной области конкретной дисциплины (или несколько онтологий различных дисциплин), которая представляет собой в общем виде структурированный по темам учебный материал, взаимосвязанный между собой и представленный в виде лекций, практических задач и тестов.

На рис. 4 изображен обобщенный алгоритм взаимодействия пользователей (преподавателей и обучаемых) с ЗОС с помощью предлагаемого интерфейса. Кратко опишем алгоритм.

Подготовка преподавателя к учебному занятию происходит на основании составленного учебного плана, программой дисциплины и методическими указаниями по проведению занятия. Он осуществляет выбор фрагмента курса (предмет), определяет обучающую задачу и формирует план занятий фрагмента курса. ИИ_ЗОС оказывает ему в этом помощь, предоставляя соответствующие сервисы, в том числе и для работы с онтологией.

Преподаватель с помощью онтологии предметной области курса может выбрать лекцию, тесты или практические задачи для будущего занятия или на основании имеющегося в онтологии материала составить свой вариант занятия в соответствии с уровнем подготовки и запросами обучаемых. Кроме этого, он может корректировать учебный материал по темам занятий курса исходя их новых достижений науки и техники. Также ИИ_ЗОС предоставляет возможность редактирования тем будущих занятий в онтологии предметной области и с помощью автоматизированного наполнения создать новый курс.

Для успешного проведения учебного занятия преподавателю необходимо познакомиться с индивидуальными особенностями будущих обучаемых. Для этого ИИ_ЗОС предоставляет ему следующие возможности: информацию об индивидуальных характеристиках обучаемых, влияющих на обучение по данному предмету; составление тестов для определения этих характеристик; составление психологического портрета обучаемого; моделирование результатов будущего обучения, возможных ошибок и их причин; составление рекомендаций по методу преподавания (например, интеллектуальные способности определяют темп и уровень сложности предлагаемого материала, а когнитивный стиль определяет форму предоставления материала).

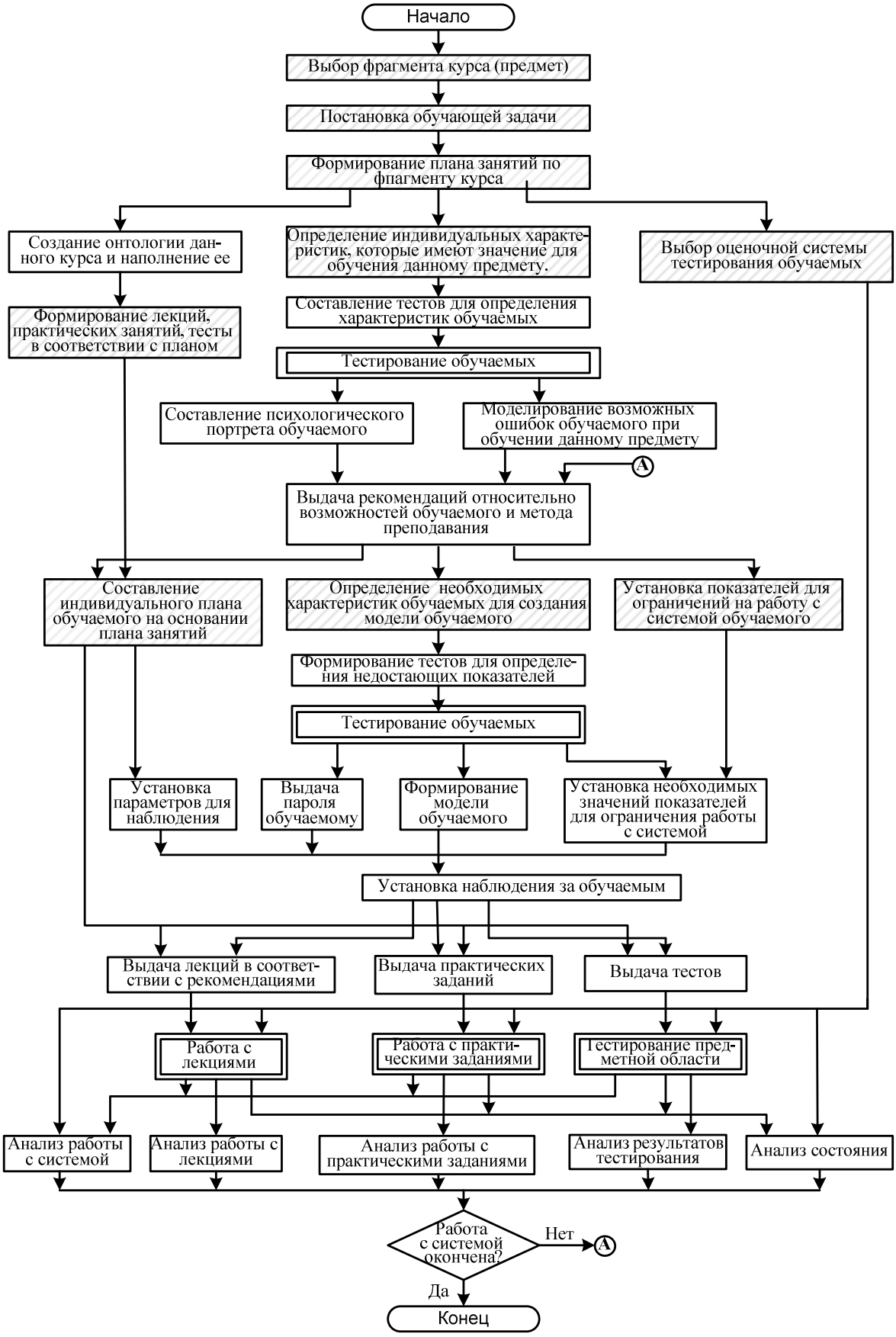


Рисунок 4 – Обобщенный алгоритм взаимодействия пользователей с ЗОС

Психологический портрет обучаемого создается с использованием методов психодиагностики. При этом тестируются психолого-педагогические характеристики обучаемого, влияющие на его способности к обучению.

По полученным результатам определяется психологический портрет обучаемого, в соответствии с которым для достижения цели обучения преподавателем осуществляется выбор индивидуального плана изучения дисциплины.

Преподаватель подготавливает учебные материалы (текст, изображения, мультимедиа и пр.) для курса в соответствии с целями обучения (лекции, тесты, практические работы и т.д.) с учетом адаптации к способу восприятия и уровню знаний конкретного обучаемого или группы обучаемых. Для проверка и оценки знаний обучаемого преподавателю предлагаются различные методики оценки.

Также он задает необходимые для реализации этого плана параметры модели обучаемого, устанавливает параметры для наблюдения за обучаемым и необходимые значения показателей на ограничение его работы с ЗОС. ИИ_ЗОС при создании модели обучаемого определяет наличие всех его характеристик, и при отсутствии показателей составляет тесты для их нахождения, проводит анализ полученных данных. Далее формируется модель, в соответствии с которой будет осуществляться процесс обучения. Кроме этого, ИИ_ЗОС выдает пароль обучаемому для входа в систему и устанавливает наблюдение за его действиями в течение всего сеанса работы его с системой.

Обучаемый с помощью системы изучает лекционный материал, выполняет тестирование и практические задания, как для перехода к другой теме курса, так и для самоконтроля.

На каждом этапе обучаемому в соответствии с его особенностями предоставляются подсказка и помощь, объясняются причины ошибок, указывается материал, который необходимо самостоятельно изучить.

Наблюдение за действиями обучаемого предполагает фиксацию каждого действия и его времени, а также полученные при этом правильные или неправильные результаты. Далее проводится анализ его результатов работы как по теме курса для определения его уровня знаний, ошибок, умений, так и с системой для определения его навыков работы с информационными технологиями. Результатом проведенной работы будет как корректировка модели пользователя, так и выдача рекомендаций преподавателю по повышению успешности данного обучаемого.

Корректировка модели обучаемого в ходе работы с системой осуществляется на основе анализа истории работы обучаемого с системой. При анализе возможен учет следующих показателей: время нахождения учебного материала на мониторе перед обучаемым, время, затраченное на изучение учебного материала, время первого ответа, количество сделанных в процессе работы ошибок и пр. При этом также определяется утомляемость обучаемого и выдаются соответствующие рекомендации, например, для снятия напряжения и повышения концентрации внимания делаются перерывы и предлагается набор упражнений для восстановления.

ИИ_ЗОС предоставляет преподавателю различные возможности контроля за процессом обучения: просмотр результатов тестирования обучаемых; проведение анализа выполненных ими практических работ и тестирования; внесение определенных корректировок в модель пользователя, управление настройками; выдача рекомендаций по повышению успешности обучаемого, непосредственное руководство.

Выводы

Анализ различных решений в области создания как информационных обучающих систем, так и человеко-машинных интерфейсов с использованием онтологического подхода показал высокую эффективность его применения. В частности, обучающие системы обеспечивают наглядность представления знаний, формируют целостное представление предметной области, развивают способности к обобщению, повышают объективность оценки знаний. Применение этого подхода при построении интерфейса расширяет функциональность за счет предоставления пользователю новых возможностей, сокращает трудоемкость разработки.

Однако, ИИ_ЗОС должен обладать качественно новыми признаками по сравнению с другими категориями интерфейсов, а весь набор реализуемых им функций, должен быть значительно расширен, что объясняется назначением и возможностями ЗОС, требующей интеллектуальную поддержку работы пользователя с системой. Полная реализация требований возможна при решении ряда задач, поставленных перед педагогической, социальной, когнитивной отраслями психологии и психолингвистикой.

Применение данного интерфейса для знаниеориентированных обучающих систем позволит в процессе работы с системой изменять параметры предметной области, в рамках которой обучаемый решает задачу, при изменении характеристик обучаемого в процессе работы с системой или по требованию преподавателя динамически изменять как состав, так и структуру модели обучаемого; формировать и корректировать индивидуальный курс для каждого обучаемого в соответствии с целями и

задачами обучения, управлять процессами интеллектуальной, психологической и физиологической нагрузки и т.д., что не только улучшит качество взаимодействия «обучающийся – преподаватель – система», но и выведет его на совершенно новый уровень обучения.

Список литературы

1. Гаврилова Т.А. Использование онтологий в системах управления знаниями / Гаврилова Т.А. – Режим доступа: http://kmtec.ru/publications/library/authors /use_ontology_ in_suz.shtml.

2. Лаптев В.В. Модель предметной области и оценка ее сложности в обучающей системе по программированию / Лаптев В.В. – Режим доступа: http://cyberleninka.ru/article/n/model-predmetnoy-oblasti-i- otsenka ee- slozhnosti-v-obuchayuschey-sisteme-po-programmirovaniyu.

3. Балакирев В.С., Маркелов А.Ю. Создание и использование онтологии предметной области «надежность систем автоматизации» / Балакирев В.С., Маркелов А.Ю. Режим доступа: – http://cyberleninka.ru/article/ n/ sozdanie- i -ispolzovanie -ontologii- predmetnoy -oblasti -nadezhnost-sistem-avtomatizatsii.

4. Слепцова Ю., Малиновская О. Интегрированный комплекс по моделированию пользователя и адаптации СДО/ Слепцова Ю., Малиновская О. – Режим доступа: http://foibg.com /ibs_isc /ibs-06/IBS-06-p18.pdf.

5. Малиновский В.П. Применение онтологий при построении тестов для проверки уровня подготовки обучаемых. / Малиновский В.П. – Режим доступа: http:// www.raai.org/about/ persons/ malinovski/ mfi.doc.

6. Палагин А.В., Кривой С.Л., Петренко Н.Г. Онтологические методы и средства обработки предметных знаний: монография / Палагин А.В., Кривой С.Л., Петренко Н.Г. – Луганск: ВНУ им.В.Даля, 2012. – 323с.

7. Косова А.М., Косова Г.Ф. Анализ стандартизации обучающих систем на основе информационных технологий / Косова А.М., Косова Г.Ф. – Режим доступа: http://dspace.nbuv.gov.ua/handle/123456789/52141.

8. Коростелёва О.Н., Коростелёв А.И. Процесс подготовки преподавателя высшей школы к занятиям / Коростелёва О.Н., Коростелёв А.И. – Режим доступа: http://cyberleninka.ru/article/n/protsess-podgotovki-prepodavatelya-vysshey-shkoly-k-zanyatiyam.

9. Блинова Н. Г., Лурье С. Б., Васина Е. В. Психофизиологическое развитие учащихся в условиях профильного обучения / Блинова Н. Г., Лурье С. Б., Васина Е. В. – Режим доступа: http://cyberleninka.ru/article/n/psihofiziologicheskoe-razvitie-uchaschihsya-v-usloviyah-profilnogo-obucheniya.

10. Валькман Ю.Р., Скляр А.В. Сравнительный анализ моделей знаний обучаемого в интеллектуальных обучающих системах. / Валькман Ю.Р., Скляр А.В. – Режим доступа: http://dspace.nbuv.gov.ua/bitstream/ handle/123456789/ 28561/12-Valkman.pdf?sequence=1

11. Грибова В.В. Методы автоматизированного проектирования и сопровождения пользовательских интерфейсов / Грибова В.В. – Режим доступа: http:// oibg.com/ ibs_isc/ibs-04/IBS- 04-p19.pdf.

12. Соболева Н.В. Подходы к проектированию когнитивного интерфейса к базам данных / Соболева Н.В. – Режим доступа: http://digital-mag.tti.sfedu.ru/8.htm.

13. Ломов П.А., Путилов В.А., Маслобоев А.В. Поддержка интеллектуальности пользовательского интерфейса системы распределенного семантического поиска: проблемы и решения / Ломов П.А., Путилов В.А., Маслобоев А.В. – Режим доступа: http://vestnik.mstu.edu.ru /v13_3_n40/articles/09_maslo.pdf.

14. Загорулько Ю.А., Боровикова О.И., Загорулько Г.Б. и др. Разработка портала знаний по компьютерной лингвистике / Загорулько Ю.А., Боровикова О.И., Загорулько Г.Б. и др. – Режим доступа: http://www.raai.org/conference/cai-08/files/cai-08_exhibition_24.doc.

15. Попова М.А. Модель онтологического интерфейса агрегации информационных ресурсов и средств ГИС / Попова М.А. – Режим доступа: http://www.foibg.com/ijitk/ijitk-vol07/ijitk07-04-p06.pdf.

Сведения об авторах

Курзанцева Лариса Игоревна – к.т.н., с.н.с., тел. 526-32-07, Институт кибернетики им. В.М. Глушкова НАН Украины, 03680, ГСП, Київ-187, пр. Академика Глушкова, 40.

УДК 621.339

О. М. РОЇК, Ю. В. МІРОНОВА, О. П. ВОЛКОТРУБ

Вінницький національний технічний університет, Вінниця

ОСОБЛИВОСТІ ЗАХИСТУ ВІДЕОМАТЕРІАЛІВ НА WEB – РЕСУРСІ

Анотація. В статті розглянуто питання розробки захисту відеоматеріалів на web- ресурсі. Було розглянуто та проаналізовано актуальність захисту відеоматеріалів, проведено порівняння з існуючими аналогами, проведено підбір оптимальних технологій та засобів реалізації. Також було приведено програмно-апаратні вимоги до комп`ютера користувача та сервера, санітарно-гігієнічні норми для робочого місця адміністратора сервера.
В статті було розроблено програмний модуль захисту відеоматеріалів на web-ресурсі, використовуючи технології компанії Microsoft. Це дало змогу реалізувати функцію програмного захисту потокового відео.
Ключові слова: потокове відео, захист відеоматеріалів, технологія захисту.
Аннотация. В статье рассматривается вопрос о разработке защиты видеоматериалов на web - ресурсе. Было рассмотрено и проанализировано актуальность защиты видеоматериалов, проведено сравнение с существующими аналогами, проведен подбор оптимальных технологий и средств реализации. Также было приведено программно-аппаратные требования к компьютеру пользователя и серверу, санитарно-гигиенические нормы для рабочего места администратора сервера.
В статье был разработан программный модуль защиты видеоматериалов на web-ресурсе, используя технологии компании Microsoft. Это позволило реализовать функцию программной защиты потокового видео.
Ключевые слова: потоковое видео, защита видеоматериалов, технология защиты.
Abstract. The article discusses the development of the protection of video on web - resource. It was reviewed and analyzed the relevance of protection of video, a comparison with existing analogues, conducted selection of optimal technologies and implementation. The requirements to the user's and server's soft- and hardware, the sanitary and hygienic norms for the server administrator's working space are given.
The paper was developed software for video security module web-website using technology of Microsoft. This made it possible to implement software protection feature streaming.
Keywords: streaming video, protection video, protection technology.

Вступ

Передача даних через мережу Інтернет справедливо вважається ненадійною і саме тому існує необхідність аналізувати, розробляти та створювати методи і технології захисту відеоматеріалів. Хоча даною проблемою цікавляться вже доволі довгий час, адміністратори web-ресурсів зазвичай займаються лише поверхневим захистом відео, в межах свого web-сайту, наприклад, обмежуючи доступ до контенту засобами мови розмітки HTML.
Максимально всебічно захистити відеоматеріали, не сильно навантажуючи обчислювальні потужності ЕОМ можливо завдяки новітнім технологіям, а саме – технології потокового мультимедіа. Завдяки цій технології користувач безперервно отримує потік інформації від сервера потокового мовлення.
В основному, мультимедіа інформація займає великі обсяги, так що витрати на зберігання та передачу такої інформації завжди великі, тому, в більшості випадків, передана в потік інформація стискається при передачі в мережу мовлення. Мультимедіа потоки бувають двох видів: за запитом або живими. Потоки інформації, що викликається за запитом користувача, зберігаються на серверах тривалий період часу. Живі потоки доступні короткий період часу, наприклад, при передачі відео зі спортивних змагань.

Отже, потокове відео – нова технологія, яка ще розвивається, але, тим не менш, стрімко завойовує позиції на ринку мультимедіа і має безліч аспектів, що потребують покращення та вдосконалення. [1]

Актуальність

За сучасних умов стрімкий розвиток інформаційних технологій індустрія on-line навчання та розваг досягла високого рівня. Все більше і більше компаній різноманітних напрямків переходять на цифровий формат передачі даних. При цьому слід відмітити, що одним з основних типів контенту є відеоматеріали та відео презентації, так як це один з найпростіших та найбільш ефективних способів надати користувачу інформацію про себе, про свої товари та послуги. В зв`язку з підвищеним попитом на технології захисту відео контенту, постає необхідність створення нових засобів захисту відео інформації, яка могла б використовуватись у web-додатках. Тому тема статті, що присвячена розв’язанню подібних задач, є актуальною.

Мета досліджень

Метою дослідження є розв’язання задачі захисту відеоматеріалів на web-ресурсі за допомогою технології потокового мультимедіа.

Постановка задачі

Для вирішення проблеми захисту відеоматеріалів на web-ресурсі необхідно розв’язати та дослідити такі задачі: проаналізувати існуючі загрози, проаналізувати існуючі методи захисту відео, описати метод захисту відеоматеріалу на web-ресурсі.

Однією із задач, які потрібно при цьому розв’язати, є задача розробки методу захисту відеоматеріалу на web-ресурсі. Аналіз показав, що найбільш зручною технологією для написання відео програвача було обрано технологію Microsoft Silverlight. Вона потребує використання мови ASP.NET для написання того ресурсу, на якому буде розміщено сервіс, а отже потрібен хостинг, який має інтерпретатор ASP.NET. Саме

таким хостингом є сервери під керуванням IIS6 або вище. В таблиці 1.1 наведені системні та програмні вимоги для сервера.[2]

Таблиця 1.1 – Системні та програмні вимоги для сервера

Програми	Вимоги для сервера
ЦП	Intel Celeron Dual-Core G530 2.4GHz/2MB/5GT/s чи подібний
ОЗУ	Не менше 4 Гб
Відеокарта	Будь-яка з підтримкою DirectX9.0 і роздільною здатністю не менше 1280x1024
ОС	Windows Vista\7,Windows Server 2008\2008 R2, MacOS
Браузер	IE9.0, чи подібний
Швидкість Інтернет-з`єднання	4Гбіт\с

Окрім комп`ютерів на базі MacOS, технологія Silverlight працює і з різноманітними портативними пристроями компанії Apple, а саме з iPhone, iPod touch: iOS 3.0, iOS 3.0.1, iOS 3.1, iOS 3.1.2, iOS 3.1.3, iOS 4.0, iOS 4.1, iOS 5.0; та iPad: iOS 3.2, iOS 4.2.1, iOS 5.0.
Враховуючи, що Silverlight може не коректно працювати з операційними системами Linux, до комп`ютера користувача також висувається ряд вимог, наведених в таблиці 1.2[3].

Таблиця 1.2 – Апаратно-програмні вимоги для користувача

Програми	Вимоги для користувача
ОЗУ	Не менше 1 Гб
Відеокарта	Будь-яка з підтримкою DirectX9.0 і роздільною здатністю не менше 1280x1024
ОС	Windows Vista\7, MacOS
Браузер	IE9.0, чи подібний
Швидкість Інтернет-з`єднання	Не менше 512 Кбіт\с

Таким чином видно, що вимоги до робочих станцій є не високими, адже на сьогоднішній день подібне обладнання не вважається найновішим та найдорожчим, але є достатньо продуктивним.

Розв'язання задачі

Технологія Microsoft Silverlight, яка пропонується в даній статті, базується для написання самого програвача, ASP.NET для розробки сайту-контейнера. Крім цього, було визначено важливість створення даного сервісу через відсутність достатнього захисту. З'ясовано параметри, які повинен мати сервер, для коректної роботи та відображення сайту, і визначенні мінімальні вимоги до ЕОМ користувачів, які виявились доволі низькими і тому не викликатимуть дискомфорту.
Для реалізації потокового відео необхідно перш за все закодувати його для передачі по каналу, створити можливість захищеної передачі закодованого сигналу, та забезпечити користувача засобами для декодування та перегляду. Компанія Microsoft надає необхідні засоби. Повна підтримка такої медіа-платформи містить наступні технології: підготовка – Microsoft Expression Encoder, доставка - IIS Media Services, відтворення - Microsoft Silverlight.
Microsoft Expression Encoder – це програмний додаток для Microsoft Windows для транскодування і нелінійного відео-редагування. Він може створювати відео-потоки для поширення та відтворення за допомогою Microsoft Silverlight .[4]
Expression Encoder дозволяє експортувати відео у форматі H.264 або VC-1. Він може підготувати відеопотоки для поширення через Microsoft Silverlight, підтримує управління плеєром Silverlight і Silverlight шаблони.
Деякі з можливостей Microsoft Expression Encoder включають:

1. Смарт-кодування і перекодування для WMV, скорочення часу редагування, послідовне кодування групи відеофайлів, Live кодування з web-камер та відеокамер;
2. Smooth Streaming (720p+ відео з використанням HTTP) з оптимізованим клієнтом (Silverlight) і сервером (IIS з Smooth Streaming)
3. API, які були використані для створення плагінів для Silverlight Streaming , Amazon S3, і YouTube;
4. Імпорт XAML накладень, фільтрів, створених в Expression Design та налаштування їх часу, анімації, прозорості, розміщення та циклів;
5. Додавання підпису до відео за допомогою SAMI або W3C Timed Text формату;

6. Попередній перегляд і порівняння параметрів кодування в реальному часі та створення зображень з екрану;
7. Об'єктна модель для кодування двигуна, SDK для окремого завантаження [5].

Перш за все, IIS (Internet Information Services, до версії 5.1 – Internet Information Server) – це набір серверів для декількох служб Інтернету від компанії Майкрософт.[6]

Основний компонент IIS – web-сервер, який дозволяє розміщувати в Інтернеті сайти. IIS підтримує протоколи HTTP, HTTPS, FTP, POP3, SMTP,NNTP. IIS другий за популярністю web-сервер за кількістю сайтів, після Apache HTTP Server.

Основною новинкою цього сервера вважається IIS Media Services. Саме завдяки цьому модулю існує можливість реалізовувати потокове відео.

IIS Media Services являє собою набір розширень для Internet Information Services (IIS) 7, які доставляють і керують медіа-контентом. Вона є частиною повної технологічної платформи для розгортання високо доступного, високо масштабованого, медіа середовища доставки.

Стандартні засоби, доступні зі звичним набором функцій та модулів – Bit Rate Throttling, Live Smooth Streaming, Smooth Streaming.

Різноманітні модулі захисту on-line контенту розроблялись на рівні з першою версією Silverlight та були в неї включені. Пізніше ці модулі були оновлені, покращені, були виправлені різноманітні помилки, деякі можливості взагалі були видалені повністю, частково, або ж замінені на більш нові, що раніше не використовувались.[7]

На сьогоднішній день основним заходом із захисту будь-якої інформації і відеоматеріалів зокрема є аутентифікація та авторизація кінцевого користувача. Дійсно, це не дуже ефективний спосіб, так як він зазвичай вимагає постійного втручання спеціально навченої людини, або ж використання дуже розумної програми, яка б могла відслідковувати активність всіх авторизованих чи аутентифікованих користувачів. Тим не менш, для людини постійно перегляди логи активності є надто важкою роботою, а програма схильна підіймати «фальшиву» тривогу.

Однак, даний засіб захисту обмежить коло тих, хто має доступ до медіа контенту.

Наступним методом, реалізованим в Silverlight є шифрування потоку використовуючи SSL(Secure Sockets Layer) та Encrypted RTMP (RTMPE). SSL, як і його попередник TLS(Transport Layer Security), не шифрують медіа контент. Ці протоколи шифрують виключно канал передачі інформації, але вони потребують сертифікат безпеки, що не завжди є достатньо комфортним.

Encrypted RTMP (RTMPE) також шифрує виключно канал передачі даних, але він не вимагає ніяких сертифікатів.[8]

Таким чином цей метод дозволяє зашифрувати канал передачі між сервером і клієнтом, але не шифрує сам файл.

Ще один реалізований метод – Digital Rights Management (DRM). В цій технології компанія Microsoft досягла доволі високих результатів, так як вона працює з нею вже не один рік. DRM дозволяє встановлювати права доступу до медіа контенту, та надає змогу безпечно розповсюджувати цей контент. Digital Rights Management надає можливість створити ліцензію, строк придатності цієї ліцензії, та встановити її на сам медіа контент. Таким чином встановлені власником правила вбудовуються на сам файл і діятимуть навіть якщо файл буде завантажено на комп`ютер.

На рис. 1 приведено схему зміни швидкості передачі відео матеріалу з сервера клієнту. Швидкість може змінюватись від навантаження на канал передачі як з боку клієнта (відкриття додаткових web-сторінок, завантаження будь-яких файлів з мережі), як з боку сервера (велика кількість вхідних підключень, віддаленість сервера від клієнта, некоректне налаштування локальної мережі, в якій знаходиться сервер), так і з боку провайдера Інтернет-послуг.

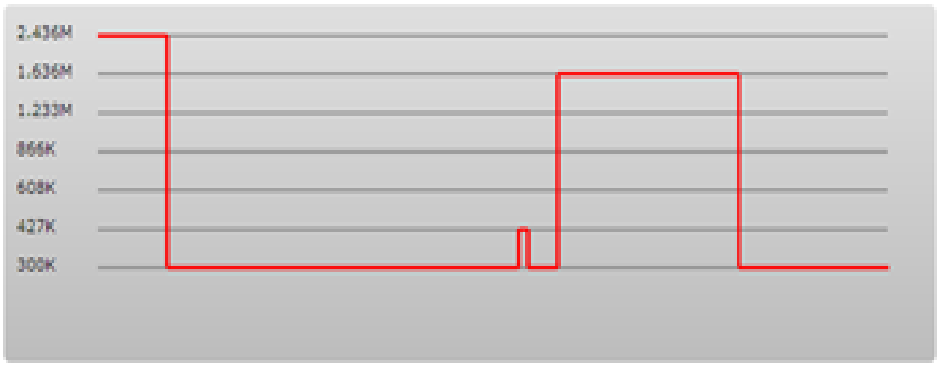


Рисунок 1 – Графік зміни швидкості передачі відео

Фактично, за допомогою даної технології можна встановлювати специфічні права доступу, що зможуть захистити медіа контент від використання будь-яким іншим чином, окрім заданого власником.

Тому, слід відмітити, що потокове відео на базі технологій компанії Microsoft дійсно зможе задовольнити користувачів за якістю, при цьому не вимагатиме надто великих та складних операцій по встановленню додаткового ПК чи модулів та оновлень до вже встановлених, матиме максимальну сумісність і при цьому буде достатньо захищеним.

Висновки

В ході написання даної статті було досліджено методи захисту відеоматеріалів на web-ресурсів та розроблено програмний модуль. Так, на сьогоднішній день все більше і більше компаній переходять на цифровий формат, представляючи різноманітні відеоматеріали на своїх web-сайтах для загального ознайомлення. Запропонований в даній статті метод вирішення цієї проблеми використовує новітні технології і може оновлюватись синхронно із оновленням використаних технологій.

Крім цього було проаналізовано сучасний стан питання, що розглядається, наведено найпоширеніші існуючі способи вирішення задачі. Також було проаналізовано існуючі аналоги та проведено порівняння з ними. В результаті було аргументовано, що захист відеоматеріалів у обраних аналогах поверхневий та недостатній, а отже розробка даного продукту є актуальною та необхідною.

Окрім цього було практично перевірено та обрано технології для реалізації потокового відео на web-ресурсі, висунуто вимоги до сервера, технічне та програмне забезпечення для коректного відображення та роботи з написаним програмним продуктом.

Результати проведених досліджень стануть основою для написання програмного продукту для захисту відеоматеріалів на web-ресурсі.

Література

1. Рихтер Дж. Программирование на платформе Microsoft .NET / Дж. Рихтер, 2009. – 420 с.
2. Мэтью Макдональд Silverlight and ASP.NET Revealed / Макдональд Мэтью 2007. – 256 с.
3. Мэтью Макдональд Pro Silverlight 4 in C# / Макдональд Мэтью 2010. – 346 с.
4. Gill Cleeren Microsoft Silverlight 4 Data and Services Cookbook / Cleeren Gill, Dockx Kevin 2010. – 593 с.
5. Nick Lecrenski Silverlight 4: Problem – Design – Solution / Lecrenski Nick 2010. – 215 с.
6. С. С. Байдачный Silverlight 4. Создание насыщенных Web-приложений / Байдачный С. С. 2010. – 374 с.
7. Вікіпедія – Потокове мультимедіа. // http://uk.wikipedia.org/wiki/ Потокове_мультимедіа – 2011. – 14 трав.
8. Ю.Є. ЯРЕМЧУК Захист інформації і безпека інформаційних систем / ЯРЕМЧУК Ю.Є. Львів, Україна: 05 – 06 червня 2014. – 159 с.
9. Вікіпедія – Microsoft Silverlight. // http://uk.wikipedia.org/wiki/ Microsoft Silverlight – 2012. – 27 бер.
10. 100 найбільш відвідуваних web-ресурсів за 2011 рік. Рейтинг по країнам світу. http://www.alexa.com/topsites/countries/UA – 2012. – 15січ.
11. Microsoft Програмно-апаратні вимоги для Silverlight, ASP.NET, SQL Server 2008 R2. // http://www.microsoft.com/visualstudio/ru-ru/products /2010-editions/professional/system require-ments - 2010. – 20 квіт.
12. Сейтвелиева С.Н. Видеоурок как элемент образовательной технологии / С.Н. Сейтвелиева // Сучасні інформаційні технології та та інноваційні методики навчання у підготовці фахівців: методологія, теорія, досвід, проблеми : зб. нак. пр. – 2010. –№ 24. – С. 131-135.А. Петраков Защитные информационные технологии аудіо видео электросвязи 2010р.

Відомості про авторів

Роїк Олександр Митрофанович – д. т. н., професор, завідувач кафедри менеджменту та безпеки інформаційних систем, Вінницький національний технічний університет, Хмельницьке шосе, 95, м. Вінниця, 21021 : тел. 598294.

Міронова Юлія Володимирівна – к. е. н., старший викладач кафедри менеджменту та безпеки інформаційних систем, Вінницький національний технічний університет, Хмельницьке шосе, 95, м. Вінниця, 21021 : тел. 598294.

Волкотруб Олена Петрівна – студентка четвертого курсу ФМ, ВНТУ.

УДК 621.38

М. А. ФІЛИНЮК, Л. Б. ЛІЩИНСЬКА, О. В. ВОЙЦЕХОВСЬКА, В. П. СТАХОВ

Вінницький національний технічний університет, м. Вінниця

МОНОІМІТАНСНИЙ ЛОГІЧНИЙ R-ЕЛЕМЕНТ «НІ»

Аннотація. Запропоновано моноімітансний логічний R-елемент «НІ», робота якого базується на використанні нечіткого імітансу, а як інформаційний параметр використовується активний імітанс. Розроблено математичну модельтакого моноімітансного логічного R-елемента«НІ»,що реалізованийв вигляді чвертьхвильовоговідрізка лінії передачі. Отримано таблицю істинності логічного R-елемента«НІ». Проведено дослідження впливу паразитних факторів на передатну характеристику моноімітансногологічного елемента, які показали, що найбільший вплив на характеристики логічного елемента здійснюють частота сигналу та паразитний реактивний опір $X_{вх}$. Рекомендовано використовувати відрізок лінії передачі з максимально можливим значенням Z_0, а значення $R_{вх}$ вибирати в діапазоні $0,5 < \overline{R}_{вх} < 1,5$.

Ключові слова: логічний елемент, відрізок лінії передачі, імітанс.

Аннотация. Предложенмоноиммитансный логический R-элемент «НЕ», работа которого основана на использовании нечеткого иммитанса, а в качестве информационного параметра используется активный иммитанс. Разработана математическая модель такого моноиммитансного логического R-элемента «НЕ», реализованного в виде четвертьволнового отрезка линии передачи. Получено таблицу истинности логического R-элемента «НЕ». Проведены исследования влияния паразитных факторов на передаточную характеристику моноиммитансного логического элемента, которые показали, что наибольшее влияние на характеристики логического элемента оказывают частота сигнала и паразитное реактивное сопротивление Хвх. Рекомендуется использовать отрезок линии передачи с максимально возможным значением Z_0, а значение $R_{вх}$ выбирать в диапазоне $0,5 < \overline{R}_{вх} < 1,5$.

Ключевые слова: логическийэлемент, отрезоклинииипередачи, иммитанс.

Annotation. Monoimmittance logic R- gate "NOT", which is based on the use of fuzzy immittance is proposed.Activeimmittance is used as an information parameter. The mathematical model of the monoimmittance logic R- gate "NOT" is designed, this R-gate "NOT" is realized in the form of quarter-wavelength transmission line. A truth table of the logical R-element "NOT" is received. We investigated the effect of parasitic factors on the transfer characteristic of monoimmittance logic gate which showed the greatest influence on the characteristics of the gate providing the signal frequency and parasitic reactance$X_{вх}$. It is recommended to use transmission line segment with the highest possible value Z_0, and select a value $R_{вх}$ in the range $0,5 < \overline{R}_{вх} < 1,5$.

Keywords: logic gate, the segment of the transmission line, immittance.

Вступ

Незважаючи на значні досягнення в галузі створення елементарної бази сучасної обчислювальної техніки постійно ведуться пошуки нових принципів побудови таких елементів, що відповідають певним специфічним вимогам. Одним з напрямків досліджень є розробка радіочастотних логічних елементів, здатних здійснювати логічні операції на несучій частоті сигналу [1, 2, 3]. Більшість технічних рішень в цій області базується на використанні нелінійних властивостей напівпровідникових приладів, що обмежує їх швидкодію і енергетичну ефективність. Як показано в [2], швидкодія радіочастотних логічних елементів в принципі не може перевищувати швидкодію відеоімпульсних логічних елементів. Новим напрямком побудови радіочастотних логічних елементів є використання при їх реалізації принципу нечіткого імітанса [4], коли логічний рівень визначається не кількісною величиною імітанса, а тільки його характером, наприклад «0» - ємнісний, «1» - індуктивний. Такі імітансні логічні елементи належать до групи «мультиімітансних елементів» і можуть бути реалізовані як на активних [5], так і тільки на пасивних [6] радіоелектронних компонентах. В останньому випадку істотно підвищується завадозахищеність та енергетична ефективність логічного елемента, тому що відбуваються тільки енергетичні втрати на частоті сигналу.

З практичної точки зору в ряді випадків доцільно характеризувати логічний стан імітансами одного характеру (або ємнісний (X_C), або індуктивний (X_L), або активний (R)). Такі імітансні логічні елементи відносяться до групи «моноімітансних елементів». На відміну від мультиімітансних логічних елементів, у них логічні рівні визначаються межами зміни імітанса, як у відеоімпульсних елементів діапазонами зміни напруги [7]. Поступаючись мультиімітансним елементам за завадозахищеністю, такі елементи зберігають їх переваги за швидкодією та енергетичною ефективністю, що визначає актуальність їх розробки. Враховуючи, що будь-який реактивний імітанс не існує без активного імітанса (якщо не використані спеціальні заходи компенсації) при розробці моноімітансних логічних елементів існує необхідність оцінювання впливу паразитних імітансів на параметри логічного елемента.

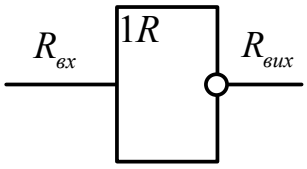
Метою роботи є розробка моноімітансного логічного R-елемента «НІ» та дослідження його основних параметрів. Для досягнення цієї мети вирішуються такі завдання:

1.Обгрунтування схеми моноімітансного логічного R-елемента «НІ» і оцінювання його потенційного швидкодії та енергетичних втрат.

2. Дослідження впливу дестабілізуючих факторів на характеристики логічного елемента.

Обґрунтування схеми моноімітансного логічного R-елемента «НІ» і оцінювання його потенційних параметрів

Виходячи з сформованої мети роботи, як інформаційний параметр використовуємо активний імітанс «R», а логічний стан елемента будемо характеризувати діапазоном значень цього параметра. Наприклад, логічній одиниці - «1» відповідає діапазон зміни $R^{(1)} > R_0$, а логічному нулю - «0» відповідає діапазон зміни $R^{(0)} < R_0$, де R_0 - резистивна межа логічного рівня. З урахуванням введених припущень, роботу моноімітансного логічного R-елемента «НІ» можна описати таблицею істинності, яка подана на рис. 1.

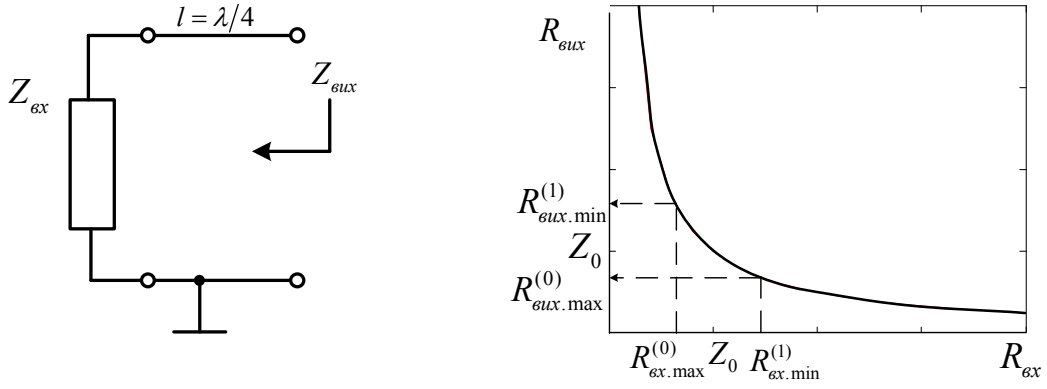


$R_{вх}$	$R_{вих}$
1 $R_{вх}^{(1)} > R_0$	0 $R_{вих}^{(0)} < R_0$
0 $R_{вх}^{(0)} < R_0$	1 $R_{вих}^{(1)} > R_0$

а) б)

Рисунок 1 – Схематичне позначення (а) і таблиця істинності (б) моноімітансного логічного R-елемента «НІ»

Для реалізації логічної функції, що описується таблицею на рис. 1б, запропоновано використовувати властивості чвертьхвильового відрізка лінії передачі, що реалізує «чвертьхвильовий трансформатор» (рис. 2а).



а) б)

Рисунок 2 – Чвертьхвильовий трансформатор (а), який реалізує моноімітансний логічний R-елемент «НІ», і його передатна характеристика (б)

Вихідний опір $Z_{вих}$ такого відрізка лінії передачі залежить від опору $Z_{вх}$, що підключається на його вході[8]

$$Z_{вих} = Z_0^2 / Z_{вх}, \quad (1)$$

де Z_0 - хвильовий опірлінії передачі.

Якщо$Z_{вх} = R_{вх}$, тоді$Z_{вих} = R_{вих} = Z_0^2 / R_{вх}$.Вважаючи, що Z_0 має дійсне фіксоване значення, передатна характеристика такого елемента має вигляд, поданий на рис. 2б. З графіка видно, що при$R_{вх} > Z_0$, $R_{вих} < Z_0$і навпаки, що відповідає вище сформульованим умовам побудови моноімітансного логічного R-елемента «НІ».

Найважливішими параметрами логічного елемента є його швидкодія і споживана потужність.

Швидкодія моноімітансного логічного R-елемента «НІ» визначається затримкою встановлення імітанса$R_{вих}$ виходу логічного елемента при зміні значення імітанса$R_{вх}$ кола, що підключено до його входу. Ця затримка визначається часом проходження електромагнітної хвилі від входу до виходу елемента і залежить від довжини відрізка лінії передачі l і швидкості v електромагнітної хвилі в лінії передачі. Для розглянутого логічного елемента «НІ» маємо:

$$\tau = l / v = \sqrt{\varepsilon} / 4 f_0, \quad (2)$$

де ε - відносна діелектричнапроникність заповнення відрізка лінії передачі, f_0 – робоча частота сигналу. З (2) видно, що чим вища частота f_0 сигналу, тем більшашвидкодіяімітансного логічного елемента. Наприклад, при використанні діелектричної підкладки з ε =1 на частоті 1 ГГц маємо τ (1 ГГц) =0,25 нс, при f = 10 ГГц – маємо τ (10 ГГц) = 25 пс.

Розглянутий логічний елемент не використовує джерело постійної енергії ($P_0 = 0$). Основні витрати енергії $P_{\sim}$ пов'язані з енергетичними втратами в лінії передачі $P_Л$ і поглинанням сигналу P_R на омічному опорі кола, що підключається до входу логічного елемента $P_{\sim} = P_Л + P_R$. Внаслідок малої довжини відрізка лінії передачі і низького значення діелектричних втрат підкладки, втратами $P_Л$ в лінії передачі можна знехтувати. Потужність сигналу, що поглинається на опорі $R_{вх}$ залежить від потужності сигналу $P_{\sim}$ несучої частоти. Її мінімальне значення $P_{\sim min}$ повинно бути більше потужності шумів, створюваних омічним опором $R_{вх}$, що визначається виразом [9]$P_ш$=4kTΔf,де k – сталаБольцмана, Δf– смуга робочих частот, T – абсолютна температура.

Мінімальна смуга робочих частот визначається стабільністю частоти опорного генератора. Наприклад, при відносній стабільності частоти генератора $\Delta f / f = 10^{-6}$ на несучій частоті 1 ГГц маємо – Δf =1 кГц, а на частоті 10 ГГц - Δf = 10 кГц. В цьомувипадку$P_ш$(1ГГц)=1,6x10^{-19} Вт и $P_ш$(10ГГц)=1,6x10^{-18} Вт. Отже $P_{\sim}$(1ГГц)>1,6x10^{-19}Вт, а $P_{\sim}$(10 ГГц)>1,6x10^{-18} Вт. Використовуючизагальноприйняте поняття фактора якості $Q = \tau P_{\sim}$ [7], для ідеального моноімітансного логічного R-елемента «НІ», що розглядається, знаходимо: Q(1 ГГц) = Q(10 ГГц) =4x10^{-5}пДж.

Вплив дестабілізуючих факторів на характеристики логічного елемента

Параметри реального моноімітансного логічного R-елемента «НІ» будуть відрізнятися від параметрів вище розглянутого «ідеального» елемента через вплив дестабілізуючих факторів. По-перше, для забезпечення завадостійкості в реальному елементі границя між логічним нулем та одиницею визначається не Z_0 , а діапазоном значень $R_{вх}$та $R_{вих}$ (рис. 2б):

$$R^{(0)}_{вх.\max} < R_{вх} < R^{(1)}_{вх.\min} \quad , \tag{3}$$

$$R^{(0)}_{вих.\max} < R_{вих} < R^{(1)}_{вих.\min} \, . \tag{4}$$

В цьому випадку діапазон можливих значень активного імпеданса, що відповідає логічному «0» та «1» дорівнює:

- для вхідного імпедансу

$$0 < R^{(0)}_{вх} \le R^{(0)}_{вх.\max}; \quad R^{(1)}_{вх.\min} < R^{(1)}_{вх} < \infty \, , \tag{5}$$

- для вхідного імпедансу

$$0 < R^{(0)}_{вих} \le R^{(0)}_{вих.\max}; \quad R^{(1)}_{вих.\min} < R^{(1)}_{вих} < \infty \, . \tag{6}$$

В реальному моноімітансному логічному елементі вихіднийімітанс $Z_{вых}$ пов’язаний з вхідним імітансом $Z_{вх}$ співвідношенням [8]

$$\overline{Z}_{вих} = \frac{\overline{Z}_{вх} + j \cdot tg\beta}{1 + j \cdot \overline{Z}_{вх} \cdot tg\beta} = \mathrm{Re}\,\overline{Z}_{вих} + j \cdot \mathrm{Im}\,\overline{Z}_{вих} \, , \tag{7}$$

де:

$$\mathrm{Re}\,\overline{Z}_{вих} = \overline{R}_{вих} = \frac{\overline{R}_{вх} \cdot (1 - \overline{X}_{вх} \cdot tg\beta) + \overline{R}_{вх} \cdot tg\beta \cdot (\overline{X}_{вх} + tg\beta)}{(1 - \overline{X}_{вх} \cdot tg\beta)^2 + (\overline{R}_{вх} \cdot tg\beta)^2} \, , \tag{8}$$

$$\mathrm{Im}\,\overline{Z}_{вих} = \overline{X}_{вих} = \frac{(1 - \overline{X}_{вх} \cdot tg\beta) \cdot (\overline{X}_{вх} + tg\beta) - \overline{R}^2_{вх} \cdot tg\beta}{(1 - \overline{X}_{вх} \cdot tg\beta)^2 + (\overline{R}_{вх} \cdot tg\beta)^2} \, , \tag{9}$$

$\beta = 2\pi l / \lambda$ - фазова стала; $\overline{R}_{вих} = \dfrac{R_{вих}}{Z_0}$, $\overline{X}_{вих} = \dfrac{X_{вих}}{Z_0}$, $\overline{R}_{вх} = \dfrac{\mathrm{Re}\,Z_{вх}}{Z_0}$, $\overline{X}_{вх} = \dfrac{\mathrm{Im}\,Z_{вх}}{Z_0}$.

Залежність (8) є передатною характеристикою реального моноімітансного логічного R-елемента «НІ». З її аналізу видно, що основними дестабілізуючими факторами, що впливають на виконання нерівностей (5) та (6) є:

- нестабільність хвильового опору Z_0 відрізка лінії передачі;
- нестабільність фазової сталої β ;

- поява у вхідному імітансі паразитної реактивної складової $X_{вх}$.

Оцінимо вплив нестабільності хвильового опору Z_0 лінії передачі на чутливість вихідного активного опору логічного елемента за умови, що всі інші дестабілізуючі фактори відсутні. В цьому випадку передатна характеристика логічного елемента описується рівнянням (1) і на його основі знаходимо

$$S_{Z_0}^{R_{вих}} = \frac{\partial R_{вих}}{\partial Z_0} \Big/ \frac{R_{вих}}{Z_0} = 2 . \qquad (10)$$

З (10) видно, що чутливість вихідного активного опору від зміни хвильового опору лінії передачі стала та дорівнює 2.

При оцінюванні чутливості вихідного активного опору $R_{вих}$логічного елемента від впливу паразитного реактивного опору $X_{вх}$ будемо вважати, що Z_0, f та l_0 відповідають ідеальному логічному елементу. В цьому випадку, використовуючи (8), знаходимо

$$S_{X_{вх}}^{R_{вих}} = \frac{\partial R_{вих}}{\partial X_{вх}} \Big/ \frac{R_{вих}}{X_{вх}} = -\frac{2\bar{X}_{вх}^2 - 2\bar{X}_{вх} ctg\beta}{\bar{R}_{вх}^2 + \bar{X}_{вх}^2 - 2\bar{X}_{вх} ctg\beta + ctg^2\beta} . \qquad (11)$$

На основі (11) отримано розрахункові залежності чутливості $S_{X_{вх}}^{R_{вих}}$ при різних значеннях $\bar{R}_{вх}$ (рис. 3).

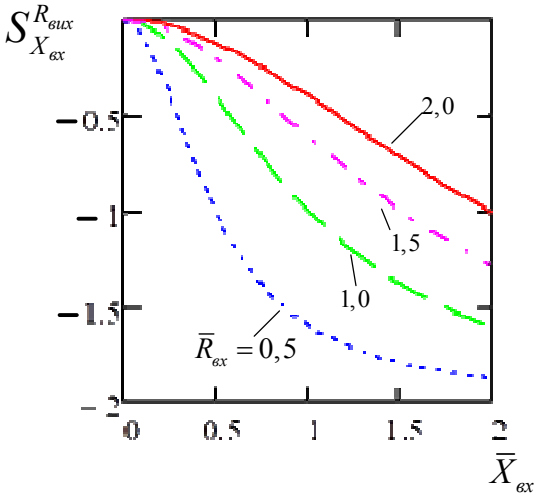


Рисунок 3 – Розрахункові залежності чутливості $S_{X_{вх}}^{R_{вих}}$ при різних значеннях $\bar{R}_{вх}$ (а) та Q (б)

З графіків видно, що чим менші приведені значення $\bar{X}_{вх}$ та $\bar{R}_{вх}$, тим менша величина $S_{X_{вх}}^{R_{вих}}$. Це дозволяє зробити висновок про можливість зменшення цієї нестабільності шляхом використання лінії передачі з великим хвильовим опором Z_0. Інший шлях зниження цієї похибки полягає в накладанні обмеження на добротність $Q = X_{вх}/R_{вх}$ імпеданса, що визначає логічний рівень $R_{вх}$. При $Q_R < 1$ маємо $S_{X_{вх}}^{R_{вих}} < -1$.

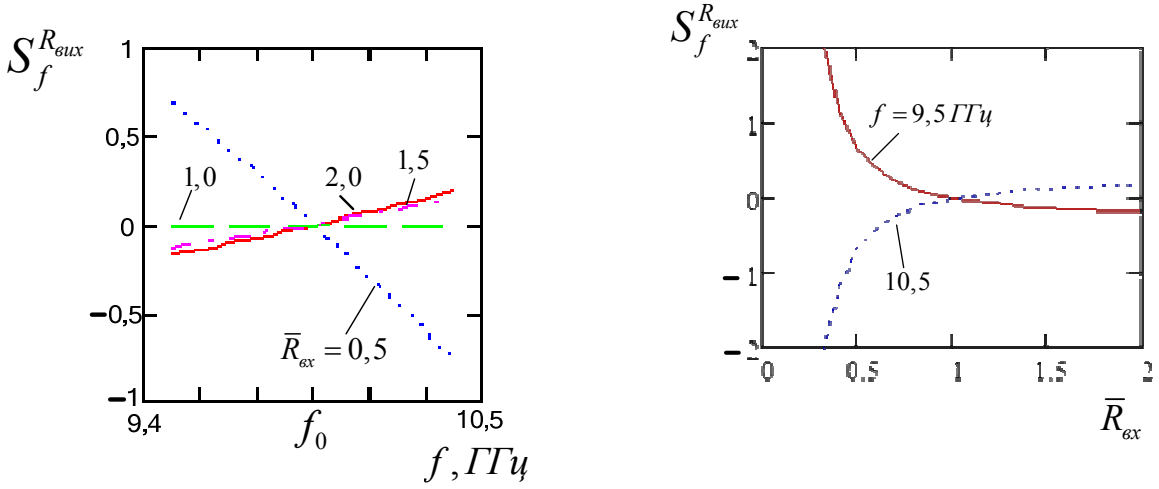
Фазова стала β залежить від опорної частоти сигналу f_0. Враховуючи, що $l = \lambda_0/4 = c/4f_0$, та використовуючи (8), знаходимо

$$S_f^{R_{вих}} = -\frac{\pi f \lambda_0 \bar{R}_{вх}^2 tg\left(\frac{\pi f \lambda_0}{2c}\right) + \pi f \lambda_0 \bar{X}_{вх}^2 tg\left(\frac{\pi f \lambda_0}{2c}\right) + \pi f \lambda_0 \bar{X}_{вх} tg\left(\frac{\pi f \lambda_0}{2c}\right)^2 -}{c\bar{R}_{вх} tg\left(\frac{\pi f \lambda_0}{2c}\right)^2 + c\bar{X}_{вх}^2 tg\left(\frac{\pi f \lambda_0}{2c}\right)^2 -} \quad (12)$$

$$\frac{-\pi f \lambda_0 \bar{X}_{вх} - \pi f \lambda_0 tg\left(\frac{\pi f \lambda_0}{2c}\right)}{-2c\bar{X}_{вх} tg\left(\frac{\pi f \lambda_0}{2c}\right) + c},$$

де f– частота опорного сигналу вдіапазоніїї нестабільності, с – швидкість електромагнітної хвилі в вакуумі.

Використовуючи (12) на рис. 4 подано розрахункові залежності чутливості $\bar{R}_{вих}$ до зміни несучої частоти сигналу f при різних значеннях $\bar{R}_{вх}$, незмінному Z_0 та $\bar{X}_{вх}$=0.



а) б)

Рисунок 4 – Розрахункові залежності чутливості $S_f^{R_{вих}}$ до зміни несучої частоти сигналу f при різних значеннях $\bar{R}_{вх}$ та незмінному Z_0 та $\bar{X}_{вх}$=0

Як видно з графіків, при нестабільності несучої частоти в діапазоні $\pm 5\%$, величина $S_f^{R_{вих}}$ не перевищує ± 1. Зменшення величини цієї нестабільності можливе шляхом вибору зміни $\bar{R}_{вх}$ в діапазоні $0,5 < \bar{R}_{вх} < 1,5$.

Висновки

1. Запропонована схема моноімітансного логічного R-елемента «НІ», що реалізується в вигляді чвертьхвильового відрізка лінії передачі, за умови використання як логічної «1» значення активного імітанса $R^{(1)} > Z_0$, а логічного «0» - $R^{(0)} < Z_0$.

2. Проведені дослідження впливу паразитних факторів (зміни хвильового опору Z_0 лінії передачі та опорної частоти f_0 сигналу, а також наявність паразитного вхідного імітанса $X_{вх}$) на передатну характеристику логічного елемента показали, що величина Z_0 не впливає на значення $S_{Z_0}^{R_{вих}}$. Найбільший вплив здійснюють зміна f та $\bar{X}_{вх}$. Для їх зменшення рекомендується використовувати відрізок лінії передачі з максимально можливим значенням Z_0, а значення $R_{вх}$ вибирати в діапазоні $0,5 < \bar{R}_{вх} < 1,5$.

Список літератури

1. Электронные радиоимпульсные вычислительные системы дискретного действия / Под ред. Неймана М. С. . // Труды московського авиационного института. Том 149. – М. : Оборонгиз, 1962. – 186 с.
2. Кнорре К. Г. Фазовые и частотне информационные СВЧ элементы / Кнорре К.Г., Тузов В.М., Шур Г.И.. – М.: Сов. радио, 1975. – 352 с.
3. Долгий И. Д. Введение в безопасную радиочастотную логику / Долгий И. Д., Кулькин А. Г., Кулькин С. А., ПономаревЮ.Э., Розенберг И.Н. // Известия Южного федерального университета. Технические науки. Выпуск № 6 (155). – 2014. – С. 229-238.
4. Лищинская Л.Б. Обоснование концепции «нечёткогоиммитанса» / Лищинская Л. Б. // Вимірювальна та обчислювальна техніка в технологічних процесах. – 2010. – №1. – С. 20–25.
5. Ліщинська Л. Б. Імітансна логіка / Л. Б. Ліщинська, М. А. Філинюк // Інформаційні технології та комп'ютерна інженерія. – 2010. – № 2(18). – С. 25–31.
6. Пасивний імітансний LC-логічний елемент АБО-НІ / М. А. Філинюк, Л. Б. Ліщинська, О. В. Войцеховська, Р. Ю. Чехместрук // Інформаційні технології та комп'ютерна інженерія. - 2014. - № 3. - С. 72-77.
7. Мальцева Л. А. Основы цифровой техники. / Мальцева Л. А., Фромберг Э. М., Ямпольский В. С.. – М. : Радиоисвязь, 1986. – 128 с.
8. Лебедев И. В. Техника и приборы СВЧ : учебное пособие. Том 1 / Лебедев И.В. – М. : Высшая школа, 1970 – 439 с.
9. Савицкая М.П. Аналоговые электронные устройства : учебное пособие / М. П. Савицкая, Л. Б. Ботнарь.– Одесса: ОНАС им. А. С. Попова, 2008. – 108 с.

Відомості про авторів

Філинюк Микола Антонович - доктор технічних наук, професор, завідувач кафедри проектування комп'ютерної та телекомунікаційної апаратури ВНТУ, м. Вінниця, Хмельницьке шосе, 95.

Ліщинська Людмила Броніславівна - доктор технічних наук, професор, завідувач кафедри економічної кібернетики та інформаційних систем ВТЕІ КНТЕУ, м. Вінниця, вул. Соборна, 87.

Войцеховська Олена Валеріївна - кандидат технічних наук, доцент, доцент кафедри проектування комп'ютерної та телекомунікаційної апаратури ВНТУ, м. Вінниця, Хмельницьке шосе, 95.

Стахов Володимир Петрович - аспірант кафедри проектування комп'ютерної та телекомунікаційної апаратури ВНТУ, м. Вінниця, Хмельницьке шосе, 95, (0432)598-075.

УДК 004.75

В. В. ЯЦКІВ

Тернопільський національний економічний університет, Тернопіль

ВИЯВЛЕННЯ ТА ВИПРАВЛЕННЯ БАГАТОКРАТНИХ ПОМИЛОК НА ОСНОВІ МОДУЛЯРНИХ КОРЕКТУЮЧИХ КОДІВ

Анотація. У вступі описано обмеження коректуючих кодів в системі залишкових класів. Визначено актуальність розробки коректуючих кодів на основі модулярної арифметики призначених для захисту від помилок даних представлених у двійковій системі числення. Сформульована мета та задачі досліджень для даної статті. Розроблено метод виявлення та виправлення двохкратних помилок за допомогою двох перевірочних символів. Розроблено алгоритм кодування та декодування даних на основі запропонованих модулярних кодів. Приведено приклад виправлення помилок в двох інформаційних символах. В кінці статті наведено список відомих наукових публікацій по даній темі.

Ключові слова: модулярна арифметика, система залишкових класів, коректуючі коди безпровідні сенсорні мережі.

Аннотация. Во введении описано ограничения корректирующих кодов в системе остаточных классов. Определены актуальность разработки корректирующих кодов основанных на модулярной арифметике и предназначенных для защиты от ошибок данных представленных в двоичной системе счисления. Сформулирована цель и задачи исследований для данной статьи. Разработан метод выявления и исправления двукратных ошибок с помощью двух проверочных символов. Разработан алгоритм кодирования и декодирования данных на основе предложенных модулярных кодов. Приведен пример исправления ошибок в двух информационных символах. В конце статьи приведен список известных научных публикаций по данной теме.

Ключевые слова: модулярная арифметика, система остаточных классов, корректирующие коды, беспроводные сенсорные сети.

Abstract. The restrictions of corrective codes in the Residue Number System is described in introduction. The relevance of development correcting codes based on modular arithmetic and designed to protect the data presented in the binary system from errors is determined. The objective and tasks of the research for this article is stated. A method for two errors detection and correction with the use of two check symbols is elaborated. The algorithm coding and decoding based on the proposed modular codes is elaborated. An example of two errors correction is shown. At the end of the article is a list of scientific publications on this topic

Key words: modular arithmetic, residue number system, corrective codes, wireless sensor networks.

Вступ

В безпровідних сенсорних мережах для підвищення надійності передачі даних використовують циклічний контроль парності (CRC – коди), який забезпечує ефективне виявлення помилок [1]. В [2] проведено аналіз методів виявлення та виправлення помилок та показано недоліки застосування існуючих коректуючих кодів в безпровідних сенсорних мережах (БСМ). БСМ є розділеною та самоорганізованою мережею сенсорів і виконавчих пристроїв, які об’єднані між собою безпровідними радіоканалами зв’язку.

Враховуючи переваги системи залишкових класів (СЗК) при обробці та передаванні даних, особливий інтерес, для використання в БСМ, представляють коректуючі коди на основі СЗК та модулярної арифметики [3, 4]. Однак для їх використання в існуючих цифрових системах оброблення та передавання даних, зокрема в БСМ, необхідно попередньо перетворити дані в СЗК, що потребує додаткових затрат часу та обчислювальних ресурсів [5, 6]. Коректуючі коди СЗК мають також обмеження, які накладаються умовою вибору зростаючої послідовності взаємно простих модулів. При цьому необхідно щоб перевірочні модулі були однакової розрядності. В іншому випадку виникає проблема ефективного (оптимального) збереження перевірочних символів.

Актуальність

З розвитком та широким впровадженням безпровідних технологій задача підвищення надійності передачі даних набуває все більш важливого значення. Відповідно, контроль цілісності даних, які передаються та обробляються сучасними цифровими системами є актуальною науковою задачею, зокрема в безпровідних сенсорних мережах.

В [7] розроблено та досліджено модулярні коректуючі коди, які забезпечують виправлення однократних помилок. Дані коди зберігають переваги коректуючих кодів СЗК, але обробляють вхідні дані представлені в позиційній системі числення (двійковій, десятковій), що значно спрощує процедури кодування / декодування та розширює область їх застосування. Отже, розробка коректуючих кодів на основі модулярної арифметики, які призначені для захисту даних представлених у двійковій системі числення від багатократних помилок є актуальною науковою задачею.

Мета

Метою статті є розробка коректуючих кодів на основі модулярної арифметики для виявлення та виправлення багатократних помилок.

Задачі

1. Для досягнення вказаної мети необхідно довести можливість виправлення помилок в двох інформаційних символах при використанні двох перевірочних модулів.
2. Розробити алгоритм виявлення та виправлення багатократних помилок.

Виявленння та виправлення помилок в одному інформаційному символі

Нехай $X \equiv (x_1, x_2, \ldots, x_i, \ldots, x_k)$ - кортеж символів заданої розрядності представлених в двійковій системі числення, які необхідно передати; $E_i \equiv (0,\ 0, \ldots, e_i, \ldots, 0)$ – код помилки; $X' \equiv (x_1', x_2', \ldots, x_i', \ldots, x_k')$ – кортеж символів в результаті дії завад:

$$(x_1', x_2', \ldots, x_i', \ldots, x_k') = (x_1, x_2, \ldots, x_i, \ldots, x_k) + \\ + (0, 0, \ldots, e_i, \ldots, 0) = (x_1, |x_2 + e_i|_P, \ldots, x_i, \ldots, x_k).$$

Значення контрольного символу обчислюється за формулою [8]

$$x_{k+1} = |(v_1 \cdot x_1 + v_2 \cdot x_2 + \ldots + v_i \cdot x_i + \ldots + v_k \cdot x_k)|_P, \quad (1)$$

де v_i – коефіцієнти взаємно прості з P, $|\bullet|_P$ – операція отримання залишку по модулю P.

Декодер по прийнятих даних $(x_1', x_2', \ldots, x_i', \ldots, x_k')$ обчислює значення контрольного символу:

$$x_{n+1}' = \left|(v_1 \cdot x_1' + v_2 \cdot x_2' + \ldots + v_i \cdot x_i' + \ldots + v_k \cdot x_k')\right|_P. \quad (2)$$

Для виявлення помилки обчислимо синдром δ, який представляє різницю між отриманим перевірочним символом і перевірочним символом обчисленим на приймальній стороні (в декодері):

$$\delta = \left|x_{k+1}' - x_{k+1}\right|_P,$$

якщо синдром дорівнює нулю, $\delta = 0$ то помилки немає, так як при відсутності помилки $x_i' = x_i$, відповідно і $x_{k+1}' = x_{k+1}$, якщо $\delta \neq 0$ – є помилка, відповідно $x_i' \neq x_i$ і, як наслідок $x_{k+1}' \neq x_{k+1}$.

Для виправлення помилок необхідно щоб значення синдрому δ було унікальне для всіх можливих варіантів помилки. Дана умова виконується при дотриманні правил: 1) значення контрольного модуля $P > 2 \cdot k \cdot (2^m - 1)$, де k – кількість інформаційних символів; 2) коефіцієнти v_i – взаємно прості з P.

Локалізація та виправлення помилки в одному інформаційному символі здійснюється шляхом порівняння синдрому δ з таблицею попередньо знайдених розв'язків рівняння $S_{ij} = \left|v_i \cdot e_{ij}\right|_P$, для всіх $-2^m - 1 < e_{ij} < 2^m - 1$, де $i = 1, \ldots, n$, $j = -2^m - 1, \ldots, 2^m - 1$.

На основі аналізу значень S_{ij} визначаємо символ в якому відбулася помилка.

Для виправлення помилки в одному символі необхідно розв'язати рівняння

$$\delta = \left|x_{k+1}' - x_{k+1}\right|_P = \left|v_i \cdot \left|x_i' - x_i\right|_P\right|_P,$$

так як $x_i' = \left|x_i + e_j\right|_P$ то $\left|v_i \cdot \left|\left|x_i + e_j\right|_P - x_i\right|_P\right|_P = S_{ij}$,

$$\left|v_i \cdot x_i + v_i \cdot e_j - v_i \cdot x_i\right|_P = S_{ij},$$

звідки слідує

$$S_{ij} = \left| v_i \cdot e_j \right|_P . \quad (3)$$

З рівняння (3) знаходимо значення помилки e_j. При цьому, правильне значення інформаційного символу дорівнює:

$$x_i = \left| x_i^{'} - e_j \right|_P .$$

Виявлення та виправлення помилок в двох інформаційних символах

В даній роботі розроблено метод виявлення та виправлення багатократних помилок на основі модурярних коректуючих кодів. Для забезпечення можливості виправлення помилок в двох інформаційних символах вибираємо взаємно прості коефіцієнти w_i і обчислимо додатковий перевірочний символ x_{k+2}.

Значення перевірочного символу обчислюється за формулою:

$$x_{k+2} = \left| (w_1 \cdot x_1 + w_2 \cdot x_2 + \ldots + w_i \cdot x_i + \ldots + w_k \cdot x_k) \right|_P . \quad (4)$$

В декодері по прийнятих даних обчислюється перевірочний символ $x_{k+2}^{'}$:

$$x_{k+2}^{'} = \left| (w_1 \cdot x_1 + v_2 \cdot x_2 + \ldots + w_i \cdot x_i^{'} + \ldots + w_k \cdot x_k) \right|_P .$$

В результаті отримали систему рівнянь:

$$\begin{cases} \left| (v_1 \cdot x_1 + v_2 \cdot x_2 + \ldots + v_i \cdot x_i + \ldots + v_k \cdot x_k) \right|_P = x_{k+1} \\ \left| (w_1 \cdot x_1 + w_2 \cdot x_2 + \ldots + w_i \cdot x_i + \ldots + w_k \cdot x_k) \right|_P = x_{k+2} \end{cases} .$$

Обчислюємо синдроми δ_1 і δ_2:

$$\delta_1 = \left| x_{k+1}^{'} - x_{k+1} \right|_P , \quad \delta_2 = \left| x_{k+2}^{'} - x_{k+2} \right|_P ,$$

або

$$\delta_1 = \left| v_1 \cdot (x_1^{'} - x_1) + v_2 \cdot (x_2^{'} - x_2) + \ldots + v_i \cdot (x_i^{'} - x_i) + \ldots + v_k \cdot (x_k^{'} - x_k) \right|_P , \quad (5)$$

$$\delta_2 = \left| w_1 \cdot (x_1^{'} - x_1) + w_2 \cdot (x_2^{'} - x_2) + \ldots + w_i \cdot (x_i^{'} - x_i) + \ldots + w_k \cdot (x_k^{'} - x_k) \right|_P . \quad (6)$$

При $\delta_1 = 0$ і $\delta_2 = 0$ – помилки відсутні, в іншому випадку наявні помилки.

Припустимо, що помилки відбулися в двох символах. Так як при відсутності помилки $x_i^{'} - x_i = 0$, то рівняння (5) і (6) набудуть вигляду:

$$\left| v_i \cdot (x_i^{'} - x_i) + v_{i+1} \cdot (x_{i+1}^{'} - x_{i+1}) \right|_P = \delta_1 , \quad (7)$$

$$\left| w_i \cdot (x_i^{'} - x_i) + w_{i+1} \cdot (x_{i+1}^{'} - x_{i+1}) \right|_P = \delta_2 . \quad (8)$$

З врахування виразів $e_1 = \left|x_1^{'} - x_1\right|_P$ і $e_2 = \left|x_2^{'} - x_2\right|_P$ рівняння (7) і (8) набудуть вигляду:

$$\left|v_1 \cdot e_1 + v_2 \cdot e_2\right|_P = \delta_1,$$

$$\left|w_1 \cdot e_1 + w_2 \cdot e_2\right|_P = \delta_2,$$

де e_1, e_2 – різниця між прийнятим і переданим значенням інформаційного символу.

Розв'язавши рівняння (7) і (8) отримаємо правильні значення інформаційних символів x_i і x_{i+1}.

Отже, даний коректуючий код виправляє помилки в t символах, якщо його мінімальна кодова відстань дорівнює: $d_{\min} \geq t+1$.

Алгоритм локалізації помилки складається з наступних кроків:

1. Обчислення перевірочних символів по прийнятих інформаційних символах: x_{k+1}, x_{k+2};

2. Обчислення синдрому δ_1, δ_2;

3. Якщо $\delta_1 = 0$, $\delta_2 = 0$ – помилки немає, зупинка і вихід. В іншому випадку продовження.

4. Припускаємо, що помилка відбулася в символах x_i і x_{i+1}, $i = 1,\ldots k$. Збільшуємо i, $i = i+1$.

5. Розв'язуємо рівняння (7) і (8);

6. Якщо знайдені $x_i < 2^m - 1$ і $x_{i+1} < 2^m - 1$. Перехід до пункту 2;

7. Збільшуємо i, $i = i+1$. Перехід до пункту 4.

Для виявлення помилки кратної r в кортежі даних, який складається із k інформаційних символів необхідно здійснити $C_k^r = \dfrac{k!}{r!(k-r)!}$ ітерацій, наприклад для $k = 8$, $r = 2$, $C_8^2 = 28$ ітерацій.

Розроблений метод та алгоритм виправлення помилок в двох інформаційних символах може бути використаний для виправлення багатократних помилок для цього необхідно збільшити кількість перевірочних символів.

Особливістю розроблених коректуючих кодів є можливість зміни кількості перевірочних символів x_{k+i}, $i = 1, \ldots, h$, де h – максимальна кількість перевірочних символів, без зміни принципів кодування. Тобто, при збільшенні ймовірності помилки в каналі зв'язку достатньо збільшити кількість перевірочних символів. Вказана можливість забезпечує високу надійність передавання даних при зміні характеристик каналу зв'язку.

Можливість виправлення двох помилок показано в наступному прикладі.

Приклад. Розглянемо коректуючий код, який складається з восьми інформаційних і двох перевірочних символів, і забезпечує виправлення помилок в будь-яких двох інформаційних символах. Розрядність інформаційних символів 4 біти. Кортеж даних, які необхідно передати має вигляд: $X = (5, 8, 10, 3, 7, 14, 12, 1)$.

Вибираємо модуль $P = 1021$ та взаємно прості коефіцієнти для обчислення першого перевірочного символу: $v_1 = 13$, $v_2 = 17$, $v_3 = 19$, $v_4 = 23$, $v_5 = 29$, $v_6 = 31$, $v_7 = 37$, $v_8 = 43$; взаємно прості коефіцієнти для обчислення другого перевірочного символу: $w_1 = 7$, $w_2 = 61$, $w_3 = 73$, $w_4 = 83$, $w_5 = 103$, $w_6 = 199$, $w_7 = 239$, $w_8 = 313$.

Значення перевірочних символів знаходимо за формулою (1):

$$x_{k+1} = \left|13 \cdot 5 + 17 \cdot 8 + 19 \cdot 10 + 23 \cdot 3 + 29 \cdot 7 + 31 \cdot 14 + 37 \cdot 12 + 43 \cdot 1\right|_{1021} = 563;$$

$$x_{k+2} = \left|7 \cdot 5 + 61 \cdot 8 + 73 \cdot 10 + 83 \cdot 3 + 103 \cdot 7 + 199 \cdot 14 + 239 \cdot 12 + 313 \cdot 1\right|_{1021} = 22.$$

Нехай прийняли дані $X^{'} = (5, 4, 6, 3, 7, 14, 12, 1)$. Обчислимо перевірочні символи по прийнятих даних за формулою (2):

$$x^{'}_{k+1} = \left|13 \cdot 5 + 17 \cdot 4 + 19 \cdot 6 + 23 \cdot 3 + 29 \cdot 7 + 31 \cdot 14 + 37 \cdot 12 + 43 \cdot 1\right|_{1021} = 419;$$

$$x^{'}_{k+2} = \left|7 \cdot 5 + 61 \cdot 4 + 73 \cdot 6 + 83 \cdot 3 + 103 \cdot 7 + 199 \cdot 14 + 239 \cdot 12 + 313 \cdot 1\right|_{1021} = 507.$$

Обчислюємо значення синдрому:

$$\delta_1 = \left|x^{'}_{k+1} - x_{k+1}\right|_P = \left|419 - 563\right|_{1021} = 877;$$

$$\delta_2 = \left|x^{'}_{k+2} - x_{k+2}\right|_P = \left|507 - 22\right|_{1021} = 485.$$

Так як синдроми δ_1 і δ_2 не дорівнюють нулю, отже наявні помилки.

Для виявлення помилки в двох інформаційних сиволах необхідно розв’язати рівняння (7) і (8) послідовно припускаючи, що помилки відбулися в двох символах.

Припустимо, що помилки в символах x_2 і x_3. Тоді рівняння (7) і (8) з врахуванням числових значень набудуть вигляду:

$$\left|17 \cdot e_2 + 19 \cdot e_3\right|_P = 877, \tag{9}$$

$$\left|61 \cdot e_2 + 73 \cdot e_3\right|_P = 485. \tag{10}$$

Помножимо рівняння (9) на 61 а рівняння (10) на 17:

$$\left|16 \cdot e_2 + 138 \cdot e_3\right|_P = 405, \tag{11}$$

$$\left|16 \cdot e_2 + 220 \cdot e_3\right|_P = 77, \tag{12}$$

і віднімемо рівняння (12) від рівняння (11): $\left|939 \cdot e_3\right|_P = 328$, за алгоритмом Евкліда знаходимо: $e_3 = 1017$.

Для знаходження e_2 підставимо значення e_3 в рівняння (10):

$$\left|61 \cdot e_2 + 73 \cdot 1017\right|_P = 485,$$

$\left|61 \cdot e_2\right|_P = 777$, отже $e_2 = 1017$.

Отже, правильні значення інформаційних символів дорівнюють:

$$x_2 = \left|x^{'}_2 - 1017\right|_{1021} = \left|4 - 1017\right|_{1021} = 8;\ x_3 = \left|x^{'}_3 - 1017\right|_{1021} = \left|6 - 1017\right|_{1021} = 10.$$

В таблиці 1 приведені результати обчислення значень інформаційних символів при всіх можливих варіантах помилок в двох інформаційних символах. Як видно з таблиці 1 значення інформаційних сим-

волів, в межах робочого діапазону, знаходяться тільки в першому рядку, що відповідає позиції помилок (u_i, u_j) в першому і другому інформаційному символі.

Таблиця 1 – Виявлення помилок в двох інформаційних символах

№	Позиція помилки		Значення інформаційних символів	
	u_i	u_j	x_i	x_i
1	1	2	871	131
2	1	3	616	724
3	1	4	990	74
4	1	5	860	192
5	1	6	988	331
6	1	7	621	103
7	1	8	108	176
8	**2**	**3**	**8**	**10**
9	2	4	978	44
		. . .		
28	7	8	944	754

Висновки

Розроблено коректуючі коди на основі модулярної арифметики, які дозволяють виявлення та виправлення багатократних помилок в пакетах даних. Дані коди забезпечують виправлення помилок в t - інформаційних символах при використанні t перевірочних символів. Представлено Алгоритм локалізації помилки.

Список літератури

1. IEEE Standard for Part 15.4: Wireless Medium Access Control Layer (MAC) and Physical Layer (PHY) specifications for Low Rate Wireless Personal Area Networks (LR-WPANs), IEEE Std 802.15.4- 2006.
2. Howard S. L. Error control coding in low-power wireless sensor networks: When is ECC energy-efficient? / Howard, S. L., Schlegel C., Iniewski K. // EURASIP Journal on Wireless Communications and Networking, №2. – 2006. – P.29-29.
3. Omondi A. Residue Number System: Theory and Implementation. / A.Omondi, B. Premkumar // Imperial College Press, vol. 2, 2007. – P. 296.
4. Акушский И.Я. Машинная арифметика в остаточных классах / И. Я.Акушский, Д.И. Юдицкий. – М.: Сов. радио. – 1968. – 460 с.
5. Goh Vik Tor. Multiple error detection and correction based on redundant residue number systems. /Goh, Vik Tor, and Mohammad Umar Siddiqi // Communications, IEEE Transactions, 2008. – P.325-330.
6. Tay Thian Fatt. A new algorithm for single residue digit error correction in Redundant Residue Number System / Tay Thian Fatt, Chang Chip-Hong //Circuits and Systems (ISCAS), IEEE International Symposium IEEE, 2014. – P. 1748-1751.
7. Яцків В.В. Модифіковані коректуючі коди системи залишкових класів та їх застосування / В.В. Яцків // Інформаційні технології та комп'ютерна інженерія. – 2013. – №2. – С.39-45.
8. Hu Zhengbing. Increasing the Data Transmission Robustness in WSN Using the Modified Error Correction Codes on Residue Number System / Hu Zhengbing, V. Yatskiv, A. Sachenko // *Elektronika ir Elektrotechnika.* Vol 21, No 1 (2015). Pp. 76-81.

Відомості про авторів

Яцків Василь Васильович, к.т.н., доцент, доцент кафедри спеціалізованих комп'ютерних систем Тернопільського національного економічного університету.

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ТА ОБЧИСЛЮВАЛЬНІ МЕТОДИ

УДК 004.272

Т. Б. МАРТИНЮК, А. В. КОЖЕМ’ЯКО, Н. О. ДЕНИСЮК, Т. Ю. ПОЗДНЯКОВА

Вінницький національний технічний університет, м. Вінниця

АНАЛІЗ ОПЕРАЦІЙНОГО БАЗИСУ ДЛЯ НЕЙРОМЕРЕЖЕВИХ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ

Анотація. Розглянуто особливості реалізації комп’ютерних засобів у складі інтелектуальних систем різного призначення, зокрема за нейромережевою технологією. Обгрунтовано вибір базових обчислювальних процедур та операцій для інтелектуальних систем. Показано, що широко розповсюдженою операцією в задачах обробки, аналізу сигналів і зображень та розпізнавання образів є векторно-матричне перемноження. Ця операція реалізується шляхом формуванням парних добутків та багатоопераднього підсумовуваня. Наведено перспективні концепції реалізації операційного базису для інтелектуальних систем, з використанням різницево-зрізової обробки векторних масивів даних.

Ключові слова: інтелектуальні системи, операційний базис, обчислювальні процедури та операції.

Аннотация. Рассмотрены особенности реализации компьютерных средств в составе интеллектуальных систем различного назначения, в часности по нейросетевой технологии. Обоснован выбор базовых вычислительных процедур и операций для интеллектуальных систем. Показано, что широко распространенной операцией в задачах обработки, анализа сигналов и изображений и распознавания образов является векторно-матричное умножение. Эта операция реализируется путем формированием парных произведений и многоопераднього суммирования. Приведены перспективные концепции реализации операционного базиса для интеллектуальных систем, с использованием разносно-срезовой обработки векторных массивов данных.

Ключевые слова: интеллектуальные системы, операционный базис, вычислительные процедуры и операции.

Abstract. It is dealt with computer assets **of** intelligent systems for various purpose, in particular with neural network tecnology. It is spoken in detail basic calculation procedures and operations for intelligent systems. In problems of processing, analysis of signals and images and pattern recognition the most common operation is vector-matrix multiplication. It is realization by formation pair of products and many operant summing. The promising concept operational basis for the implementation of intelligent systems is shown.

Key words: intelligent systems, operational basis, processing procedures and operations

Вступ

Концептуальною базою для інтелектуалізації комп’ютерних систем є методи штучного інтелекту з використанням нейроінформаційних технологій [1, 2]. З іншого боку в роботі [1] визначено, що штучні нейромережі (ШНМ), в свою чергу, представляють собою окремий клас інтелектуальних систем. Причому, саме універсальні апроксимуючі властивості ШНМ можуть бути використані при створенні нового покоління інтелектуальних комп’ютерних систем [3].

Актуальність

З публікацій видно, що потреба в інтелектуальних системах різного призначення не зменшується [1-8]. Про це свідчать розробки в області створення інтелектуальних телекомунікаційних систем, інтелектуальних систем керування мобільними роботами, засобів захисту інформації в інтелектуальних системах, моделювання "нервової діяльності" в інтелектуальних системах, інтелектуальних систем ідентифікаціїї, геоінформаційної інтелектуальної системи моніторингу стану сільськогосподарських культур, системи інтелектуального керування трафіком у комп’ютерних мережах тощо [2, 5, 6, 8]. При цьому обгрунтування вибору конкретної нейротехнології в процесі розроблення інтелектуальних засобів визначається класом розв’язуваних задач [1].

Мета

Метою даної роботи є аналіз операційного базису нейромережевих інтелектуальних систем для обгрунтування вибору функціонально-повного набору обчислювальних операцій на нових принципах їх реалізації.

Постановка задачі

Відомо, що ШНМ представляють собою "величезний розподілений паралельний процесор, що складається з елементарних одиниць оброблення інформації, які накопичують експериментальні знання і передають їх для подальшого оброблення" [1,2]. І тільки властивості ШНМ, пов’язані з їх здатністю до навчання та узагальнення отриманих знань, наділяють їх рисами штучного інтелекту [1].

Отже, зважаючи на різноплановість застосування нейротехнологій, зокрема ШНМ, у вирішенні практичних задач, правомірним є використання такого позначення, як "обчислювальні засоби інтелектуальних систем", орієнтуючись саме на застосування нейроструктур. На таких же засадах застосовують поняття "обчислювальний інтелект" [7].

Серед найбільш актуальних задач, що розв’язуються інтелектуальними системами, варто відмітити такі [1,2,5,6,8]: ефективне керування дистанційними об’єктами (наприклад, керування мобільними роботами); реалізація "технічного мозку" для навчених систем керування промисловими роботами; природньо наближене моделювання функцій нервової системи при біомедичних дослідженнях; медичне експрес-діагностування; розпізнавання зорових та мовленнєвих образів;

інтелектуалізація систем ідентифікації; підвищення ефективності систем захисту інформації в комп'ютерних та телекомунікаційних мережах.

При розробці методів оброблення та аналізу даних для обчислювальних засобів інтелектуальних систем необхідно враховувати такі особливості та обмеження, що пов'язані з характером наведених вище розв'язуваних задач: задачі мають яскраво виражений природний паралелізм подання та оброблення даних; апріорна інформація про характеристики розподілу великих масивів цифрових даних відсутня або містить приблизні значення; оброблення та аналіз даних і розпізнавання образів повинні виконуватися в реальному часі, що обумовлено необхідністю оперативного прийняття рішень; пріоритетною повинна бути апаратна реалізація засобів через компактність і мобільність їх застосування.

Отже, основним напрямком покращення алгоритмічних показників багатьох процесів оброблення та аналізу сигналів і зображень, а також розпізнавання образів є збільшення рівня паралелізму при обробленні масивів даних і, зокрема векторних масивів даних, в процесі виконання базових операцій множення і багатоопераднього підсумовування.

Аналіз базових операцій обчислювальних систем

Аналіз сучасних комп'ютерних систем показав, що вони реалізують у сукупності всі базові інформаційні технології, зокрема оброблення та аналіз сигналів і зображень, а також розпізнавання образів. З іншого боку, аналіз складових найбільш розповсюджених інформаційних технологій [1, 2, 9, 10] дозволяє навести класифікаційну схему базових обчислювальних процедур та операцій (рис. 1) і визначити такі базові процедури, як фільтрація, згортка, кореляція, класифікація, паралельний пошук і нейрообчислення. Ці процедури потребують виконання таких базових операцій, як векторно-матричне перемноження, сортування, визначення екстремальних значень, ранжування, порівняння з еталоном, паралельне порівняння.

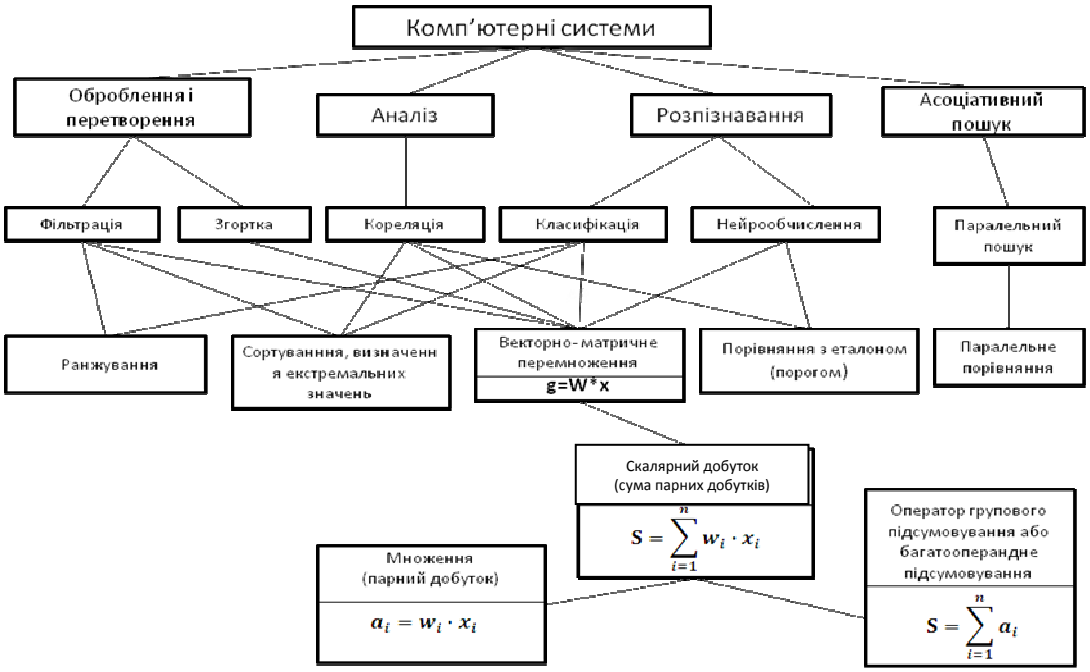


Рисунок 1 – Класифікаційна схема базових обчислювальних процедур та операцій

Разом з тим, серед особливостей нейрообчислень слід відзначити [10]:
– великий обсяг обчислень з переважанням обчислювальних операцій над логічними;
– доцільність використання векторного оброблення даних;
– використання двох видів паралелізму – просторового і часового;
– доцільність розпаралелювання та конвеєризації процесу оброблення;
– базові операції типу множення-підсумовування.

Отже, наочним є факт, що найбільшу складову серед нейрообчислень має формування добутку вектора біжучого стану мережі на матрицю міжнейронних зв'язків. Тому для прискорення цієї нейрооперації застосовують апаратні нейроакселератори [1, 2, 10], наприклад, багатопроцесорний варіант потокового нейрокомп'ютера на базі трансп'ютерів IMS T805.

Все це дозволяє серед наведених вище базових операцій визначити особливе місце операції множення матриці на вектор, яка широко використовується не тільки при моделюванні НМ, але й при

розв’язанні прикладних задач у системах оброблення й аналізу зображень, при розпізнаванні образів, у машинній графиці, для реалізації обчислень з плаваючою комою.

У свою чергу, базову операцію векторно-матричного перемноження вигляду

$$\mathbf{g} = \mathbf{W} \cdot \mathbf{x}, \tag{1}$$

можна розкласти на такі складові, як скалярний добуток або суму парних добутків вигляду

$$S = \sum_{i=1}^{n} w_i \cdot x_i, \tag{2}$$

а в подальшому відповідно на операцію множення (парний добуток) вигляду

$$a_i = w_i \cdot x_i \tag{3}$$

та оператор групового підсумовування (ОГП) вигляду

$$S = \sum_{i=1}^{n} a_i. \tag{4}$$

Про актуальність реалізації операції (1) свідчать наведені у роботі [11] результати паралельно-конвеєрної реалізації задачі множення матриці на поток векторів на реконфігурованих обчислювальних системах (РОС), тобто на ПЛІС. Так, експериментальні дослідження показали, що РОС на базі 16 ПЛІС XC4VLX40 з тактовою частотою 160 МГц мають продуктивність 21,7 ГФлопс, а на базі 16 ПЛІС XC4VLX80 з тактовою частотою 250 МГц - 70,9 ГФлопс, що у порівнянні з продуктивністю ПК Intel(R) Core(TM) 2 CPU 4400@2.00GHz складає виграш у часі відповідно у 54 рази і 176 разів.

У свою чергу, паралельні методи, алгоритми та НВІС-структури обчислення сум парних добутків (2) достатньо докладно розглянуто у роботі [10], де наведено аналітичні вирази для оцінювання основних характеристик пристроїв обчислення сум парних добутків з врахуванням кількості операндів та їх розрядності.

Разом з тим не менш актуальними є задачі ефективної реалізації операції множення (3) та ОГП (4). У роботах [12, 13] наведено переваги апаратної реалізації операції множення, зокрема на ПЛІС. У статті [13] наведено оцінки складності в базисі ПЛІС FPGA серії XC4000E для типів помножувачів з послідовною, паралельно-послідовною та табличною архітектурами. Перевагу мають помножувачі з паралельно-послідовною архітектурою для чисел великої розмірності ($n = 16$; 32). У роботі [12] обґрунтовано вибір для реалізації на ПЛІС схеми помножувача на базі алгоритму Бута і доведено, що найбільш придатним в цьому випадку для згортки часткових добутків є ієрархічне "дерево" багаторозрядних масштабуючих суматорів.

У порівнянні з множенням операція багатоопérандного підсумовування (4) має значно менше можливостей для підвищення рівня паралелізму [14]. У роботі [14] детально розглянуто і досліджено розпаралелювання реалізації ОГП (4). Найбільш прийнятним способом є почисловий паралельно-послідовний, відомий як алгоритм логарифмічного підсумовування [9, 15]. Він зорієнтований на використання відомої структури суматорів у вигляді "дерева", що дозволяє організувати виконання будь-якої операції типу "згортки" за рахунок її максимального розпаралелювання. В даному випадку довжина критичного шляху у відповідних інформаційних графах зводиться до величини порядку $t \cdot \log_2 n$, де n – розмірність масиву даних, t – час виконання елементарної операції [9]. Для цього способу характерним є виконання сукупності бінарних операцій підсумовування, що відрізняє його від порозрядного паралельно-послідовного способу, який використовує обчислення порозрядних сум (розрядних зрізів) сукупності доданків [15]. Останній відомий як спосіб багатоопérандного підсумовування, в основі якого лежить перетворення багаторядного коду в однорядний [16].

При апаратній реалізації швидкодіючих багатовхідних суматорів можливий вибір або комбінація двох варіантів: на базі цифрового сигнального процесора (ЦСП) або на ПЛІС. При використанні ЦСП DSP48E маємо тактову частоту 400 МГц і логічну ємність – 1 блок, а при використанні ПЛІС XilinxVirtex 5 маємо тактову частоту 154 МГц і логічну ємність – 44 комбінаційних логічних блоків (КЛБ, LUT) [17]. В даному випадку виграш у швидкодії досягається за рахунок ефективного трасування сигнальних ліній в рамках ЦСП.

Крім того, особливу увагу варто приділити саме багатооперандному підсумовуванню, оскільки доведено, що пара операцій (min, +) утворює напівкільце на розширеній множині дійсних чисел (включно із "∞") [18]. Пара операцій (min, +) задовольняє умовам, що визначають напівкільце, а саме [18]:

1) $\min(x,y) = \min(y,x)$;
2) $\min(\min(x,y),z) = \min(x, \min(y,z))$;
3) $x + y = y + x$;
4) $(x + y) + z = x + (y + z)$;
5) $x + \min(y,z) = \min(x + y, x + z)$;
6) множина $RU\{\infty\}$ містить "∞", для якого при будь-якому x виконуються рівності $\min(x,\infty)=x$, $x + \infty = x$;
7) множина $RU\{\infty\}$ містить 0, яке при будь-якому x задовольняє рівність $x + 0 = x$.

Таким чином, подальше дослідження з метою підвищення ефективності виконання базових процедур для інтелектуальних систем базується на розробленні та вдосконаленні базових обчислювальних операцій на нових принципах організації такої операції, як багатооперандне підсумовування. З цією метою серед перспективних концепцій розвитку інформаційних технологій оброблення і аналізу даних та розпізнавання образів в інтелектуальних системах необхідно виділити застосування принципів око-процесорного оброблення цифрових даних, що дозволить [3, 19]: досягти глибокого розпаралелювання при обробленні одно- і двовимірних масивів даних; сумістити виконання базових нейроопераций (підсумовування зважених вхідних сигналів і порогове порівняння); розширити функціональні можливості оброблення елементів одно- і двовимірних масивів даних, наприклад, за рахунок їх сортування та ранжування, суміщених зі згорткою елементів масиву.

Висновки

1. На основі здійсненого аналізу операційного базису для інтелектуальних систем відзначено, що найбільш розповсюдженою операцією в задачах оброблення, аналізу сигналів та зображень і розпізнавання образів є операція векторно-матричного перемноження, яка, в свою чергу, реалізується в процесі формування парних добутків та їх групового підсумовування (багатооперандного підсумовування)
2. Шляхом обґрунтування базових операцій для комп'ютерних засобів інтелектуальних систем показано обмеженість в алгоритмічно-стуктурному аспекті класичного варіанта розпаралелювання операції групового підсумовування.
3. При проведенні аналізу визначено актуальну задачу розроблення такого методу багатооперандного підсумовування, який би забезпечив не тільки високий рівень паралелізму, але й можливість суміщення при виконанні декількох операцій та отримання проміжних результатів.

Список літератури

1. Осовский С. Нейронные сети для обработки информации / С. Осовский; пер. с польск. И.Д. Рудинского. – М.: Финансы и статистика, 2004. – 344 с. – ISBN 5-279-02567-4.
2. Комашинский В.И. Нейронные сети и их применение в системах управления и связи / В.И. Комашинский, Д.А. Смирнов. – М.: Горячая линия-Телеком, 2003. – 94 с. – ISBN 5-93517-094-9.
3. Кожем'яко В.П. Наукова концепція образного відео-комп'ютера око-процесорного типу в контексті сучасної методології штучного інтелекту / В.П. Кожем'яко, А.А. Яровий // Оптико-електронні інформаційно-енергетичні технології. – 2001. – № 2. – С. 84-89. – ISSN 1681-7893.
4. Палагин А.В. Системная интеграция средств компьютерной техники / А.В. Палагин, Ю.С. Яковлев. – Вінниця: УНІВЕРСУМ-Вінниця, 2005. – 680 с. – ISBN 966-641-140-7.
5. Гаврилов А.В. Архитектура гибридной интеллектуальной системы управления мобильного робота / А.В. Гаврилов, В.В. Губарев, К. Х. Джо, Х. Х. Ли // Вестник Новосибирского ГТУ. – 2004. – № 2. – С. 3-13.
6. Гильгурт С.Я. Программно-аппаратная защита данных в распределённых интеллектуальных системах / С. Я. Гильгурт, А.К. Гиранова // Искусственный интеллект. – 2010. – № 3. – С. 706-711. – ISSN 1561-5359.
7. Rutkowski L. Computational Intelligence. Methods and Techniques. Berlin-Heidelberg: Springer-Verlag, 2008, 514 p.
8. Буков А.А.Технические нервные системы. Обучаемые системы управления со зрением для промышленных роботов / А.А. Буков. – Липецк: Изд-во Липецк. гос. техн. ун-та, 2001. – 223 с.
9. Прэтт У. Цифровая обработка изображений / У. Прэтт; пер. с англ. – В 2-х кн. – М.: Мир, 1982. – 792 с.
10. Цмоць І.Г. Інформаційні технології та спеціалізовані засоби обробки сигналів і зображень у реальному часі / І.Г. Цмоць. – Львів: Видавництво УАД, 2005. – 228 с. – ISBN 966-322-024-4.

11. Трунов Г.Л. Параллельно-конвейерная реализация задач и умноження матрицы на поток векторов на реконфигурируемых вычислительных системах / Г.Л. Трунов, А.Г. Коваленко // Искусственный интеллект. – 2008. – № 3. – С.742-749. – ISSN 1561-5359.

12. Стешенко В.Б.ПЛИС фирмы Altera: элементная база, система проектирования и языки описания аппаратуры / В.Б. Стешенко. – М.: Издательский дом "Додэка-XXI", 2002. – 576 с. – ISBN 5-94120-033-1.

13. Особенности реализации операции умножения на ПЛИС / МоЧжо Чо // Современные наукоемкие технологии. – 2008. – № 4 [Електронний ресурс] – Режим доступу: http://www.rae.ru/fs/?section=content&op=show_article&article_id=7791584.

14. Справочник по цифровой вычислительной технике; под ред. В.Н. Малиновского. – К.:Техника, 1980. – 320 с.

15. Мартинюк Т.Б. Рекурсивні алгоритми багатоопераднної обробки інформації / Т.Б. Мартинюк. – Вінниця: УНІВЕРСУМ-Вінниця, 2000. – 216 с. – ISBN 966-7199-98-3.

16. Гамаюн В.П. Способ ускоренного преобразования многорядного кода в однорядный / В.П. Гамаюн // Управляющие системы и машины. – 1995. – № 4/5. – С. 10-14. – ISSN 0130-5395.

17. Зианбетов Э.И. Исследование специализированных структур цифровой обработки сигналов в составе ПЛИС серии Virtex / Э.И. Зианбетов, В.П. Малахов, В.С. Ситников // Современные информационные и электронные технологии: Девятая междунар. науч.-практ. конф., 19-23 мая 2008 г.: труды. – Одесса, 2008. – С. 133.

18. Шлезингер М. Десять лекций по статистическому и структурному распознаванию / М. Шлезингер, В. Главач. – К.: Наук. думка, 2004. – 536 с. – ISBN 966-00-0341-2.

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

Відомості про авторів

Мартинюк Тетяна Борисівна – д.т.н., доцент, професор кафедри лазерної та оптикоелектронної техніки, Вінницький національний технічний університет.

Кожем’яко Андрій Вікторович - к.т.н., доцент, доцент кафедри лазерної та оптикоелектронної техніки, Вінницький національний технічний університет.

Денисюк Наталія Олексіївна – здобувач кафедри лазерної та оптикоелектронної техніки, Вінницький національний технічний університет.

Позднякова Тетяна Юріївна - студентка магістартури кафедри лазерної та оптикоелектронної техніки, Вінницький національний технічний університет.