



ISSN 2788-5518

ІНФОКОМУНІКАЦІЙНІ ТА КОМП'ЮТЕРНІ ТЕХНОЛОГІЇ

№ 1 (05) 2023

НАУКОВИЙ ЖУРНАЛ

**Київ
Університет «Україна»**

2023



ISSN 2788-5518

INFOCOMMUNICATION AND COMPUTER TECHNOLOGIES

№ 1 (05) 2023

SCIENTIFIC JOURNAL

**Kyiv
University «Ukraine»**

2023

ПРЕЗИДЕНТУ УНІВЕРСИТЕТУ «УКРАЇНА» ПРОФЕСОРУ ТАЛАНЧУКУ П.М. 85!

DOI 10.36994 / 2788-5518-2023-01-05-01



Першого липня святкував свій 85 річний ювілей Президент Відкритого міжнародного університету розвитку людини «Україна» Петро Таланчук — постать, знана не лише в Україні, а й далеко за її межами.

Ми вітаємо незвичайну людину з надзвичайними досягненнями. Справжнього Українця! Головне кредо Петра

Михайловича Таланчука – утвердження державності та українського духу у світі. Підкреслюємо свідому національно-патріотичну позицію освітнього діяча, яку він утілює все своє життя в роботі та у своїх наукових і публіцистичних працях.

Обдарований вчений, політик, талановитий організатор та керівник народився 1 липня 1938 року в селі Городище-Косівське Володарського району на Київщині. З самого дитинства він був палкий, працьовитий, відмінником в школі. Після військової служби пройшов конкурс до Військово-морського політичного училища імені Жданова в Ленінграді (нині Санкт-Петербург), а через рік був переведений до Вищого військово-морського училища імені Фрунзе. У 1960-му вступив на механіко-машинобудівний факультет Київського політехнічного інституту, який з відзнакою закінчив 1965 року. Педагогічну і наукову діяльність розпочав асистентом на кафедрі, захистив кандидатську дисертацію, в якій досліджував динамічні характеристики приладів, і в подальшому розширив цю тему в докторській дисертації.

Петро Таланчук на демократичній основі був обраний у 1987 році ректором Київському політехнічному інституті, де до цього обіймав посаду проректора з навчально-виховної роботи, а під час горбачовської «перебудови» і демократизації. Патріот України та відчайдушна людина, тодішній ректор КПІ, спокушаючи долю надав у 1989 році приміщення для проведення I-го з'їзду Народного Руху України. На з'їзді якому було затверджено Програму Народного Руху України за незалежність України

та її перебудову. Як відмічав у своєму інтерв'ю Перший голова Народного Руху України Іван Драч : «На той час... згода на це була рівноправна героїському вчинкові...».

Політична діяльність ювіляра розпочалася у 1989-му як депутата Верховної Ради СРСР. Як вважає Петро Таланчук «Життя, щоб досягати високості... Інші варіанти — для слабких духом». Тому він у найважчі для молоді Української держави часи у 1992 році ризикнув очолити одну з найбільш зруйнованих галузей — став першим міністром освіти незалежної України і свій трудовий шлях у міністерстві розпочав з реформи у освіт. Перший крок задля об'єднання освітян, за ініціативи міністра освіти, проведено перший З'їзду працівників освіти, з якими обговорили та з правками прийняли Державну національну програму «Освіта» (Україна ХХІ століття). Саме ця програма першою відкрито й перспективно визначила «патріотичну національну ідею, заклала формування особистості в контексті української державності, чим започаткувала перехід від колишньої «радянської» освіти до розробки дієвих механізмів реформування освіти. Зокрема для збереження її потенціалу та обсягу підготовки фахівців, посилення державної підтримки пріоритетних напрямів освіти і науки, приведення у відповідність із найновішими світовими досягненнями сучасної науки та з урахування змін в економіці. Вона визначила спрямованість стратегії розвитку освіти в перші десятиріччя розвитку української незалежної держави, яку можна звести принаймні до трьох основних завдань: зміцнення національної системи освіти; її адаптація до ринкових і демократичних перетворень; повномірне входження освіти України в європейський освітній простір. Але, на жаль, реалізація цієї програми не була доведена до кінця. Проте це не зупинило його. Зважаючи на обставини, він відшукав інший шлях її впровадження — створив і очолив вищий навчальний заклад нового покоління — Університет «Україна» заснований як безаналоговий для України інклюзивний заклад відкритого типу, який відомий не лише в нашій країні, а й за її межами. Концепція інклюзивної освіти була реалізована вперше в Україні саме в Університеті «Україна». За 25 років своєї діяльності університету було підготовлено майже 195 тис. фахівців, із них осіб з інвалідністю понад 10 тисяч.

Від народження Петру Михайловичу Богом даний талант не тільки до самовідданої праці, а й до постійного пошуку ідей: інженер-новатор, учений-дослідник, автор 60 винаходів та патентів, має понад 480 друкованих праць, серед яких – 18 монографій, підручників і навчальних посібників. Дійсний член Міжнародної академії наук вищої школи, член Національної спілки журналістів України. Є почесним ректором

Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського», почесним членом Всесвітньої інноваційної фундації та чотирьох міжнародних академій, член Міжнародної Асоціації Президентів Університетів (IAUP). Удостоєний ордена «За заслуги» III ступеня, ордена «Дружби народів», ордена ім. Князя Ярослава Мудрого V ступеня. Нагороджений Хрестом Пошани «За духовне відродження», медаллю «Григорій Сковорода», ювілейною медаллю «25 років незалежності України», Почесною грамотою Верховної Ради України, має знак «Відмінник освіти України», знак «За наукові та освітні досягнення», знак «Ушинський К. Д.», заслужений діяч науки і техніки України. За подвижництво в організації діяльності мережі територіально відокремлених навчальних підрозділів Університету в регіонах України, упровадження в освітній процес методик дистанційного та інклюзивного навчання Петра Михайловича Таланчука визнано кращим освітянином року та лауреатом відзнаки «Лиш Храм збудуй».

Петро Михайлович 01.07.23 року був нагороджений орденом святого Архистратига Михаїла I ступеня за заслуги перед Українською Православною Церквою та побожним народом.

Колектив бажає Петру Таланчуку міцного здоров'я, злагоди, добра, невичерпної енергії, нових плідних здобутків на ниві освіти і науки в державі, щоб ви добудували омріяний освітянський комплекс та Україна здобула Перемогу над ворогами-загарбниками, а справедливість узяла гору над невіглаством».

Засновник, видавець та виробник журналу: «Відкритий міжнародний університет розвитку людини «Україна».

Свідоцтво про державну реєстрацію:
Серія KB № 24782-14722 Р від 06.04.2021 р.

Журнал включений до переліку наукових фахових видань ВАК України- категорія «Б», додаток 4 до наказу МОН України № 735 від 29.06. 2021.

Журнал відповідає вимогам наказу МОН України №32 від 15.01.2018 р.

Має міжнародні коди: ISSN 2788-5518 та DOI 10.36994/2788-5518.

Журнал індексується у міжнародних базах та каталогах Google Scholar, Index Copernicus, Vernadsky National Library of Ukraine.

Мова видання: українська, англійська

© Університет «Україна», 2023.

Рекомендовано до друку вченою радою відкритого міжнародного університету розвитку людини «Україна» (протокол №6 від 30.06.2023 р.)

URL-адреса видання: visn-icct.uu.edu.ua/

Тематика журналу:

- Теоретичні основи інфокомунікацій та комп'ютерних наук
- Інфокомунікаційні системи
- Безпроводові технології
- Технічна електродинаміка та антени
- Волоконно-оптичні системи
- Радіорелейні та супутникові системи
- Телевізійні системи
- Системи відеоконференцзв'язу та відеоспостереження
- Комп'ютерні системи та мережі
- Комп'ютерна і програмна інженерія
- Системи SCADA
- Захист інформації та кібербезпека
- Прикладна математика та моделювання
- Метрологія та інформаційно-вимірювальні системи
- Медична та екологічна електроніка
- Автоматизація управління технологічними процесам

В журналі публікуються статті для захисту дисертацій зі спеціальностей :121 Інженерія програмного забезпечення. 122 Комп'ютерні науки та інформаційні технології. 123 Комп'ютерна інженерія. 151 Автоматизація та комп'ютерні інтегровані технології. 152 Метрологія та інформаційно-вимірювальна техніка. 172 Телекомунікації та радіотехніка

Редколегія

Головний редактор: Таланчук Петро Михайлович, д. т. н., проф., університет «Україна», ORCID 0000-0001-8748-6127, Київ, Україна. talanshuk_pm@ukr.net

Перший заступник головного редактора: Семенко Анатолій Іларіонович, д.т.н, проф., університет «Україна», ORCID 0000-0002-7043-7801, Київ, Україна. setel@ukr.net

Заступники головного редактора: Забара Станіслав Сергійович, д.т.н., проф., університет «Україна», ORCID 0000-000-9554-7575, Київ, Україна. staszabara37@gmail.com

Відповідальний секретар: Павленко Володимир Іванович, к.ф.-м.н., доц., університет «Україна», ORCID 0000-0002-3958-0415, Київ, Україна. pavlenko.v@i.ua

Члени редколегії

Безрук Валерій Михайлович, д.т.н., проф., Харківський Національний університет радіоелектроніки, ORCID 0000-0003-2349-7788, Харків, Україна. valeriy_bezruk@ukr.net

Блаунштейн Натан Шаєвич, д.ф.-м. н., проф., університет Бен-Гуріона, ORCID 0000-0003-2945-9379, Бір-Шева, Ізраїль. nathan.blaunstein@hotmail.com

Бойко Юлій Миколайович, д.т.н., проф., Khmel'nitsky national University, Khmel'nitsk, Ukraine. ORCID 000-0003-0603-7827, Хмельницький, Україна. boiko_julius@ukr.net

Бескровний Олексій Іванович, к.т.н, доц., Університет «Україна», ORCID 0000-0002-7043-7801, Київ, Україна. obezkrovnyy@gmail.com

Гепко Ігор Олександрович, д.т.н., проф., Український державний центр радіочастот, Х, Київ, Україна. gerko@ucrf.gov.ua

Климаш Михайло Миколайович, д.т.н., проф., Національний університет «Львівська політехніка», ORCID 0000-0003-2867-1482, Львів, Україна. mklimash@polynet.lviv.ua

Козловський Валерій Валерійович, д.т.н., проф., Національний авіаційний університет, ORCID 0000-0002-8301-5501, Київ, Україна. vv_k@nau.edu.ua

Кушнір Микола Ярославович, к.ф.-м.н, доц., Чернівецький Національний університет ім. Юрія Федьковича, ORCID 0000-0001-9480-3856, Чернівці, Україна. kushnirnick@gmail.com

Лунтовський Андрій Олегович, д.т.н., проф., Державна академія Саксонії «Беруфсакадемія», ORCID 0000-0001-7038-6955, Дрезден, Німеччина. andriy.luntovskyy@ba-sachsen.de

Мельник Ігор Віталійович, д.т. н., проф., НТУУ «Київський політехнічний інститут ім. Ігоря Сікорського», ORCID 0000-0003-0220-0615, Київ, Україна. imelnik@phbme.kpi.ua

Наконечний Олександр Григорович, д.ф.-м.н., проф., Київський Національний університет ім. Тараса Шевченка, ORCID 3070, Київ, Україна. a.nakonechniy@gmail.com

Наритник Теодор Миколайович, к.т.н., проф, НТУУ «Київський політехнічний інститут ім. Ігоря Сікорського», ORCID 0000-0001- 8071-6356, Київ, Україна. t.narytnyk@ukr.net

Петренко Анатолій Іванович, д.т.н., проф., НТУУ «Київський політехнічний інститут ім. Ігоря Сікорського», ORCID 0000-0001- 6712-7792, Київ, Україна. tolja.petrenko@gmail.com

Попов Валентин Іванович, д.ф-м.н., проф., Ризький технічний університет, ORCID 0000-0002-2040-9319. Рига, Латвія. popovs@latnet.lv

Почерняєв Віталій Миколайович, д.т.н., проф., Державний університет інтелектуальних технологій і зв'язку, ORCID 0000 0001 7130 8668, Одеса Україна. vpochernyayev@gmail.com

Рогоза Валерій Станіславович, д.т.н., проф., НТУУ «Київський політехнічний інститут ім. Ігоря Сікорського», ORCID 0000-0003-2327-156X, Київ, Україна. rosvetnik@gmail.com

Самарай Валерій Петрович, к.т.н., с. н. п., університет «Україна», ORCID 0000-0003- 4419-1366, Київ, Україна. samaraj@ukr.net

Сидоренко Анатолій Сергійович, д.ф-м.н., проф., академік Академії наук Молдови, ORCID 0000-0001-7433-4140, Кишинів, республіка Молдова. anatoli.sidorenko@kit.edu

Стрелковська Ірина Вікторівна, д.т.н., проф., міжнародний гуманітарний університет, ORCID 0000-0002-1813-0554, Одеса, Україна. irina7000370@gmail.com

Тимошенко Анатолій Григорович, к.т.н, доц., університет «Україна», ORCID 0000-0003-0954-3186, Київ, Україна. timoshag@i.ua

Ящишін Євген Михайлович, д.т.н., проф., Варшавський технологічний університет, ORCID 0000-0001-8378-1399, Варшава, Польща. e.jaszczyszyn@ire.pw.edu.pl

Founder, publisher and producer of the journal: "Open International University of Human Development" Ukraine".

Certificate of state registration:

Series KV № 24782-14722 R from 06.04.2021.

The journal is included in the list of scientific professional publications of the Higher Attestation Commission of Ukraine - category "B", Annex 4 to the order of the Ministry of Education and Science of Ukraine № 735 from 29.06. 2021.

The journal meets the requirements of the order of the Ministry of Education and Science of Ukraine №32 from 15.01.2018.

It has international codes: ISSN 2788-5518 and DOI 10.36994 / 2788-5518.

Publication URL: visn-icct.uu.edu.ua/

Subject journal:

- ☐ **Theoretical foundations of infocommunications and computer science**
- ☐ **Infocommunication systems**
- ☐ **Wireless technology**
- ☐ **Technical electrodynamics and antennas**
- ☐ **Fiber optic systems**
- ☐ **Radio relay and satellite systems**
- ☐ **Television systems**
- ☐ **Video conferencing and video surveillance systems**
- ☐ **Computer systems and networks**
- ☐ **Computer and software engineering**
- ☐ **SCADA systems**
- ☐ **Information security and cybersecurity**
- ☐ **Applied mathematics and modeling**
- ☐ **Metrology and information-measuring systems**
- ☐ **Medical and environmental electronics**
- ☐ **Automation of process control**

The journal publishes articles for the defense of dissertations in the following specialties: 121 Software engineering. 122 Computer science and information technologies. 123 Computer engineering. 151 Automation and computer integrated technologies. 152 Metrology and information and measurement technology. 172 Telecommunications and radio engineering

The journal is indexed in international databases and directories of Google Scholar, Index Copernicus, Vernadsky National Library of Ukraine

Language of publication: Ukrainian, English

© University "Ukraine", 2023.

Recommended for publication by the Academic Council of the Open International University of Human Development "Ukraine" (Protocol №6 dated 30.06.2023.)

Publication URL: visn-icct.uu.edu.ua/

Editorial board

Chief Editor: **Peter Talanchuk**, Dr. habil., Prof., University "Ukraine", ORCID 0000-0001-8748-6127, Kyiv, Ukraine. *talanshuk_pm@ukr.net*

First deputy editor-in-chief: **Anatoliy Semenko**, Dr. habil., Prof., University "Ukraine", ORCID 0000-0002-7043-7801, Kyiv, Ukraine. *setel@ukr.net*

Deputy Chief Editor: **Stanislav Zabara**, Dr. habil., Prof., University "Ukraine", ORCID 0000-000-9554-7575, Kyiv, Ukraine. *staszabara37@gmail.com*

Executive secretary: **Vladimir Pavlenko**, Ph.D., Ass.Prof. University "Ukraine", ORCID 0000-0002-3958-0415, Kyiv, Ukraine. *pavlenko.v@i.ua*

Members of the Editorial Board

Valeriy Bezruk, Dr. habil., Prof., Kharkiv national University of Radioelectronics, ORCID 0000-0003-2349-7788, Kharkiv, Ukraine. *valeriy_bezruk@ukr.net*

Natan Blaunstein, Dr. habil.Phys., Prof., Ben-Gurion University, ORCID 0000-0003-2945-9379, Beer Sheva, Israel. *nathan.blaunstein@hotmail.com*

Julius Boiko, Dr. habil., Prof., Khmelnytsky national University, ORCID 0000-0003-0603-7827, Khmelnytskyi, Ukraine. *boiko_julius@ukr.net*

Alexey Beskrovnyy, Ph.D., Ass.Prof., University "Ukraine", ORCID 0000-0002-7043-7801, Kyiv, Ukraine. *obezkrovnyy@gmail.com*

Igor Gepko, Dr. habil., Prof., Ukrainian State Center of Radio Frequencies, ORCID 0000-0002-4138-021X, Kyiv, Ukraine. *gepko@ucrf.gov.ua*

Mykhailo Klymash, Dr. habil., Prof., Lviv polytechnic national University, ORCID 0000-0003-2867-1482, Lviv, Ukraine. *mklmash@polynet.lviv.ua*

Valeriy Kozlovsky, Dr. habil., Prof., National Aviation University, ORCID-0000-0002-8301-5501, Kyiv, Ukraine. *vv_k@nau.edu.ua*

Mykola Kushnir, Ph.D., Phys., Ass.Prof., Yuriy Fedkovych Chernivtsi national University, ORCID 0000-0001-9480-3856, Chernivtsi, Ukraine. *kushnirnick@gmail.com*

Andriy Luntovsky, Dr. habil., Prof., Saxon Study Academy "Berufsakademie", ORCID 0000-0001-7038-6955, Dresden, Germany. *andriy.luntovskyy@ba-sachsen.de*

Igor Melnyk, Dr. habil., Prof., National technical University of Ukraine "Igor Sikorsky Kyiv polytechnic Institute", ORCID 0000-0003-0220-0615, Kyiv, Ukraine. *imelnik@phbme.kpi.ua*

Alexander Nakonechny, Dr. habil. Phys., Prof., Taras Shevchenko national University of Kyiv, ORCID 0000-0002-8705-3070, Kyiv, Ukraine. *a.nakonechniy@gmail.com*

Teodor Narytnyk, Ph.D. Prof. National technical University of Ukraine "Igor Sikorsky Kyiv polytechnic Institute", », ORCID 0000-0001- 8071-6356, Kyiv, Ukraine. *t.narytnyk@ukr.net*

Anatoliy Petrenko, Dr. habil., Prof., National technical University of Ukraine "Igor Sikorsky Kyiv polytechnic Institute", ORCID 0000-0001- 6712-7792, Kyiv, Ukraine. *tolja.petrenko@gmail.com*

Valentin Popov, Dr. habil. Phys., Prof., Riga technical University, ORCID 0000-0002-2040-9319. Riga, Latvia. *popovs@latnet.lv*

Vitaliy Pochernyaev, Dr. habil., Prof., University of Intellectual Technologies and Communications, ORCID 0000-0001-7130-8668, Odesa, Ukraine. *vpochernyaev@gmail.com*

Valeriy Rogoza, Dr. habil., Prof., National technical University of Ukraine "Igor Sikorsky Kyiv polytechnic Institute", ORCID 0000-0003-2327-156X Kyiv, Ukraine. *rosvetnik@gmail.com*

Valeriy Samaraj, Ph.D., SSR., University "Ukraine", ORCID 0000-0003-4419-1366, Kyiv, Ukraine. *samaraj@ukr.net*

Anatoliy Sidorenko, Dr. habil. Phys., Prof., Academician of AS of Republic of Moldova, ORCID 0000-0001-7433-4140, Kishinev, Republic of Moldova. *anatoli.sidorenko@kit.edu*

Irina Strelkovskaya, Dr. habil., Prof., International Humanitarian University, ORCID 0000-0002-1813-0554, Odessa, Ukraine. *irina7000370@gmail.com*

Anatoliy Tymoshenko, Ph.D., Ass. Prof., University "Ukraine", ORCID 0000-0003-0954-3186, Kyiv, Ukraine. *timoshag@i.ua*

Eugeniy Yashchyshyn, Dr. habil., Prof., Warsaw University of Technology, ORCID 0000-0001-8378-1399, Warsaw, Poland. *e.jaszczyszyn@ire.pw.edu.pl*

ЗМІСТ

ПЕЗИДЕНТУ УНІВЕРСИТЕТУ «УКРАЇНА» ПРОФЕСОРУ ТАЛАНЧУКУ П.М. 85!	3
ДО 100-РІЧЧЯ ПРОФЕСОРА ЦИБІЗОВА К.М. НАРИС ЖИТТЯ ТА НАУКОВОЇ ДІЯЛЬНОСТІ ДОКТОРА ТЕХНІЧНИХ НАУК ПРОФЕСОРА КОСТЯНТИНА МИКОЛАЙОВИЧА ЦИБІЗОВА	16
НАУКОВІ ТВОРИ ВДЯЧНИХ УЧНІВ ПРОФЕСОРА ЦИБІЗОВА К.М. ПЕРЕДМОВА	20
Почерняєв В.М., Магомедова М.С. ПОЛЯРИЗАТОР НА СЕКТОРНОМУ ХВИЛЕВОДІ, ЧАСТКОВО ЗАПОВНЕНОМУ ДІЕЛЕКТРИКОМ	21
Почерняєв В.М., Сивкова Н.М. ЕЛЕКТРОМАГНІТНЕ ПОЛЕ БОЧКОПОДІБНОГО ПОЛОГО РЕЗОНАТОРА	28
Почерняєв В.М., Сушанова В.С. МОДЕЛЮВАННЯ БАГАТОПРОМЕНЕВОГО КАНАЛУ ЗВ'ЯЗКУ	35
Наритник Т.М., Кузьмичов І.К., Авдєєнко Г.Л., Музичишин Б.І., Май О.В., Алісов А.О. СУМАТОР ПОТУЖНОСТЕЙ НА БАЗІ КВАЗІОПТИЧНОГО ВІДКРИТОГО РЕЗОНАТОРА ЯК ВИСОКОСТАБІЛЬНИЙ ГЕНЕРАТОР ТЕРАГЕРЦЕВОГО ДІАПАЗОНУ	43
Сайко В.Г., Ільїнов М.Д., Нестеренко І.К. МЕТОД СИНТЕЗУ ВІБРАТОРНИХ АНТЕН З РЕАКТИВНИМИ ВСТАВКАМИ	55

ІНФОКОМУНІКАЦІЙНІ ТЕХНОЛОГІЇ

Гайдаманчук В.А., Матвієнко М.В., Пастушенко І.М., Семенко А.І. РЕГІОНАЛЬНА НАВІГАЦІЙНА СУПУТНИКОВА СИСТЕМА ЯК ДЖЕРЕЛО PNT ПІД ЧАС НАВМИСНИХ ТРИВАЛИХ ЗАГРОЗ СИГНАЛАМ GNSS	66
Панченко Б.Є., Буката Л.М., Багачук Д.Г. МАТЕМАТИЧНА МОДЕЛЬ ВЗАЄМОДІЇ СТАЦІОНАРНИХ SH-ХВИЛЬ З СИСТЕМОЮ КРИВОЛІНІЙНИХ ТРІЩИН У НАПІВПРОСТОРІ	73
Уривський Л.О., Косогор А.В. КОМПЛЕКСНИЙ АНАЛІЗ ВЛАСТИВОСТЕЙ КОДУ ХЕММІНГА В СУПЕРПОЗИЦІЇ ТЕОРІЙ КОДУВАННЯ, ЗАВАДОСТІЙКОСТІ ТА ІНФОРМАЦІЇ	85
Тихонов В.А., Безрук В.М. ДЕКОМПОЗИЦІЯ МОДЕЛІ НЕСТАЦІОНАРНОГО ВИПАДКОВОГО ПРОЦЕСУ	94
Єзерський Н.В., Мовчанюк А.В. ОПТИМІЗАЦІЯ АНАЛОГОВОГО МОДУЛЮ ЗА ШУМОВИМИ ХАРАКТЕРИСТИКАМИ У БОРТОВОМУ ДЕТЕКТОРІ РЕЄСТРАТОРА- АНАЛІЗАТОРА ДЛЯ ІДЕНТИФІКАЦІЇ ЗАРЯДЖЕНИХ ЧАСТИНОК	101

КОМП'ЮТЕРНІ ТЕХНОЛОГІЇ

Корчинський В.В., Стайкуца С. В., Виноградов І.В., Швець О. В., Бєлова Ю. В. МЕТОДИ ПІДВИЩЕННЯ ЯКОСТІ ПЕРЕТВОРЕННЯ МОВИ НА ТЕКСТ В СИСТЕМАХ БІОМЕТРИЧНОЇ АУТЕНТИФІКАЦІЇ	114
--	-----

Ткачук А.В., Булах Б.В. ОПИС ЗНАНЬ ПРО ВИХІДНИЙ КОД З ВИКОРИСТАННЯМ ОНТОЛОГІЇ.....	123
Одегов М.А., Гаджисев М.М., Буката Л.М., Глазунова М.В., Кочеткова М.В. ПОРІВНЯННЯ АЛГОРИТМІВ КЛАСИФІКАЦІЇ BIG DATA МЕТОДАМИ ІМІТАЦІЙНОГО МОДЕЛЮВАННЯ.....	134
Одегов М.А., Гаджисев М.М., Буката Л.М., Глазунова М.В., Кочеткова М.В. ОБҐРУНТУВАННЯ ШВИДКИХ АЛГОРИТМІВ КЛАСИФІКАЦІЇ НА МНОЖИНАХ BIG DATA ЗА КРИТЕРІЯМИ НАДІЙНОСТІ І ПРОДУКТИВНОСТІ.....	148
Корнага Я.І., Барабаш А.О. МАТЕМАТИЧНА МОДЕЛЬ МОНІТОРИНГУ ПОДІЙ В СИСТЕМІ УПРАВЛІННЯ РОЗПОДІЛЕНИМИ БАЗАМИ ДАНИХ НА ОСНОВІ СЕНСОРІВ ПОДІЙ.....	161
Середа А.В., Самарай В. П., Даценко І. П., Павленко В. І., Бровченко Є. М., Таланчук П.М. РОЗВИТОК МОБІЛЬНОЇ ОСВІТИ ДЛЯ ПОКРАЩЕННЯ ЯКОСТІ НАВЧАННЯ.....	167
Бутенко М.Є., Павленко В. І. ПЕРЕДБАЧЕННЯ ВАРТОСТІ ЗЕМЛІ СІЛЬСЬКОГОСПОДАРСЬКОГО ПРИЗНАЧЕННЯ ЗА ДОПОМОГОЮ АЛГОРИТМІВ МАШИННОГО НАВЧАННЯ	175
Вахранов І.А., Булах Б.В. КОМБІНОВАНІ ПІДХОДИ ДО СТАТИЧНОГО АНАЛІЗУ КОДУ З ВИКОРИСТАННЯМ НЕЙРОННИХ МЕРЕЖ	183
Бровченко Є. М., Самарай В. П., Даценко І. П., Павленко В. І., Середа А.В. ЗАХИСТ НЕСТРУКТУРОВАНОЇ ІНФОРМАЦІЇ НА МОБІЛЬНОМУ ПРИСТРОЮ	194
Левченко В.В., Павленко В.І., Таланчук П.М. ХВИЛІ ЗСУВУ В ШАРУВАТО–ПЕРІОДИЧНІЙ СТРУКТУРІ «П'ЄЗОЕЛЕКТРИК - ЩІЛИНА»	201
ПРАВИЛА ОФОРМЛЕННЯ СТАТЕЙ	205

CONTENT

TO THE PRESIDENT OF "UKRAINE" UNIVERSITY TO PROFESSOR TALANCHUK P.M. 85!	3
TO THE 100TH ANNIVERSARY OF PROFESSOR K.M. CYBIZOV ESSAY OF THE LIFE AND SCIENTIFIC ACTIVITY OF THE DOCTOR OF TECHNICAL SCIENCES PROFESSOR KOSTYANTYN MYKOLAYOVYCH CYBIZOVA	16
WORKS OF GRATEFUL STUDENTS OF PROFESSOR CYBIZOV K.M PREFACE.....	20
Vitaly Pochernyaev, Mariia Mahomedova. POLARIZER ON A SECTOR WAVEGUIDE PARTIALLY FILLED BY DIELECTRIC	21
Vitaly Pochernyaev, Nataliia Syvkova. ELECTROMAGNETIC FIELD OF A BARREL-SHAPED HOLLOW RESONATOR	28
Vitaly Pochernyaev, Viktoriia Sushanova. SIMULATION OF A MULTIPATH COMMUNICATION CHANNEL	35
Teodor Narytnyk, Igor Kuzmichev. Bohdan Muzychishin. Volodymyr Saiko, Mykhailo Ilyinov, Igor Nesterenko. RESEARCH OF A QUASIOPTICAL OPEN RESONATOR FOR POWER SUMMATION IN A HIGHLY STABLE GENERATOR OF THE TERAHERTZ RANGE	43
Volodymyr Saiko, Mykhailo Ilyinov, Igor Nesterenko. SYNTHESIS METHOD OF VIBRATOR ANTENNAS WITH REACTIVE INSERTS	55

INFOCOMMUNICATION TECHNOLOGIES

Viktor A. Gaidamanchuk, Mykola V. Matvienko, Ihor M. Pastushenko, Anatoliy Semenko. REGIONAL NAVIGATION SATELLITE SYSTEM AS PNT SOURCE DURING INTENTIONAL LONG-TERM THREATS TO GNSS SIGNALS	66
Borys Panchenko. Liudmyla Bukata. Denys Bahachuk. MATHEMATICAL MODEL OF THE INTERACTION OF STATIONARY SH-WAVES WITH A SYSTEM OF CURVILINEAR CRACKS IN A HALF-SPACE	73
Leonid Uryvsky, Anastasia Kosogor. COMPREHENSIVE ANALYSIS OF THE HAMMING CODE PROPERTIES BASED ON SUPERPOSITION OF THE CODING, NOISE IMMUNITY AND INFORMATION THEORIES	85
Vyacheslav Tikhonov, Valeriy Bezruk. DECOMPOSITION OF THE NON-STATIONARY RANDOM PROCESS MODEL	94
Nikita Yezerskyi, Andriy Movchanuk. OPTIMIZATION OF THE ANALOG MODULE ACCORDING TO NOISE CHARACTERISTICS IN THE ON-BOARD DETECTOR OF THE RECORDER-ANALYZER FOR THE IDENTIFICATION OF CHARGED PARTICLES	101

COMPUTER TECHNOLOGIES

Korchynskyi V.V., Staikuca S.V., Vynogradov I.V., Shvets O.V., Bielova I.V. IMPROVING THE QUALITY OF SPEECH-TO-TEXT CONVERSION IN BIOMETRIC AUTHENTICATION SYSTEMS	114
Andrii Tkachuk, Bogdan Bulakh. DESCRIBING THE KNOWLEDGE ABOUT THE SOURCE CODE USING AN ONTOLOGY	123
Nick Odegov, Matin Hadzhyiev, Liudmyla Bukata. Liudmyla Glazunova, Marina Kochetkova. COMPARISON OF BIG DATA CLASSIFICATION ALGORITHMS BY SIMULATION METHODS.....	134
Nick Odegov, Matin Hadzhyiev, Liudmyla Bukata. Liudmyla Glazunova, Marina Kochetkova. JUSTIFICATION OF FAST CLASSIFICATION ALGORITHMS ON BIG DATA SETS WITH RELIABILITY AND PERFORMANCE CRITERIA	148
Yaroslav Kornaga, Andrii Barabash. MATHEMATICAL MODEL OF EVENTS MONITORING IN THE MANAGEMENT SYSTEM OF DISTRIBUTED DATABASES BASED ON EVENT SENSORS.....	161
Andrii Sereda, Valery Samarai, Ivan Datsenko, Volodymyr Pavlenko, Evgen Brovchenko, Peter Talanchuk. EVOLUTION OF MOBILE EDUCATION TO IMPROVE THE QUALITY OF LEARNING	167
Maksym Butenko, Volodymyr Pavlenko. AGRICULTURAL LAND PRICE PREDICTION USING MACHINE LEARNING ALGORITHMS.....	175
Illia Vokhranov, Bogdan Bulakh. COMBINED APPROACHES TO THE STATIC CODE ANALYSIS USING NEURAL NETWORKS	183
Evgen Brovchenko, Valery Samarai, Ivan Datsenko, Volodymyr Pavlenko, Andrii Sereda. PROTECTION OF UNSTRUCTURED INFORMATION ON A MOBILE DEVICE.....	194
Volodymyr Levchenko, Volodymyr Pavlenko, Peter Talanchuk. BULK SHEAR WAVES IN A LAYER--PERIODIC STRUCTURE "PIEZOELECTRIC - SLOT"	201
ARTICLE DESIGN RULES	205

ДО 100-РІЧЧЯ ПРОФЕСОРА ЦИБІЗОВА К.М.

НАРИС ЖИТТЯ ТА НАУКОВОЇ ДІЯЛЬНОСТІ ДОКТОРА ТЕХНІЧНИХ НАУК
ПРОФЕСОРА КОСТЯНТИНА МИКОЛАЙОВИЧА ЦИБІЗОВА.....

DOI 10.36994 / 2788-5518-2023-01-05-02

Костянтин Миколайович Цибізов народився 28 липня 1923 р. у селі Нижня Добринка Молотівського (Жирнівського) району Сталінградської (Волгоградської) області на батьківщині. Його батько Микола Іванович Цибізов народився також на батьківщині 1898 р. та був різноробом. Тільки у двох дітей класу, у якому навчався майбутній науковець, батьки не були репресовані. Під час війни Микола Іванович був призваний на фронт, воював у званні червоноармійця та загинув 26.11.1942 р. у Псковській області. Його мати Цибізова (уроджена Черв'якова) Ганна Василівна народилася там же 28 липня 1899 (до речі, народивши сина Костянтина у день свого народження). Костянтин Миколайович в анкетах про своє соціальне походження писав "з селян".

На початку 1930-х років сім'я переїхала до Ашхабада, рятуючись від голоду у Поволжі. Середню школу Костянтин Миколайович успішно закінчив 1940 р. в Ашхабаді та вступив до Ленінградського електротехнічного інституту інженерів зв'язку ім. М.О. Бонч-Бруєвича. Після року навчання в інституті у липні 1941 р. був призваний до лав Червоної Армії і направлений як курсант на прискорене навчання до Військового училища повітряного спостереження, сповіщення та зв'язку (ВУПССЗ), яке формувалося на території Ленінградського училища зв'язку і було передислоковане в м.Бірьск (Башкирія). Тут Костянтин Миколайович познайомився зі своїми майбутнім науковим керівником Г.В. Кисунько, який також був курсантом ВУПССЗ. Навчання тривало 10 місяців, і у травні 1942 р. вже у первинному офіцерському званні був направлений на Далекий Схід. Служив у радіотехнічних військах (в Хабаровську, в Олександрівську-Сахалінському, в Миколаївську-на-Амурі) на посадах начальника станції радіоуправління літаків (РУЛ), командира взводу окремої роти РУЛ та заступника командира окремої роти РУЛ (так називалися станції радіолокаційного виявлення). Брав участь у бойових діях у війні з Японією в Південно-Сахалінській операції у серпні 1945 р. У вересні 1946 р. вступив у Ленінграді до Військової академії зв'язку ім. С.М. Будьонного (ВАЗ), яку закінчив з відзнакою 1952 р. і цього ж року вступив до ад'юнктури академії.



Науковим керівником К.М. Цибізова став д.т.н., професор Г.В. Кисунько, майбутній член-кореспондент АН СРСР. Докторську дисертацію Г.В. Кисунько захистив 1951 р. у науково-дослідному інституті, директором якого був академік АН СРСР А.І. Берг. Слід зазначити, що 1950 р. Г.В. Кисунько був направлений до «закритого» СБ-1, підпорядкованого Третьому Головному Управлінню при Раді Міністрів СРСР і створеного для розробки системи ППО (спочатку система "Беркут", потім – С-25). Згодом Г.В. Кисунько став Генеральним конструктором першої системи ПРО "Система "А" та Генеральним конструктором системи ПРО "А-35". У роботах зі створення антено-фідерних трактів цих систем Костянтин Миколайович брав найактивнішу участь.

У 1955 р. у ВАЗ Костянтин Миколайович захистив кандидатську дисертацію, а 1965 р. там же – докторську дисертацію. Цікавою виявилася наукова дискусія з опонентом його докторської дисертації, відомим вченим у галузі класичної електродинаміки д.ф.-м.н. Б.З. Каценеленбаумом

(докторська дисертація захищена 1961 р.), який врешті-решт погодився з підходом Костянтина Миколайовича. Цибізов К.М. узагальнив метод власних функцій та хвильоводних рівнянь, вперше запропонований для циліндричних хвильоводів Г.В. Кисунько, на косокутні криволінійні координати. Цим методом здійснено аналіз поля у хвильоводах зі змінним перерізом уздовж поздовжньої координати та наближено визначено власні функції та власні значення циліндричного хвильоводу складного поперечного перерізу. На відміну від методу плоских поперечних перерізів, розвинутого Б.З. Каценеленбаумом, поле у хвильоводі описується за допомогою власних векторних функцій, які окремо задовольняють граничним умовам. Для амплітуд поля при цих функціях отримана нескінченна система диференціальних рівнянь першого порядку зі змінними коефіцієнтами, в якій взаємопов'язана вся безліч типів хвиль.

Слід зазначити, що з наукової школи Г.В. Кисунько вийшли такі видатні вчені, як д.т.н., проф. Б.М. Машковцев та д.т.н., проф. Б.Ф. Ємелін, з якими Костянтин Миколайович пов'язувала не тільки наукова робота, а й велика людська дружба. Ними були виконані масштабні науково-дослідні та дослідно-конструкторські роботи «з тематики Г.В. Кисунько» у 1950-1960 р.р.. Костянтин Миколайович до Г.В. Кисунько завжди ставився з надзвичайною повагою і навіть захопленням. Зауважимо, що зі школи Г.В. Кисунько вийшов ще один видатний учений у галузі електродинаміки – д.т.н., проф. В.В. Нікольський, який однак мав інший підхід до вирішення внутрішніх задач електродинаміки (докторська дисертація захищена 1962 р.).

У 1966 р. Костянтин Миколайович був призначений заступником начальника з навчально-наукової роботи (ННР) щойно створеного Київського вищого військового інженерного двічі Червонопрапорного училища зв'язку ім. М.І. Калініна (КВВІДЧПУЗ). Нині це Військовий інститут телекомунікацій та інформатизації імені Героїв Крут. На той час він як голова викладацького «десанту» був направлений з ВАЗ для організації навчально-наукового процесу до КВВІДЧПУЗ. Цибізов-адміністратор очолив групу талановитих викладачів ВАЗ, які «десантувалися» в Києві з метою організації навчально-наукового процесу вищого військового навчального закладу інженерного рівня в «дусі» ВАЗ. Цибізовим-організатором були закладені основи підготовки військових інженерів-зв'язківців, які з перших випусків мали надвисоке поціновування у військах зв'язку. Така повага підтверджується тим, що на всіх вузлах зв'язку Генерального Штабу МО СРСР або польових частинах центрального підпорядкування займали відповідні високі посади випускники КВВІДЧПУЗ. Про фундаментальність організації навчально-наукового процесу свідчить той факт, що через 30 років у 1990-ті та 2000-ті р.р. вищі посади багатьох провідних телекомунікаційних компаній, підприємств та організацій в Україні обіймали і досі обіймають випускники КВВІДЧПУЗ.

На посаді заступника начальника училища з ННР розкрився талант Костянтина Миколайовича не тільки як організатора навчально-наукового процесу, а й як прогресивного адміністратора. Доповідати заступнику начальника училища з ННР Цибізову К.М. було дуже важко, оскільки він вказував час який давався на доповідь, і якщо доповідач не дотримувався часового ліміту, доповідь переривалася і переносилася на наступний раз. При цьому час потрібно було узгодити.

Проте «тяга» до наукової та педагогічної діяльності як у педагога від Бога, узяла «своє». У 1973 р. Костянтин Миколайович написав рапорт про звільнення з посади заступника начальника училища з ННР з призначенням на будь-яку педагогічну посаду в училищі. То справді був вчинок: він відмовився від високої посади в Києві.

Наукова діяльність К.М. Цибізова була присвячена широкому колу теоретичних та прикладних питань у галузі електродинаміки, теорії електромагнітного поля, хвильоводної та антенної техніки. Отримані ним результати в теорії та техніці НВЧ набули світове визнання. Так, наприклад, після виходу у світ монографії «Складні хвильоводні системи» була спроба запросити Костянтина Миколайовича за кордон для читання спеціального курсу по антено-хвильоводним трактам, що здійснити тоді було зовсім неможливо. Цибізов-вчений написав близько 100 наукових праць, видав 3 монографії, понад 20 підручників та навчальних посібників. Для видатного вченого це невелика кількість наукових праць, але цьому є цілком чітке пояснення. Обов'язковість перед людьми, готовність прийти на допомогу не тільки учням, а й будь-кому, хто звертався з науковим чи навчальним питанням, принциповість і чесність, доброчесність й уважність не дозволяли йому працювати «тільки для себе».

З кінця 1940-х років освоєння сантиметрового діапазону хвиль стало завданням №1 для радянської радіолокації. Костянтин Миколайович опинився на піку затребуваності фахівців з

теорії та техніки НВЧ і в цьому велике щастя для вченого. Можна стверджувати, що в цій царині він був щасливою людиною. Він жив роторами, дивергенціями, градієнтами, але ніщо людське йому не було чужим. На багатьох «формальних» зборах і нарадах на той час він часто працював із ручкою в руці, виводячи нові формули електромагнітного поля на листочках «у клітинку». Цибізовим-вченим всебічно досліджено секторно-гвинтові хвилеводи, збудники хвилі H_{01} , секторіальний рупор, міжхвилеводні резонатори зі змінним перетином. Під керівництвом Цибізова-вченого виконані роботи з дослідження щілинних антен на смужкових лініях, низькопрофільних антен, опромінювачів дзеркальних антен з автосупроводом, багатодіапазонних опромінювачів, спіральних антен для станції космічного зв'язку «Кристал», багатопроменевих фазованих антенних решіток, дзеркальних антен для станцій «Злива» та «Окуляр», багатодіапазонних антен з опромінювачами на секторних і коаксіально-секторних хвилеводах, антен конструкцій, що трансформуються, адаптивних антен з автокомпенсаторами перешкод. Костянтин Миколайович брав активну участь у створенні антенно-фідерних трактів РЛС А-100 (система далекого радіолокаційного виявлення), РЛС Б-200 (система наведення зенітних ракет). Особливою складністю відрізнялися хвилеводні тракти з хвилеводними комутаторами на зчленуваннях, що обертаються, розроблені за його безпосередньої участі.

У створеній Цибізовим-вченим теорії складних хвилеводів подальший розвиток отримав апарат функцій Ейрі. Після використання функцій Ейрі в теорії розповсюдження радіохвиль академіком АН СРСР В.А. Фоком це найбільш повне застосування апарату функцій Ейрі в електродинаміці. До застосування апарату функцій Ейрі призводить рішення диференціального рівняння другого порядку зі змінним коефіцієнтом, отримане в діагональному наближенні методом приєднаних рівнянь.

Наукова школа К.М. Цибізова налічує більше 20 учнів, а Л.В. Скрипник (1994р.), М.В. Гряник (1997р.), В.М. Почерняєв (1999 р.) стали докторами наук. Його учні стали лауреатами державних премій СРСР та України та мають наукове звання професора. Він підготував близько 200 дипломників за спеціальністю «військового радіоінженера». Цікавим є досвід К.М. Цибізова у підготовці кандидатів та докторів наук. Він любив розповідати учням, що зі своїм вчителем Г.В. Кисунько в ході роботи над кандидатською дисертацією зустрічався три рази: при постановці наукового завдання, за поданням до Ради закінченої дисертаційної роботи та на захисті дисертації, включаючи банкет. Однак сам він не дотримувався цього тренду, якщо говорити сучасною мовою, а щедро дарував свій талант Педагога та Вченого всім, хто до нього звертався. Зазначимо, що вищезгадані три зустрічі, можливо, були зумовлені тим, що Г.В. Кисунько вже тоді був «сильно засекречений».

Характерною рисою Цибізова-педагога було те, що він завжди готувався до кожної лекції, навіть перебуваючи на посаді заступника начальника училища з ННР або начальника кафедри «Антенно-фідерних пристроїв та поширення радіохвиль». Він став професором цієї кафедри після 1983 р., коли у віці 60 років пішов у відставку. Це повчальний приклад для всіх викладачів, незалежно від наукового звання, наукового ступеня та педагогічного стажу. Аудиторія, незалежно від свого складу, любила слухати його лекції завдяки образно-точній манері викладу складного матеріалу фізико-математичного змісту. І тут розкривався Цибізов-методист, який крейдою списував дошку так, що йому завжди її вистачало на навчальне питання. Конспектувати за ним було не важко, і, на щастя, збереглися конспекти лекцій у його учнів. Ці конспекти лекцій з електродинаміки, напевно, можуть бути видані за прикладом фейнманівських лекцій з фізики. Костянтин Миколайович любив повторювати приблизно таке: «Крейда для педагога не бруд, а шляхетний атрибут шляхетної професії».

У назві цього Нарису показана актуальність наукової діяльності Костянтина Миколайовича і в епоху цифрових трансформацій. Блискучий опонент на захисті докторських та кандидатських дисертацій, прихильник «гарної» математики, під якою розумів аналітичне розв'язання електродинамічних завдань. Вважав, що одна коротка наближена формула, отримана при розв'язанні електродинамічної задачі, прикрашає це рішення більше, ніж рішення, отримане чисельними методами.

Цибізов-методист вмів ясно сформулювати будь-яку життєву проблему, чітко розумів «новинні» події і міг доступно їх роз'яснити будь-якій аудиторії слухачів.

До кінця своїх днів він залишався чинним професором створеної ним кафедри.



Д.т.н., професор Почерняєв В.М.

НАУКОВІ ТВОРИ ВДЯЧНИХ УЧНІВ ПРОФЕСОРА ЦИБІЗОВА К.М

ПЕРЕДМОВА

Частину журналу присвячено 100-річчю з дня народження доктора технічних наук, професора Костянтина Миколайовича Цибізова, видатного українського Вченого та Педагога.

Добірка статей включає роботи учнів К.М. Цибізова, в тому числі вже його «наукових онучок», колег по кафедрі, яку він очолював, а потім на якій працював до кінця життя.

Стаття Почерняєва В.М. і Магомедової М.С. присвячена частково заповненому діелектриком секторному хвилеводу, а як відомо, власні векторні функції полого секторного хвилеводу вперше записані К.М. Цибізовим.

Стаття Почерняєва В.М. і Сивкової Н.М. присвячена електромагнітному полю резонатора зі змінним перерізом, а саме рішення подібних задач з використанням апарату функцій Ейрі було запропоновано К.М. Цибізовим у його відомих монографіях.

Стаття Почерняєва В.М. і Сушанової В.С. присвячена моделюванню багатопроменевого каналу зв'язку, в якій використовується апарат спеціальних функцій математичної фізики, яким повною мірою володів К.М. Цибізов.

Стаття Наритника Т.М. і колег присвячена актуальному питанню дослідження квазіоптичного відкритого резонатора терагерцового діапазону частот, що відповідає напрямку дослідження К.М. Цибізова по хвилеводним НВЧ резонаторам.

Стаття Сайка В.Г., Ільїнова М.Д., Нестеренка І.К. присвячена питанню синтезу вібраторних антен, що відповідає тематиці науково-педагогічної діяльності К.М. Цибізова на кафедрі «Антенно-фідерних пристроїв та поширення радіохвиль».

УДК 621.372.8

ПОЛЯРИЗАТОР НА СЕКТОРНОМУ ХВИЛЕВОДІ, ЧАСТКОВО ЗАПОВНЕНОМУ ДІЕЛЕКТРИКОМ

DOI 10.36994 / 2788-5518-2023-01-05-03

Почерняєв В.М., д.т.н., проф. Національна академія СБ України, Київ, Україна.
vpochernyaev@gmail.com

Магомедова М.С., Київський фаховий коледж зв'язку, Київ, Україна. kkz.praktika@ukr.net

Анотація: У статті досліджено поляризатор на секторному хвилеводі, частково заповненому діелектриком. Розроблено конструкцію поляризатора на секторному хвилеводі з меншими габаритами діелектричної пластини та кращим коефіцієнтом стоячої хвилі зчленування прямокутний хвилевод – поляризатор – прямокутний хвилевод порівняно з поляризатором на круглому хвилеводі. Показано схеми зчленування хвилеводних відрізків антено-фідерного тракту на прямокутних хвилеводах з поляризатором на секторному хвилеводі та з поляризатором на круглому хвилеводі. Наведено формулу постійної поширення секторного хвилеводу, частково заповненого діелектриком та ефективної діелектричної проникності секторного хвилеводу. Наведено залежність відносних діелектричних проникностей хвилеводу з секторною діелектричною пластиною, паралельною та перпендикулярною площині поляризації падаючої хвилі від усередненої товщини секторної діелектричної пластини. У роботі проаналізовано секторний хвилевод, частково заповнений діелектриком, який має $\theta = \pi/3$ і $\Phi = \pi/6$. У цьому випадку використовується власна скалярна функція секторного хвилеводу для хвилі H_{31} . Представлені відповідні залежності зміни диференціального фазового зсуву в діапазоні хвиль, що розглядається, від усередненої товщини секторної пластини. При малих товщинах фазуючих пластин зменшення широкосмугового поляризатора обумовлено збільшенням фазового зсуву в довгохвильовій частині діапазону, а при великих - в короткохвильовій. Підбором діелектричної проникності матеріалу можна збільшити широкосмуговий поляризатор за рахунок деякого збільшення довжини пластини. Наведено залежність коефіцієнта стоячої хвилі зчленування прямокутний хвилевод – поляризатор – прямокутний хвилевод від величини λ/λ_0 . Коефіцієнт стоячої хвилі для зчленування прямокутний хвилевод – поляризатор на секторному хвилеводі – прямокутний хвилевод на 3...5 % менше, ніж зчленування прямокутний хвилевод – поляризатор на круглому хвилеводі – прямокутний хвилевод. Поляризатор на секторному хвилеводі з діелектричною пластиною має перевагу перед поляризатором на круглому хвилеводі з діелектричною пластиною при побудові антено-фідерних трактів на прямокутних хвилеводах. Ця перевага полягає у досягненні меншого значення коефіцієнта стоячої хвилі на хвилеводному відрізку тракту з поляризатором та меншою довжиною діелектричної пластини – $1,25\lambda_0$.

Ключові слова: поляризатор, секторний хвилевод, секторна діелектрична пластинка, ефективна діелектрична проникність, постійна поширення.

POLARIZER ON A SECTOR WAVEGUIDE PARTIALLY FILLED BY DIELECTRIC

Vitaly Pochernyaev, Dr.habil. Prof. National Academy of Security of Ukraine, Kyiv, Ukraine.
vpochernyaev@gmail.com

Mariia Mahomedova, Kyiv Professional College of Communications, Kyiv, Ukraine. kkz.praktika@ukr.net

Abstract: In the article a polarizer based on a sector waveguide partially filled by dielectric is studied. The design of a polarizer based on a sector waveguide with smaller dimensions of the dielectric plate and a better coefficient of the standing wave rectangular waveguide - polarizer on sector waveguide - rectangular waveguide junction compared with rectangular waveguide-polarizer based on a circular waveguide -rectangular waveguide junction has been developed. The diagrams of the articulation of the waveguide segments of the antenna-feeder path on rectangular waveguides with a polarizer on a sector waveguide and with a polarizer on a circular waveguide are shown. Formulas for the propagation constant of a sector waveguide partially filled with a dielectric and the effective permittivity of a sector waveguide are given. The dependences of the relative permittivities of a waveguide with a sector dielectric plate parallel and perpendicular to the plane of polarization of the incident wave on the average thickness of the sector dielectric plate are presented. The paper analyzes a sector waveguide partially filled with a dielectric, which has $\theta = \pi/3$ and $\Phi = \pi/6$. In this case, the scalar eigenfunction of the sector waveguide for the quasi- H_{31} wave is used. The corresponding dependences of the change in the differential phase

shift in the considered wave range on the average thickness of the sector plate are presented. At small thicknesses of the phasing plates the decrease in the polarizer broadband is due to an increase in the phase shift in the long-wavelength part of the range and at large thicknesses in the short-wavelength one. By choosing the permittivity of the material, it is possible to increase the broadband of the polarizer at the expense of some increase in the length of the plate. The dependence of the standing wave coefficient of the junction rectangular waveguide - polarizer - rectangular waveguide on the value λ/λ_0 . The standing wave coefficient for the junction of rectangular waveguide - polarizer on a sector waveguide - rectangular waveguide is 3...5 % less than for a junction of rectangular waveguide - polarizer on a circular waveguide - rectangular waveguide. A polarizer based on a sector waveguide with a dielectric plate has an advantage over a polarizer based on a circular waveguide with a dielectric plate when building antenna-feeder paths on rectangular waveguides. This advantage consists in achieving a lower value of the standing wave coefficient on the waveguide section of the path with a polarizer and a shorter length of the dielectric plate – $1,25\lambda_0$.

Keywords: polarizer, sector waveguide, sector dielectric plate, effective permittivity, propagation constant.

Вступ

Одним із широко поширених поляризаторів є поляризатор на круглому хвильоводі. Такий поляризатор представляє собою круглий хвильовод з розташованою в ньому діелектричною пластиною. [1]. При проектуванні та аналізі роботи поляризатора на круглому хвильоводі, насамперед, необхідно знати відносні діелектричні проникності хвильоводу з діелектричною пластиною, паралельною та перпендикулярною площині поляризації падаючої хвилі. Однак багато антено-фідерних трактів телекомунікаційних станцій НВЧ діапазону реалізовані на прямокутних хвильоводах. Зчленування поляризатора на круглому хвильоводі з прямокутними хвильоводами на вході та виході має не завжди відповідний коефіцієнт стоячої хвилі при довжині діелектричної пластини $1,5\lambda_0$.

Метою роботи є розробка конструкції поляризатора з меншими габаритами діелектричної пластини та кращим коефіцієнтом стоячої хвилі зчленування прямокутний хвильовод – поляризатор – прямокутний хвильовод.

Виконання досліджень

Альтернативним варіантом побудови поляризатора на круглому хвильоводі з діелектричною пластиною є конструкція поляризатора на секторному хвильоводі. Власні функції та критичні хвильові числа полого секторного хвильоводу визначені К.М.Цибізовим у першій монографії [2], потім результати досліджень секторного хвильоводу наводяться у монографіях [3,4]. На рис.1. показані схеми зчленування хвильоводних відрізків антено-фідерного тракту на прямокутних хвильоводах з поляризатором на секторному хвильоводі (рис.1а) та з поляризатором на круглому хвильоводі (рис.1б). Для розрахунку характеристик поляризатора на секторному хвильоводі необхідне знання ефективної діелектричної проникності секторного хвильоводу та відповідні відносні діелектричні проникності хвильоводу з секторною діелектричною пластиною, паралельною та перпендикулярною площині поляризації падаючої хвилі.

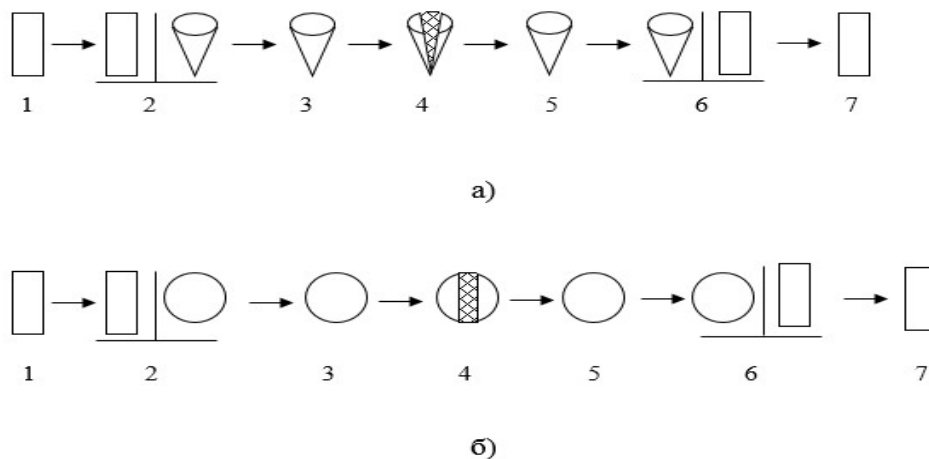


Рис.1. Схеми, що пояснюють включення поляризатора до антено-фідерного тракту: а) поляризатора на секторному хвильоводі; б) поляризатора на круглому хвильоводі

На рис.1а показано: 1,7 - прямокутний хвилевод; 2 – перехід із прямокутного хвилеводу на секторний хвилевод; 3,5 – секторний хвилевод; 4 - секторний хвилевод, частково заповнений діелектриком; 6 - перехід із секторного хвилеводу на прямокутний хвилевод.

На рис.1б показано: 1,7 - прямокутний хвилевод; 2 – перехід із прямокутного хвилеводу на круглий хвилевод; 3,5 – круглий хвилевод; 4 - круглий хвилевод з діелектричною пластиною; 6 - перехід із круглого хвилеводу на прямокутний хвилевод.

Постійне розповсюдження частково заповненого діелектриком секторного хвилеводу знайдемо у вигляді:

$$\gamma_0 = k_0 \sqrt{\varepsilon_{\text{еф}} - (\lambda/\lambda_{\text{кр}})^2}, \quad (1)$$

де k_0 – хвильове число; $\varepsilon_{\text{еф}}$ – ефективна діелектрична проникність; $\lambda, \lambda_{\text{кр}}$ – довжина хвилі та критична довжина хвилі відповідно.

Ефективна діелектрична проникність у (1) секторному хвилеводі знайдена в роботі [5]:

$$\varepsilon_{\text{еф}}^c = 1 + (\varepsilon_r - 1)(R_\delta \Phi / R^2 \theta) [1 - \sin(2\varphi_1 + \varphi) F_c],$$

де ε_r – відносна діелектрична проникність секторної пластини; R, R_δ – радіус хвилеводу та радіус пластини відповідно; θ, Φ – кут сектора поперечного перерізу хвилеводу та кут сектора поперечного перерізу пластини відповідно; φ_1 – кут, що фіксує положення діелектричної пластини; F_c – комбінація функції Бесселя першого роду, визначається залежно від застосованих власних скалярних функцій секторного хвилеводу.

Нехай секторний хвилевод, частково заповнений діелектриком, має $\theta = \pi/3$ та $\Phi = \pi/6$. У цьому випадку будемо використовувати власну скалярну функцію секторного хвилеводу для хвилі H_{31} . Така власна скалярна функція має вигляд:

$$\psi_{h31} = C_{h31} \mathfrak{I}_3(\alpha_{31} r) \cos 3(\varphi - \frac{\pi}{3}), \quad (2)$$

$$C_{h31} = \sqrt{\frac{2(2 - \delta_{03})}{\pi/3}} \frac{\mu_{31}}{R \sqrt{\mu_{31}^2 - \mathfrak{I}_3(\mu_{31})}} = \sqrt{\frac{12}{\pi}} \frac{\mu_{31}}{R \sqrt{\mu_{31}^2 - 9\mathfrak{I}_3(\mu_{31})}},$$

де α_{31} – поперечне хвильове число хвилі H_{31} ; μ_{31} – перший корінь похідної функції Бесселя першого роду третього порядку.

Відносні діелектричні проникності секторного хвилеводу з секторною діелектричною пластиною, паралельною та перпендикулярною площині поляризації падаючої хвилі мають такий вигляд:

$$\varepsilon_{\parallel} = (\frac{\mu_{31}}{k_0 R})^2 + (\frac{\gamma_{\parallel}}{k_0})^2, \quad (3)$$

$$\varepsilon_{\perp} = (\frac{\mu_{31}}{k_0 R})^2 + (\frac{\gamma_{\perp}}{k_0})^2. \quad (4)$$

Умовні постійні поширення, що враховують відносні діелектричні проникності (3) та (4) та власна скалярна функція хвилі H_{31} (2) мають вигляд:

$$\gamma_{\parallel} = \gamma_0 - \frac{k_0(\varepsilon_r - 1)\psi_{h31}}{2\pi\gamma_0},$$

$$\gamma_{\perp} = \gamma_0 - \frac{k_0(\varepsilon_r - 1)\psi_{h31}}{2\pi\gamma_0}.$$

На рис.2 наведено залежності відносних діелектричних проникностей хвилеводу від усередненої товщини s секторної діелектричної пластини: крива 1 – відносна діелектрична

проникність $\varepsilon_{\parallel}$ для $\varepsilon_r = 2,5$; крива 2 - відносна діелектрична проникність $\varepsilon_{\perp}$ для $\varepsilon_r = 2,5$; крива 3 - відносна діелектрична проникність $\varepsilon_{\parallel}$ для $\varepsilon_r = 2$; крива 4 - відносна діелектрична проникність $\varepsilon_{\perp}$ для $\varepsilon_r = 2$. З графіків видно, що з певної товщини пластини забезпечується максимальний фазовий набіг ортогональних векторів електричного поля. Для діелектрика з $\varepsilon_r = 2,5$ товщина пластини складає $0,19\lambda_0$, а для діелектрика з $\varepsilon_r = 2$ товщина пластини складає $0,21\lambda_0$.

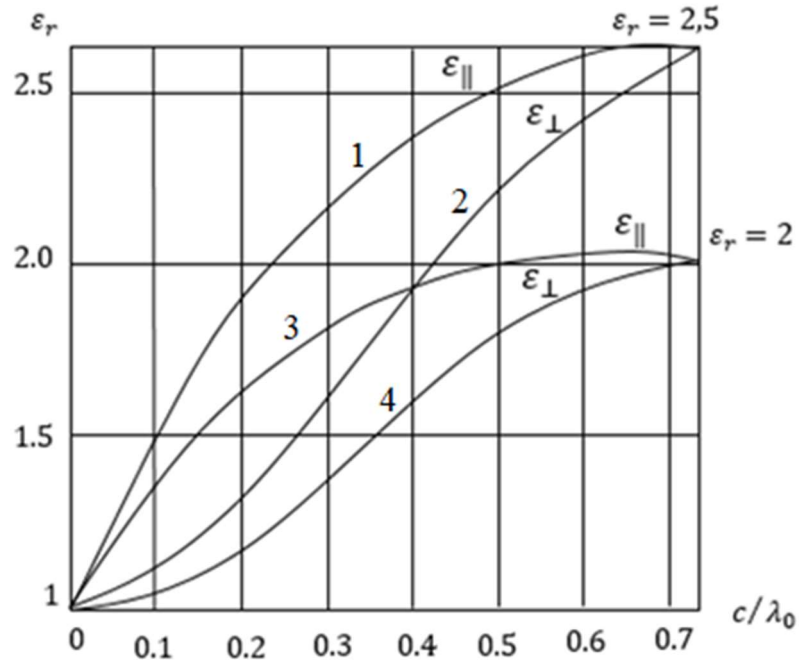


Рис. 2. Залежність відносних діелектричних проникностей хвильоводу від усередненої товщини секторної діелектричної пластини $\varepsilon_r = 2,5$ та $\varepsilon_r = 2$ за рівнем $R_d/2$

Розглянемо діапазонні можливості поляризатора. Основним параметром, що характеризує його роботу, є диференціальний фазовий зсув ортогональних векторів електричного поля. Він визначається виразом:

$$\Delta\varphi = \frac{2\pi l}{\lambda} \left(\sqrt{\varepsilon_{\parallel} - \left(\frac{\lambda}{\lambda_{\text{кр}}}\right)^2} - \sqrt{\varepsilon_{\perp} - \left(\frac{\lambda}{\lambda_{\text{кр}}}\right)^2} \right), \quad (5)$$

де l - довжина пластини.

Неважко помітити, що диференціальний фазовий зсув має екстремальну точку, яка обумовлюється максимальною швидкістю зміни набігу фаз і визначається за допомогою рівняння:

$$\frac{\partial}{\partial \lambda} \Delta\varphi = 0.$$

Таким чином, у кожному конкретному випадку існує певне співвідношення між $\varepsilon_{\parallel}$ та $\varepsilon_{\perp}$ або певна товщина пластини, що забезпечує при заданій λ_0 мінімальне значення зміни β у діапазоні хвиль. Зміна довжини хвилі в обидва боки від λ_0 призводить до збільшення фазового зсуву. Знайшовши екстремум виразу (5), отримаємо значення для λ_0 , при якій забезпечується максимальна широкосмуговість поляризатора:

$$\lambda_{0\text{max}} = \lambda_{\text{кр}} \sqrt{\frac{\varepsilon_{\parallel} \varepsilon_{\perp}}{\varepsilon_{\parallel} + \varepsilon_{\perp}}}. \quad (6)$$

Побудувавши за допомогою формул (3) та (4) залежність кореня виразу (6) від ступеня заповнення хвиеводу s при заданих середній довжині хвилі робочого діапазону λ_0 і діаметр хвиеводу, можна визначити товщину фазуючої діелектричної пластини, що забезпечує максимальну широкосмуговість. Можна також показати, що фазовий зсув на довжину хвилі для такої товщини діелектричної пластини на середній довжині хвилі визначається виразом:

$$\Delta\varphi|_{\lambda_{0max}} = \frac{2\pi\lambda_0(\varepsilon_{\parallel} - \varepsilon_{\perp})}{\lambda_{кр}\sqrt{\varepsilon_{\parallel}\varepsilon_{\perp}}}.$$

Наведемо кількісну оцінку описаних положень. На рис. 3 і 4 наведені частотні залежності диференціального фазового зсуву на довжину хвилі, розраховані для $\varepsilon_r = 2,5$ та $\varepsilon_r = 2$ за формулою (5) за допомогою графіку (рис.2).

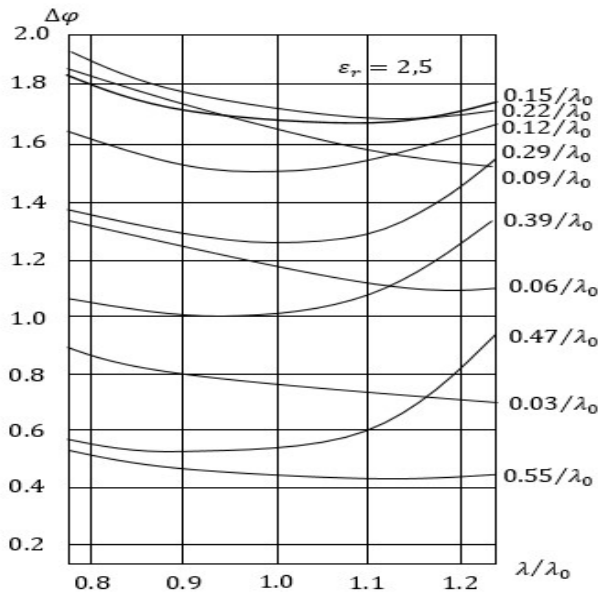


Рис.3. Частотні залежності диференціального фазового зсуву, віднесений до довжини хвилі

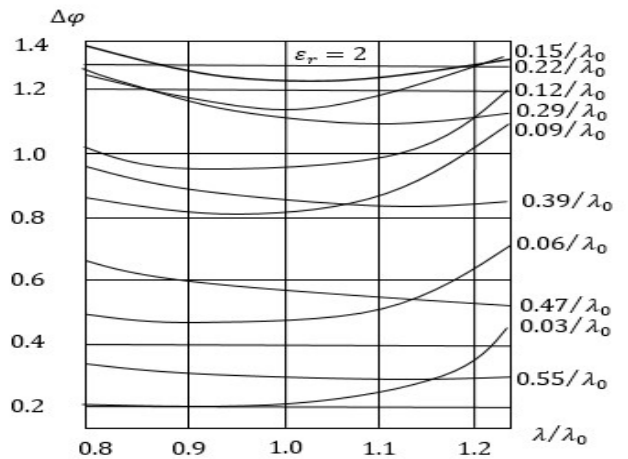


Рис.4. Частотні залежності диференціального фазового зсуву, віднесений до довжини хвилі $\varepsilon_r = 2$

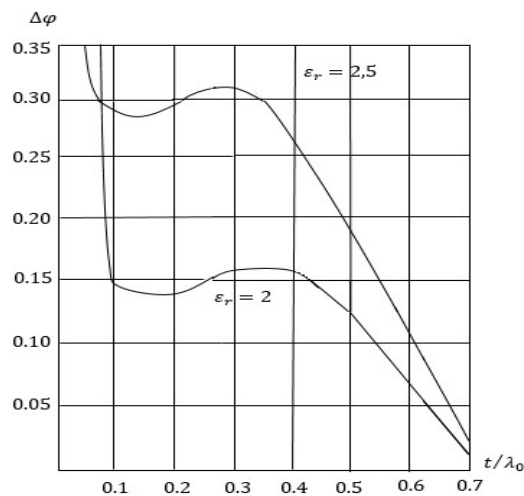


Рис.5. Відповідні рис.3 та 4 частотні залежності зміни диференціального фазового зсуву в діапазоні хвиль від усередненої товщини секторної пластини.

На рис. 5 представлені відповідні рис.3 та 4 залежності зміни диференціального фазового зсуву в аналізованому діапазоні хвиль від товщини пластини. Підбором діелектричної проникності матеріалу можна збільшити широкосмуговість поляризатора за рахунок деякого збільшення товщини пластини. Ще більшу широкосмуговість можна отримати при дотриманні нерівності ($c > \frac{\lambda_0}{2}$), однак при цьому збільшуються подовжні габарити поляризатора та ускладнюється узгодження поляризатора з антено-фідерним трактом.

При товщині пластини з відносною діелектричною проникністю $\varepsilon_r = 2$ зміна диференціального фазового зсуву на довжину хвилі в діапазоні хвиль 45% не перевищує 6° . Визначено, що при довжині фазуючої діелектричної пластини $\varepsilon_r = 2$ рівню $1,3\lambda_0$ забезпечується фазовий зсув у 90° . Зміна диференціального фазового зсуву діелектричної пластини $\varepsilon_r = 2,5$ при смузі частот 30% склало 3° .

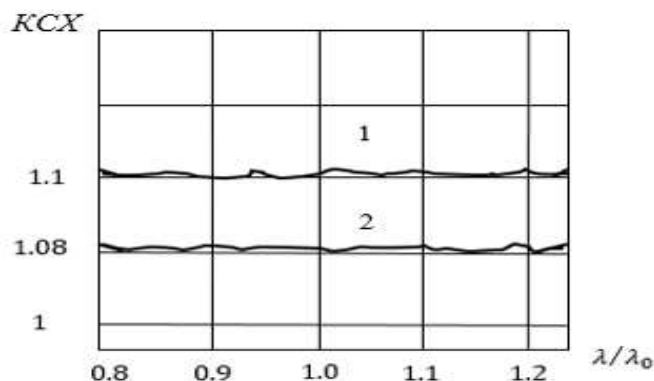


Рис.6. Залежність коефіцієнта стоячої хвилі зчленування прямокутний хвильовод – поляризатор – прямокутний хвильовод від величини λ/λ_0 : 1 – для поляризатора на круглому хвильоводі; 2 – для поляризатора на секторному хвильоводі

З рис.6 видно, що коефіцієнт стоячої хвилі для зчленування прямокутний хвильовод – поляризатор на секторному хвильоводі – прямокутний хвильовод менше, ніж для зчленування прямокутний хвильовод – поляризатор на круглому хвильоводі – прямокутний хвильовод на 3...5% у робочому діапазоні частот. Коефіцієнт стоячої хвилі визначається через модуль коефіцієнта відображення, який знаходиться через коефіцієнт трансформації, що є інтегралом від добутку відповідних поперечних електричних власних векторних функцій досліджуваного хвильоводу. Приклади таких розрахунків містяться в роботі [6].

Висновки

Поляризатор на секторному хвильоводі з діелектричною пластиною має перевагу перед поляризатором на круглому хвильоводі з діелектричною пластиною при побудові антено-фідерних трактів на прямокутних хвильоводах. Ця перевага полягає в досягненні меншого на 3...5% значення коефіцієнта стоячої хвилі на хвильоводному відрізку антено-фідерного тракту з поляризатором та меншою довжиною діелектричної пластини – $1,25\lambda_0$.

Література

- Орлеанская Э.В. Широкополосное устройство для поворота плоскости поляризации. *Радиотехника*, 1969. №8. С.23-29.
- Коган Н.Л., Машковцев Б.М., Цибизов К.Н. Сложные волноводные системы. Л.: Судпромгиз, 1963. -356 с.
- Машковцев Б.М., Цибизов К.Н., Емелин Б.Ф. Теория волноводов. М.: Наука, 1966. 351с.
- Почерняев В.Н., Цибизов К.Н. Теория сложных волноводов. Киев: Науковий світ, 2003. 224с.
- Почерняев В.Н., Скрыпник Л.В. Фазовые постоянные частично заполненных цилиндрических волноводов. *Радиоэлектроника*, 1996. №11. С.67– 70.
- Почерняев В.Н. Устройства на частично заполненных диэлектриком волноводах. Киев: УКНИПСК, 2000. 224с.

References

- Orleanskaya E.V. Shirokopolosnoe ustrojstvo dlya povorota ploskosti polarizacii. *Radiotekhnika*, 1969. №8. pp.23-29.

2. Kogan N.L., Mashkovtsev B.M., Tsibizov K.N. Slozhnie volnovodnie sistemi.L.:Sudpromgiz,1963. 356 s.
3. Mashkovtsev B.M., Tsybyzov K.N., Emelyn B.F. Teoryia volnovodov. M.:Nauka,1966.pp.351.
4. Pochernyaev V.N., Tsibizov K.N. Teoriya slozhnikh volnovodov. Kiev: Naukovii svit, 2003. 224s.
5. Pochernyaev V.N., Skripnik L.V. Fazovie postoyannie chastichno zapolnennikh tsilindricheskikh volnovodov. *Radioelektronika*, 1996. №11. S.67–70.
6. Pochernyaev V.N. Ustroystva na chastichno zapolnennyih dielektrikom volnovodah. Kiev: UKNIPS, 2000. 224c.

УДК 621.372.8

ЕЛЕКТРОМАГНІТНЕ ПОЛЕ БОЧКОПОДІБНОГО ПОЛОГО РЕЗОНАТОРА

DOI 10.36994 / 2788-5518-2023-01-05-04

Почерняєв В.М., д.т.н., проф. Національна академія Служби безпеки України, Київ, Україна.
vpochernyaev@gmail.com

Сивкова Н.М., док. філос. Національна академія Служби безпеки України, Київ, Україна.
natsivonat@gmail.com

Анотація: В статті досліджено електромагнітне поле бочкоподібного полого резонатора. Відзначимо, що у роботах К.М. Цибізова розроблено підхід визначення полів хвелеводних резонаторних пристроїв із змінним поперечним перерізом. У цій роботі цей підхід поширений для бочкоподібного полого резонатора. Рішення рівняння Ейрі, якими скористався К.М. Цибізов є наближеними до рішень, одержаних через рішення рівняння, що зводиться до однієї з форм гіпергеометричного рівняння або рівняння Вебера. Радіальна та азимутальна складові поля бочкоподібного полого резонатора відповідним складовим поля полого циліндричного резонатора, а поздовжня складова описується виродженою гіпергеометричною функцією. В роботі наведено конструкцію бочкоподібного хвелеводного резонатора з металевими торцями та записано вирази для резонансних довжин хвиль бочкоподібного резонатора з металевими торцями для H - та E -коливань. Наведені резонансні довжини хвиль та резонансні частоти для двох бочкоподібних резонаторів з металевими торцями у відповідності до циліндричного резонатора з радіусом рівним радіусу торця бочкоподібного резонатора та рівних продольних розмірів. Також в роботі, наведено хвелеводний резонатор з відкритим лівим торцем та записані перші п'ять коренів функцій Ейрі. Розгляд такого резонатора переслідує мету демонстрації узагальненості розглянутого підходу. Наведено конструкцію бочкоподібного резонатора з відкритими торцями, перерізи яких являються критичними. Вказано, що знаходження рішень рівнянь функцій, що описують поздовжній розподіл поля являється демонстрацією узагальненості підходу, розробленого К.М. Цибізовим, а в якості прикладів дослідженні резонатори змінного поперечного перерізу: бочкоподібний хвелеводний резонатор з металевими торцями, хвелеводний резонатор з відкритим лівим торцем та бочкоподібний резонатор з відкритими торцями, які вказані на рисунках.

Ключові слова: бочкоподібний полий резонатор, функції Ейрі, функції Вебера, рівняння Вебера, векторна функція, резонансні частоти.

ELECTROMAGNETIC FIELD OF A BARREL-SHAPED HOLLOW RESONATOR

Vitaly Pochernyaev, Dr.habi., Prof. National Academy of the Security Service of Ukraine, Kyiv, Ukraine.
vpochernyaev@gmail.com

Nataliia Syvkova, Ph.D. National Academy of the Security Service of Ukraine, Kyiv, Ukraine.
natsivonat@gmail.com

Abstract: In the article is investigates the electromagnetic field of a barrel-shaped hollow resonator. It is significant that in the articles of K.M. Tsibizov developed an approach for determining the fields of waveguide resonator devices with a variable cross section. In this article, this approach is extended to a barrel-shaped hollow resonator. The solution of the Airy equation, which was used by K.M. Tsibizov are approximate solutions obtained through solving an equation that reduces to one of the forms of the hypergeometric equation or the Weber equation. The radial and azimuthal components of the field of a barrel-shaped hollow resonator are the corresponding components of the field of a hollow cylindrical resonator, and the longitudinal component is described by a degenerate hypergeometric function. In this article presents the design of a barrel-shaped waveguide resonator with metal ends and writes expressions for the resonant wavelengths of a barrel-shaped resonator with metal ends for H - and E -oscillations. The resonant wavelengths and resonant frequencies for two barrel-shaped resonators with metal ends are given in accordance with a cylindrical resonator with an equal end radius of the barrel-shaped resonator and equal longitudinal dimensions. The article also presents a waveguide resonator with an open left end and writes down the first five roots of the Airy functions. The consideration of such a resonator aims at demonstrating the generality of the considered approach. The design of a barrel-shaped resonator with open ends, the sections of which are critical, is presented. It is indicated that finding solutions to the equations of functions that describe the longitudinal distribution of the field is a demonstration of the generalization of the approach developed by K.M. Tsibizov, and as examples of the study of resonators of variable cross section: a barrel-shaped waveguide resonator with metal ends, a waveguide resonator with an open left end, and a barrel-shaped resonator with open ends, which are shown in the figures.

Keywords: *barrel-shaped hollow resonator, Airy function, Weber function, Weber frequency, vector function, resonant frequencies.*

Вступ

У техніці НВЧ широке застосування знаходять резонатори з високою навантаженою добротністю, для якої необхідно мати високу власну добротність. Як відомо власні добротності циліндричних полів резонаторів для електричних та магнітних типів хвиль у першому наближенні пропорційні значенню LR де L – довжина циліндричного резонатора, R – його радіус. Полий резонатор у вигляді бочки має утворюючу, відмінну від лінійної і при фіксованому значенні L , може мати більше усереднене значення R_{cp} в порівнянні з радіусом циліндричного резонатора. Тоді, при одній і тій же довжині резонатора власна добротність бочкоподібного резонатора буде вищою за власну добротність циліндричного резонатора.

У роботах К. М. Цибізова [1-3] розглянуті циліндричні резонатори змінного поперечного перерізу з відкритими торцями, як резонатори складного профілю з утворюючою у вигляді параболи. У роботі [2, с.94] зазначено, що рішення резонаторних пристроїв складного профілю з такою твірною можна отримати, використовуючи рівняння Вебера, але "отримання рішення на цьому шляху важко, тому що воно виражається через порівняно рідко представлені в літературі функції Вебера". Якщо парабола апроксимована лінійною функцією, то рівняння Вебера переходить до рівняння Ейрі. Вирішення такого диференціального рівняння виражаються через функції Ейрі. Перевагами цього підходу стало не тільки визначення необхідних величин у явному вигляді для складного резонатора профілю з відкритими торцями, а й отримання безперервного рішення поблизу критичних перерізів і відсутність невизначеності в самому критичному перерізі. На наш погляд, після застосування академіком В. А. Фоком цих функцій у роботах з теорії дифракції та поширення хвиль, що увійшли до [4], найбільш ґрунтовне застосування функцій Ейрі в теорії та техніці НВЧ, продемонстровано в [1-3].

Метою цієї роботи є дослідження електромагнітного поля бочкоподібних полів резонаторів з металевими торцями, з відкритим лівим торцем, з відкритими торцями.

Виконання досліджень

Відомо, що рішення трьохвимірного рівняння Гельмгольца для циліндричного резонатора будь-якої форми поперечного перерізу може бути записано у вигляді [5]:

$$U_{\text{ц}} = \left[\sum_m (J_m(\chi_{mn}r) + PN_m(\chi_{mn}r))(C_m \sin m\varphi + D_m \cos m\varphi) \right] \begin{pmatrix} \sin B_z \\ \cos B_z \end{pmatrix}, \quad (1)$$

де $J_m(\chi_{mn}r)$, $N_m(\chi_{mn}r)$ – функції Бесселя першого та другого порядку відповідно, $P = \text{const}$, C_m , D_m – постійні коефіцієнти, які визначені граничними умовами. Допустимість такого підходу – прямої суперпозиції, безпосередньо слідує з лінійності та однорідності хвильового рівняння. Такий підхід можливий і у випадку бочкоподібного резонатора, як резонатора перемінного поперечного перерізу. Рішення буде мати вигляд:

$$U_{\text{б}} = (J_m(\chi_{mn}r) + PN_m(\chi_{mn}r)) \begin{Bmatrix} \sin m\varphi \\ \cos m\varphi \end{Bmatrix} \sum_p K_p \begin{Bmatrix} \sin B_{pz} \\ \cos B_{pz} \end{Bmatrix}, \quad (2)$$

де K_p – невідомі коефіцієнти, що належать визначенню.

Підхід К.М. Цибізова до вирішення резонаторних пристроїв перемінного поперечного перерізу полягає в тому, щоб підібрати відповідне диференціальне рівняння, яке описує позовжний розподіл поля. Покажемо це на прикладі трьох бочкоподібних резонаторів.

На рис.1 показаний профіль бочкоподібного хвильового резонатора з металевими торцями.

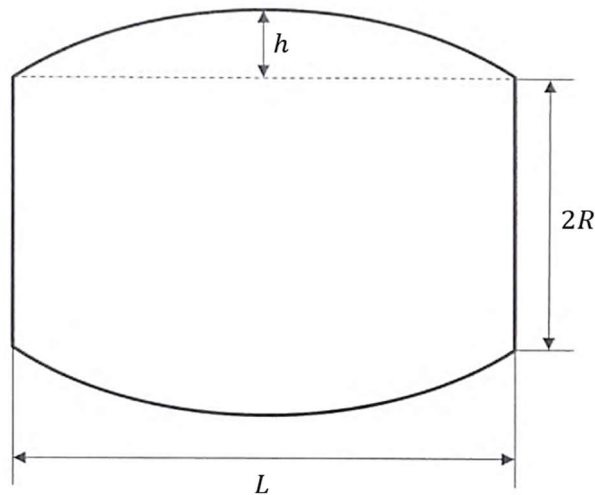


Рис.1. Профіль бочкоподібного хвильоводного резонатора з металевими торцями

Функція, що описує поздовжній розподіл поля, може бути знайдена з рівняння виду:

$$Z''(z) + G(z)Z(z) = 0. \quad (3)$$

Визначення функції $Z(z)$ виключає необхідність знаходження коефіцієнтів K_p в (2). Як показано в [1-3], при утворюючій у вигляді параболи рівняння (3) записується у вигляді:

$$Z''(z) + (c^2 - z^2)Z(z) = 0. \quad (4)$$

Коли утворюючу можна апроксимувати лінійною функцією, то К.М. Цібізов «відходить» від рішення рівняння Вебера (4), і вирішує рівняння виду:

$$Z''(z) + zZ(z) = 0. \quad (5)$$

Рішеннями рівняння (5) є функції Ейрі.

Перейдемо до рішення рівняння (4). Зробимо заміну перемінної: $\xi = z\sqrt{2p}$, $\eta(\xi) = Z(z)$ і позначив $2c^2 = 8k$, можна прийти до однієї з форм виродженого гіпергеометричного рівняння:

$$4\eta''(\xi) - (p\xi^2 - 8k)\eta(\xi) = 0. \quad (6)$$

Рішення рівняння (6) записується у вигляді:

$$\eta(\xi) = \left(\frac{1}{\sqrt{\xi}}\right) {}_1F_1\left(k, \frac{1}{4}, \frac{\xi^2}{2p}\right),$$

де ${}_1F_1$ - вироджена гіпергеометрична функція, $p = 1, 2, 3, \dots$

Зробивши зворотне перетворення отримаємо:

$$Z(z) = 2^{-\frac{1}{4}} z^{-\frac{1}{2}} {}_1F_1\left(\frac{c^2}{4}, \frac{1}{4}, pz^2\right). \quad (7)$$

Використовуючи умови першої змішаної крайової задачі для бочкоподібного полого резонатора з металевими торцями, можна побачити, що її рішення створює систему власних функцій:

$$U_6 = 0 \text{ при } r = R, \frac{\partial U_6}{\partial z} = 0 \text{ при } z = 0, L,$$

$$U_{mnp}^{(e)} = N_{mnp}^{(e)} J_m(\chi_{mn} r) \begin{Bmatrix} \sin m\varphi \\ \cos m\varphi \end{Bmatrix} Z'(z),$$

де $Z'(z)$ з формули (7).

Друга змішана крайова задача:

$$\frac{\partial U_6}{\partial z} = 0 \text{ при } r = R, U_6 = 0 \text{ при } z = 0, L \text{ і відповідно } U_{mnp}^{(h)} = N_{mnp}^{(h)} J_m(\chi'_{mn} r) \begin{Bmatrix} \sin m\varphi \\ \cos m\varphi \end{Bmatrix} Z(z).$$

Випишемо вираз для резонансних довжин хвиль бочкоподібного резонатора з металевими торцями:

- для Н-коливаний:

$$\lambda_{mnp} = \frac{2\pi}{\sqrt{\left(\frac{\mu_{mn}}{R_{cp}}\right)^2 + \left(\frac{p\pi}{2\sqrt{h^2 + \left(\frac{L}{2}\right)^2}}\right)^2}}, \quad (8)$$

де μ_{mn} - корінь рівняння $J'_m(\mu_{mn}) = 0$,

- для Е-коливаний:

$$\lambda_{mnp} = \frac{2\pi}{\sqrt{\left(\frac{\varepsilon_{mn}}{R_{cp}}\right)^2 + \left(\frac{p\pi}{2\sqrt{h^2 + \left(\frac{L}{2}\right)^2}}\right)^2}}, \quad (9)$$

де ε_{mn} - корінь рівняння $J_m(\varepsilon_{mn}) = 0$.

Другий доданок під радикалом в знаменнику виразів (8) та (9) для резонансних довжин хвиль є подовжнім хвильовим числом. Якщо покласти $h = 0, R_{cp} = R$, то «бочка» становиться «циліндром» і формули (8), (9) переходять в формули для резонансних довжин хвиль циліндричного резонатора.

Таблиця 1

Резонансні довжини хвиль та резонансні частоти для двох бочкоподібних резонаторів з металевими торцями у відповідності до циліндричного резонатора з радіусом рівним радіусу торця бочкоподібного резонатора та рівних продольних розмірів L

Тип резонатора Параметри резонатора	Циліндричний резонатор $L = 5,85$ см, $R = 2,3$ см	Бочкоподібний резонатор $L = 5,85$ см	
		$R_{cp} = \frac{9}{8} R$	$R_{cp} = \frac{7}{6} R$
Резонаторні довжини хвиль коливання H_{111}	6,52 см	7,136 см	7,3 см
Резонаторні частоти коливання H_{111}	4,6 ГГц	4,2 ГГц	4,1 ГГц

В табл.1 наведені резонансні довжини хвиль та резонансні частоти для двох бочкоподібних резонаторів з металевими торцями у відповідності до циліндричного резонатора з радіусом рівним радіусу торця бочкоподібного резонатора та рівних продольних розмірів L . Як видно з

табл.1, бочкоподібний резонатор можливо використовувати нижче по частоті практично для тих же габаритів та вимірювати резонансну частоту шляхом зміни величини R_{cp} .

Підхід К.М. Цибізова шляхом рішення рівняння (5) можливий і у випадку бочкоподібного резонатора з металевими торцями. Метод асимптотичного інтегрування виразу (5) приводить до рішення через функції Ейри:

$$Z(z) = \sqrt[4]{\frac{t}{p(z)}} [u(t) + jv(t)],$$

де $u(t)$ та $v(t)$ – функції Ейри [1-4], а параметр t визначається з співвідношення:

$$\int_{z_0}^z \sqrt{-p(z)} dz = \frac{2}{3} (-t)^{\frac{3}{2}}, (z > z_0, t < 0),$$

$$\int_z^{z_0} \sqrt{p(z)} dz = \frac{2}{3} (t)^{\frac{3}{2}}, (z < z_0, t > 0),$$

При великих t функції Ейри можуть бути замінені головними членами їх асимптотичних розкладів. Тоді маємо:

$$Z(z) = \frac{1}{\sqrt[4]{p(z)}} \exp\left(\pm \int_z^{z_0} \sqrt{p(z)} dz\right), p > 0, z - z_0 < 0,$$

$$Z(z) = \frac{1}{\sqrt[4]{-p(z)}} \exp\left(\pm j \int_{z_0}^z \sqrt{-p(z)} dz\right), p < 0, z - z_0 > 0.$$

На рис.2 наведено хвилеводний резонатор з відкритим лівим торцем. Розгляд такого резонатора переслідує мету демонстрації узагальненості розглянутого підходу. Будемо вважати, що переріз лівого торця такого резонатора – «півбочки», є критичним. Тоді, функція, яка описує продольний розподіл поля, визначається з рівняння:

$$Z_1''(z) + (az + b)Z_1(z) = 0. \quad (10)$$

Зробимо заміну змінної: $\xi = az + b$, $\eta(\xi) = Z_1(z)$. Отримаємо:

$$\eta''(\xi) + \left(\frac{1}{a^2}\right)\xi\eta(\xi) = 0.$$

Рішення має вигляд: $\eta(\xi) = \sqrt{\xi} J_{\frac{1}{3}}\left(\left(\frac{2}{3a}\right)\xi^{\frac{3}{2}}\right)$. Рішення рівняння (10) отримує вигляд:

$$Z_1(z) = \sqrt{az + b} J_{\frac{1}{3}}\left(\left(\frac{2}{3a}\right)(az + b)^{\frac{3}{2}}\right). \quad (11)$$

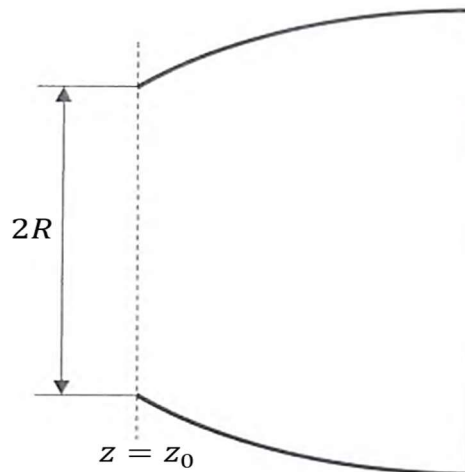


Рис.2. Хвильоводний резонатор з відкритим лівим торцем

При $a = 1, b = 0$ отримаємо результат робіт [1-3].

Крайові умови: вимога кінцевості рішення при $z < z_1$ та умова, яка впливає з рівності нулю дотичної складової електричного поля на ідеально провідниковій торцевій стінці хвильовода ($z = z_L$). Рівняння для визначення резонансних хвильових чисел для H -хвиль:

$$\int_{z_0}^{z_1} \sqrt{-p(z)} dz = \frac{2}{3} (-t_i^0)^{\frac{3}{2}},$$

де $p(z) = -1 + \chi_h^2/k^2(z)$, t_i^0 - i -корінь функції $v(t)$.

Вираз для визначення резонансних хвильових чисел для E -хвиль:

$$\int_{z_0}^{z_1} \sqrt{-p(z)} dz = \frac{2}{3} (-t_i')^{\frac{3}{2}},$$

де $p(z) = -1 + \chi_e^2/k^2(z)$, t_i' - i -корінь функції $v'(t)$.

Перші п'ять коренів функцій $v(t)$ та $v'(t)$ зведені в табл.2.

Таблиця 2
Корені функцій $v(t)$ та $v'(t)$

i	t_i^0	t_i'
1	2,338	1,019
2	4,088	3,248
3	5,52	4,82
4	6,787	6,163
5	7,944	7,372

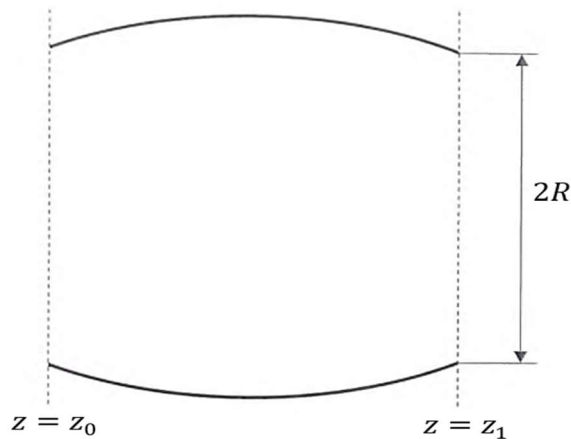


Рис.3. Бочкоподібний резонатор з відкритими торцями, перерізи яких являються критичними

На рис.3 показано бочкоподібний резонатор з відкритими торцями, перерізи яких являються критичними. Крайовою умовою є вимога кінцевості рішення на всій осі z . В цьому випадку функція $p(z)$ має два корені, які відповідають двом критичним перерізам $z = z_0$ та $z = z_1$. Якщо ці корені прості, то від рівняння (4) можна перейти до рівняння виду:

$$Z_2''(z) + (2s + 1 - z^2)Z_2(z) = 0. \quad (12)$$

Параметр s приймає точно цілі значення і рішення рівняння (12) має наступний вигляд:

$$Z_2(z) = \exp\left(\frac{-z^2}{2}\right) H_s(z), \quad (13)$$

де $H_s(z)$ – поліном Ерміта.

Рівняння для визначення резонансних хвильових чисел з результатом робіт [1-3]:

$$\int_{z_0}^{z_1} \sqrt{-p(z)} dz = \left(s + \frac{1}{2}\right) \pi, \text{ при } s = 1, 2, 3, \dots$$

$$p(z) = -1 + \chi_e^2/k^2(z) - \text{для } E\text{-хвиль},$$

$$p(z) = -1 + \chi_h^2/k^2(z) - \text{для } H\text{-хвиль}.$$

Висновки

На закінчення відмітимо, що знаходження рішень рівнянь (4), (10), (12) являється демонстрацією узагальненості підходу, розробленого К.М. Цибизовим, а в якості прикладів дослідженні резонатори змінного поперечного перерізу, вказані на рис.1, 2, 3. Поперечний розподіл електромагнітного поля в бочкоподібних поглинаючих резонаторах не відрізняється від поперечного розподілу поля в полому циліндричному резонаторі, а продольний розподіл поля визначається виразами (7), (11), (13).

Література

1. Коган Н.Л., Машковцев Ю.М., Цибизов К.Н. Сложные волноводы системы. Л: Судпромгиз, 1963. 351 с.
2. Машковцев Б.М., Цибизов К.Н., Емелин Б.Ф. Теория волноводов. М: Наука, 1966. 356 с.
3. Почерняев В.Н., Цибизов К.Н. Теория сложных волноводов. Киев: Науковий Світ, 2003. 224с.
4. Фок В.А. Проблемы теории дифракции и распространения электромагнитных волн. М: Сов. радио, 1970. 536
5. Кисунько Г.В. Электродинамика полых систем. Л: ВАС, 1949. 469 с.

References

1. Kogan N.L., Mashkovtsev Yu.M., Tsibizov K.N. Slozhnyie volnovodyi sistemyi. L: Sudpromgiz, 1963. 351 s.
2. Mashkovtsev B.M., Tsibizov K.N., Emelin B.F. Teoriya volnovodov. M: Nauka, 1966. 356 s.
3. Pochernyaev V.N., Tsibizov K.N. Teoriya slozhnyih volnovodov. Kiev: Naukoviy Svlt, 2003. 224s.
4. Fok V.A. Problemyi teorii difraktsii i rasprostraneniya elektromagnitnyih voln. M: Sov.radio, 1970. 536 s.
5. Kisunko G.V. Elektrodinamika polyih sistem. L: VAS, 1949. 469 s.

УДК 621.372.8

МОДЕЛЮВАННЯ БАГАТОПРОМЕНЕВОГО КАНАЛУ ЗВ'ЯЗКУ

DOI 10.36994 / 2788-5518-2023-01-05-05

Почерняєв В.М., д.т.н., професор, Національна академія Служби безпеки України, Київ, Україна. vpochernyaev@gmail.com

Сушанова В.С., к.т.н., Київський фаховий коледж зв'язку, Київ, Україна. povviktoriya@gmail.com

Анотація. У статті здійснено моделювання багатопроменевого каналу зв'язку з двома діапазонами або піддіапазонами (інтервалами) частот, системна функція якого виражається у замкнутому вигляді через еліптичні функції. Багатопроменевість призводить до того, що завжди існує певна кількість помилок, які характеризують нижню межу ймовірності помилок незалежно від застосовуваних методів завадостійкого кодування та необмеженого збільшення енергетичного потенціалу радіолінії. Математичне моделювання багатопроменевого каналу зв'язку передбачає спільний облік тимчасових затримок та доплерівського зсуву частоти сигналу, що приймається. Показано зв'язок між передавальною функцією $H(\omega, t)$ та системною функцією $S(\tau, \omega)$. У роботі вперше застосовано апарат еліптичних функцій та еліптичних інтегралів до визначення системної функції, яка у відомих літературних джерелах визначається за допомогою математичного апарату статистичного аналізу. Показано застосування такого підходу до мобільних радіосистем, що працюють у декількох частотних інтервалах, як, наприклад, комбінованої мобільної цифрової тропосферно-радіорелейної станції або цифрової тропосферної станції. Їхня робота здійснюється у двох діапазонах частот або двох піддіапазонах частот (передача/прийм). Тому побудувати математичну модель багатопроменевого каналу зв'язку з визначенням системних функцій у даному випадку стає досить складним завданням. Ця ситуація ускладнюється, якщо застосовується система адаптації за частотою, коли для боротьби із замираннями та зосередженими перешкодами використовується кілька інтервалів частот, у найпростішому випадку – два інтервали. У статті викладено підхід щодо подання системних функцій через еліптичні функції, що дозволяє у явному вигляді визначити передавальну функцію багатопроменевого каналу зв'язку. Для розглянутого випадку двочастотних інтервалів кінцевий вираз записано через тета-функції та повний еліптичний інтеграл першого роду.

Ключові слова: багатопроменевий канал зв'язку, частотно-рознесений прийом, частотно-селективні замирання, системна функція, передавальна функція, еліптична функція, тета-функція, комбінована мобільна цифрова тропосферно-радіорелейна станція, цифрова тропосферна станція.

SIMULATION OF A MULTIPATH COMMUNICATION CHANNEL

Pochernyaev Vitaly, Doctor of Technical Sciences, Prof., National Academy of the Security Service of Ukraine, Kyiv, Ukraine. vpochernyaev@gmail.com

Sushanova Viktoriia, Candidate of technical sciences, Kyiv Professional College of Communication, Kyiv, Ukraine. povviktoriya@gmail.com

Abstract. In the article, the simulation of a multipath communication channel with two bands or subbands (intervals) of frequencies, the system function of which is expressed in a closed form through elliptic functions. A multipath communication channel is a linear system with time various parameters. Multipath leads to the fact that there is always a certain number of errors. They characterize the lower limit of the probability of errors, regardless of the applied methods of error-correcting coding and an unlimited increase in the energy potential of the radio link. Mathematical simulation of a multipath communication channel provides the joint consideration of time delays and Doppler frequency shift of the received signal. The correlation between the transfer function $H(\omega, t)$ and the system function $S(\tau, \omega)$ is shown. In the article at the first time, the mathematical theory of elliptic functions and elliptic integrals is applied to the definition of a system function, which in known literature sources is determined using the mathematical methods of statistical analysis. In article is shown the application an approach to mobile radio systems, which are operating in several frequency bands, such as a combined mobile digital troposcatter-radiorelay station or a digital troposcatter station. Their work is carried out in two frequency bands or two frequency subbands (transmit / receive). Therefore, in this case it becomes very difficult to create a mathematical model of a multipath communication channel with the definition of system functions. This situation is complicated if a frequency adaptation system is used. At the same time, several frequency intervals are used to combat fading and interferences, in the simplest case – two intervals. The article describes the approach to the representation of system functions through elliptic functions. This allows to have close form of the transfer function for the

multipath communication channel. For the considered case of two-frequency intervals, the final expression is written in terms of theta-functions and the complete elliptic integral of the first kind.

Keywords: *multipath communication channel, frequency-spaced reception, frequency-selective fading, system function, transfer function, elliptic function, theta-function, combined mobile digital troposcatter-radiorelay station, digital troposcatter station.*

Вступ

При моделюванні багатопробієвих каналів зв'язку досліджують апарат системних функцій. Для реальних каналів зв'язку системні функції, як правило, не стаціонарні і для їх опису адекватним математичним апаратом є статистичний аналіз [1]. Однак застосування цього математичного апарату ускладнює вирішення великого числа конкретних практичних завдань. Це ускладнення посилюється, коли доводиться боротися не тільки з частотно-селективними замираннями, але і з зосередженими завадами сучасними способами, як наприклад, повторна передача однієї і тієї ж інформації на декількох частотах, що відповідають різним інтервалам частот. Тому необхідно представити системні функції багатопробієвих каналів зв'язку з повторною передачею інформації на декількох частотах за допомогою еліптичних інтегралів і еліптичних функцій. Професор К.М. Цибізов у своїх роботах використовував апарат спеціальних функцій математичної фізики. В даній роботі, що присвячена великому Вчителю, кінцевий результат з моделювання багатопробієвого каналу зв'язку отримано за допомогою апарата спеціальних функцій – еліптичних інтегралів першого роду, повних еліптичних інтегралів першого роду, еліптичних функцій Якобі, тета-функцій.

Метою роботи – є моделювання багатопробієвого каналу зв'язку з двома діапазонами або піддіапазонами (інтервалами) частот, системна функція якого виражається в замкнутому вигляді через еліптичні інтеграли та еліптичні функції Якобі.

Основний зміст

Один із ключових напрямів розвитку цифрових тропосферних систем зв'язку – збільшення швидкості інформації, що передається. Це вимагає подальшого вдосконалення моделей багатопробієвого каналу зв'язку та методик розрахунку характеристик багатопробієвості. Багатопробієвий канал зв'язку є лінійною системою з параметрами, що змінюються в часі, вихідний сигнал $y(t)$ якого пов'язаний з вхідним сигналом $x(t)$ наступною залежністю [1]:

$$y(t) = \int_{-\infty}^{\infty} h(t, \tau) x(t - \tau) d\tau, \quad (1)$$

де $h(t, \tau)$ - імпульсна реакція каналу, яка представляє відгук каналу в момент часу t на імпульс, поданий на вхід каналу в момент часу $(t - \tau)$ і яка є дійсною величиною. З фізичної точки зору імпульсна реакція каналу являє собою залежність середньої потужності сигналу, що приймається від часу τ при випромінюванні імпульсу, тому іноді званої енергетичним спектром затримок. Для багатопробієвого каналу зв'язку справедлива наступна залежність вихідного сигналу $y(t)$ від комплексного спектру $X(\omega)$ вхідного сигналу $x(t)$ [1]:

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega, t) X(\omega) \exp(j\omega t) d\omega, \quad (2)$$

де $H(\omega, t)$ – передавальна функція каналу, що характеризує залежний від часу відгук каналу на імпульс частоти $\exp(j\omega t)$.

З виразів (1) та (2) випливає:

$$h(t, \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega, t) \exp(j\omega \tau) d\omega,$$

$$H(\omega, t) = \int_{-\infty}^{\infty} h(t, \tau) \exp(-j\omega \tau) d\tau.$$

Багатопроменевість каналу зв'язку проявляється в тому, що сигнал від передавача до приймача проходить відстань різними шляхами. Сигнал для кожного шляху в момент часу t визначатиметься добутком $h_i(t, \tau)x(t - \tau)$, а сумарний сигнал виразом (1). Цей вираз можна вважати границею суми зважених та затриманих відгуків на сигнал $x(t)$, серед якої і-відгук, що має затримку в інтервалі від τ до $(\tau + d\tau)$, з'являється на виході каналу з вагою $h_i(t, \tau)d\tau$. Вираз (2) можна вважати границею суми відгуків каналу на різні спектральні компоненти сигналу, кожна з яких займає смугу частот від ω до $(\omega + d\omega)$ і множиться на відповідне значення передавальної функції $H(\omega, t)$ [1].

Зазначимо, що багатопроменевість призводить до того, що завжди існує певне число помилок, які характеризують нижню границю ймовірності помилок незалежно від застосовуваних методів завадостійкого кодування та необмеженого збільшення енергетичного потенціалу радіолінії.

Математичне моделювання багатопроменевого каналу зв'язку повинно передбачати спільне врахування часових затримок і доплерівського зсуву частоти сигналу, що приймається. Спектр доплерівських частот $\tilde{\omega}$ імпульсної реакції каналу $h(t, \tau)$ характеризується системною функцією $\tilde{S}(\tau, \tilde{\omega})$ [2]:

$$\tilde{S}(\tau, \tilde{\omega}) = \int_{-\infty}^{\infty} h(t, \tau) \exp(-j\tilde{\omega}t) dt, \quad (3)$$

$$h(t, \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{S}(\tau, \tilde{\omega}) \exp(-j\tilde{\omega}t) d\tilde{\omega}. \quad (4)$$

Якщо вважати канал зв'язку частотно-часовим фільтром [1], то вираз для сигналу на виході каналу можна одержати, підставивши (4) у (1):

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(t - \tau) \tilde{S}(\tau, \tilde{\omega}) \exp(j\tilde{\omega}t) d\tilde{\omega} d\tau. \quad (5)$$

Фізична суть виразу (5) полягає в тому, що суму затриманих і зсунутих за частотою відгуків через багатопроменевість каналу зв'язку можна інтерпретувати як відгук із затримкою в інтервалі від τ до $(\tau + d\tau)$ та доплерівським зсувом частоти в інтервалі $\tilde{\omega}$ до $(\tilde{\omega} + d\tilde{\omega})$ з вагою $\tilde{S}(\tau, \tilde{\omega})$.

Зв'язок між передавальною функцією $H(\omega, t)$ і системною функцією $\tilde{S}(\tau, \tilde{\omega})$ визначається подвійним перетворенням Фур'є:

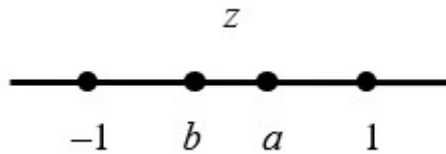
$$H(\omega, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{S}(\tau, \tilde{\omega}) \exp(-j(\omega\tau - \tilde{\omega}t)) d\tilde{\omega} d\tau. \quad (6)$$

Стан багатопроменевого каналу зв'язку випадковим чином змінюється в області визначення змінних $t, \tau, \omega, \tilde{\omega}$, які є аргументами передавальної та системної функції, що зв'язують сигнали на вході та виході каналу зв'язку.

Для визначення системної функції використовують математичний апарат статистичного аналізу. Однак, якщо робота радіотехнічної системи, такої, наприклад, як комбінованої мобільної цифрової тропосферно-радіорелейної станції (МЦТрРПС) або цифрової тропосферної станції (ЦТрС), здійснюється у двох діапазонах частот або двох піддіапазонах (інтервалах) частот (передавання/приймання), то побудувати математичну модель багатопроменевого каналу зв'язку з визначенням системних функцій у цьому разі стає вельми важким завданням.

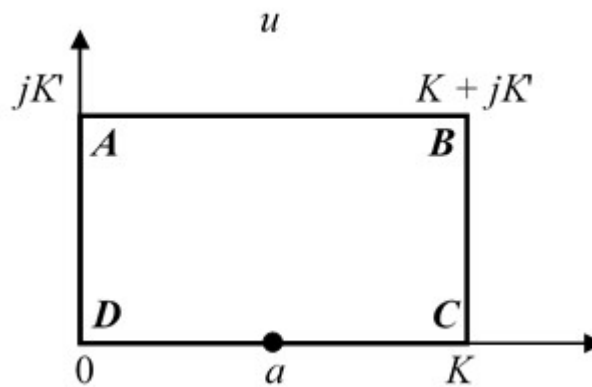
Ця ситуація ускладнюється, якщо застосовується система адаптації за частотою, коли для боротьби із завмираннями і зосередженими завадами використовується кілька інтервалів частот, у найпростішому випадку - два інтервали. Така система адаптації за частотою застосовується в деяких ЦТрС і реалізується в МЦТрРПС.

Спочатку задамо на дійсній осі площини z (рис. 1) два кінцеві інтервали: $(-1, b)$, $(a, 1)$.

Рис.1. Кінцеві інтервали дійсної осі z

Фізична інтерпретація цих двох інтервалів - частотні піддіапазони ЦТРС, в одному з яких здійснюється передавання сигналів, а в другому - їх приймання. Частотні піддіапазони рівноправні з погляду передавання і приймання сигналів. Так само функціонує і тропосферна компонента МЦТРРС.

Для введення еліптичних функцій відобразимо конформно верхню половину площини z на прямокутник площини u як це показано на рис.2.

Рис.2. Прямокутник площини u

Для цього знайдемо величину k ($0 < k < 1$) за такою формулою:

$$k^2 = \frac{2(a-b)}{(1+a)(1-b)}.$$

Враховуючи, що $b = 1 - 2cn^2a/dn^2a$, формулою конформного відображення є:

$$z = \frac{sn^2a \cdot cn^2u + cn^2a \cdot sn^2u}{sn^2a - sn^2u}, \quad (7)$$

де sna , snu , cna , cnu - еліптичні функції Якобі.

При цьому точки $-1, b, a, 1 = z$ переходять відповідно в точки $K, K + jK', jK', 0 = u$. Відповідно до [2] символи K і K' позначають повні еліптичні інтеграли першого роду для модуля k і додаткового модуля $k' = \sqrt{1 - k^2}$. Справедливі такі співвідношення:

$$k^2 = m, k'^2 = m_1$$

де m і m_1 - параметр і додатковий параметр теорії еліптичних функцій відповідно.

На підставі формули (7) отримаємо:

$$\frac{du}{dz} = \frac{\sqrt{1+a}\sqrt{1-b}}{2\sqrt{(z^2-1)(z-a)(z-b)}},$$

$$z = 1 + \frac{sna \cdot cna}{dna} \left\{ \frac{\theta'(a+u)}{\theta(a+u)} + \frac{\theta'(a-u)}{\theta(a-u)} - 2 \frac{\theta'(a)}{\theta(a)} \right\},$$

де dna - еліптична функція Якобі;

$\frac{\theta'(a+u)}{\theta(a+u)}, \frac{\theta'(a-u)}{\theta(a-u)}, \frac{\theta'(a)}{\theta(a)}$ - логарифмічні похідні тета-функцій [2].

Важливість тета-функцій полягає в тому, що кожна з еліптичних функцій Якобі може бути виражена як відношення двох тета-функцій:

$$\begin{aligned} sna &= \frac{1}{\sqrt{k}} \frac{\theta_1(a/2K)}{\theta_0(a/2K)}, \\ cna &= \sqrt{\frac{k'}{k}} \frac{\theta_2(a/2K)}{\theta_0(a/2K)}, \\ dna &= \sqrt{k'} \frac{\theta_3(a/2K)}{\theta_0(a/2K)}. \end{aligned}$$

Тета-функції можна визначити такими розкладаннями:

$$\begin{aligned} \theta_0(z) &= 1 - 2q \cos 2\pi z + 2q^4 \cos 4\pi z - 2q^9 \cos 6\pi z + \dots + (-1)^n 2q^{n^2} \cos 2n\pi z + \dots, \\ \theta_1(z) &= 2q^{\frac{1}{4}} \sin \pi z - 2q^{\frac{9}{4}} \sin 3\pi z + 2q^{\frac{25}{4}} \sin 5\pi z - \dots + (-1)^n 2q^{\left(n+\frac{1}{2}\right)^2} \sin \sin (2n+1)\pi z + \dots, \\ \theta_2(z) &= 2q^{\frac{1}{4}} \cos \pi z - 2q^{\frac{9}{4}} \cos 3\pi z + 2q^{\frac{25}{4}} \cos 5\pi z - \dots + (-1)^n 2q^{\left(n+\frac{1}{2}\right)^2} \cos \cos (2n+1)\pi z + \dots, \\ \theta_3(z) &= 1 + 2q \cos 2\pi z + 2q^4 \cos 4\pi z + 2q^9 \cos 6\pi z + \dots + 2q^{n^2} \cos 2n\pi z + \dots, \end{aligned}$$

де $z = a/2K$

Записавши системну функцію в замкненому вигляді у формулі (6), можна знайти і явний вигляд передавальної функції $H(\omega, t)$. Для цього треба знайти функцію відображення $\hat{S}(z)$, яка є відображенням системної функції $\tilde{S}(\tau, \tilde{\omega})$ у комплексній площині z . Функція $\hat{S}(z)$ має вигляд:

$$\hat{S}(z) = \int_1^z \frac{dz}{\sqrt{(z^2 - 1)(z - a)(z - b)}}.$$

Функція $\hat{S}(z)$ обертається на нескінченність лише на нескінченності, де має полюс першого порядку, а логарифмічної особливості не має. Тому, функцію $\hat{S}(z)$ можна представити лінійною комбінацією еліптичних інтегралів або через комбінацію тета-функцій. Тоді, з урахуванням перетворень маємо:

$$\hat{S}(z) = -\frac{1}{p} \left\{ \frac{\pi u}{K} + K' \left[\frac{\theta'(a+u)}{\theta(a+u)} - \frac{\theta'(a-u)}{\theta(a-u)} \right] \right\}.$$

Параметр p знаходиться за умови, що $u \rightarrow a, |z| \rightarrow \infty$. Тоді маємо:

$$p = \frac{K' dna}{sna \cdot cna}.$$

Еліптичні інтеграли можна обчислити за таблицями за допомогою рядів за допомогою номограм. Повні еліптичні інтеграли, у яких модуль k є комплексною змінною, обчислюються за допомогою таблиць і шляхом розкладання в ряди [3].

Зазначимо, що не слід плутати параметр p з параметром Якобі q [2], який дорівнює $\exp(-\pi K'/K)$ і за яким тета-функція розкладається в ряд.

Зазначимо, що цифрові радіорелейні станції (ЦРРС) працюють в одному діапазоні частот, у якому здійснюють передавання і приймання сигналів. Це ж поширюється і на багатодіапазонні ЦРРС. Завдання подання системної функції багатопроменового каналу зв'язку для одного частотного інтервалу спроститься. Припустимо, що на комплексній площині z задано інтервал (a, b) . Тоді, не порушуючи узагальненості, запишемо:

$$a = k^2 (0 < k < 1), b = 1.$$

За допомогою конформного відображення маємо:

$$a = \frac{1}{sn^2(u, k)},$$

$$\frac{du}{da} = -\frac{1}{2\sqrt{a(a-k^2)(a-1)}}$$

Тоді, функція $\hat{S}(z)$ має такий вигляд:

$$\hat{S}(z) = -\frac{\theta(0)\theta(u-a)}{\theta(u)\theta(a)} e^{\frac{aK'}{K}},$$

де $\theta(0), \theta(u-a), \theta(u), \theta(a)$ - тета-функції.

Нехай повторна передача здійснюється в інтервалах частот $[a_1, b_1], [a_2, b_2], \dots [a_n, b_n]$. У комплексній площині $\tilde{\omega}$ введемо раціональні функції $R(\tilde{\omega})$ і $P(\tilde{\omega})$:

$$R(\tilde{\omega}) = \tilde{\omega}(\tilde{\omega} - a_1)(\tilde{\omega} - b_1) \dots (\tilde{\omega} - b_n),$$

$$P(\tilde{\omega}) = (\tilde{\omega} - a_1)(\tilde{\omega} - a_1) \dots (\tilde{\omega} - a_n).$$

Системні функції можна записати в комплексній площині у наступному вигляді:

$$Q_1(\xi, \tilde{\omega}) = \cos \cos(\xi \sqrt{\tilde{\omega}}) + j \frac{\sqrt{R(\tilde{\omega})} \sin \sin(\xi \sqrt{\tilde{\omega}})}{P(\tilde{\omega}) \sqrt{\tilde{\omega}}}, \quad (8)$$

$$Q_2(\xi, \tilde{\omega}) = \cos \cos(\xi \sqrt{\tilde{\omega}}) - j \frac{\sqrt{R(\tilde{\omega})} \sin \sin(\xi \sqrt{\tilde{\omega}})}{P(\tilde{\omega}) \sqrt{\tilde{\omega}}}. \quad (9)$$

Системні функції не є цілими, але мають по одному кореню в кожному інтервалі частот. Якщо функції $\ln \ln Q_1(\xi, \tilde{\omega})$ і $\ln \ln Q_2(\xi, \tilde{\omega})$ мають простий полюс в точці, то логарифмічними особливими точками з від'ємними відніманнями відрахуваннями є тільки точки $\tilde{\omega} = \infty$. Тому, число нулів цих функцій дорівнює n , і всі лежать по одному в кожному із зазначених вище інтервалів. Ці властивості виправдовують введення такого представлення системних функцій і вказують на апарат еліптичних інтегралів і еліптичних функцій, яким необхідно скористатися [2].

Якщо функції $Q_1(\xi, \tilde{\omega})$ і $Q_2(\xi, \tilde{\omega})$ в (1, 2) є раціональними від ξ і $\tilde{\omega}$, де $\tilde{\omega}$ є многочленом третьої або четвертої ступеня від ξ , то $\int Q_1(\xi, \tilde{\omega}) d\xi$ і $\int Q_2(\xi, \tilde{\omega}) d\xi$ називаються еліптичними. Як відомо, еліптичні інтеграли в загальному випадку не виражаються через елементарні функції [2,3].

Розглянутий випадок передачі і прийому інформації на декількох частотах по суті є частотно-рознесеним прийомом, який є одним із способів боротьби з завадами сигналу і зосередженими завадами. Найбільш поширеним способом боротьби з завадами в системах тропосферного зв'язку є просторово-рознесений прийом, наприклад [4]. В цьому випадку тропосферна станція має, як правило, дві антени, що працюють на передачу і прийом, але випромінювання і прийом сигналу здійснюється у своїх підсмугах частот в кожному на одній

частоті. Для каналів зв'язку таких систем тропосферного зв'язку явний вид системної функції спрощується. Якщо прийняти $a_1 = k^2 (0 < k < 1)$ і $b_1 = 1$, то за допомогою конформного відображення:

$$\tilde{\omega} = \frac{1}{sn^2(u, k)}, \quad \frac{du}{d\tilde{\omega}} = -\frac{1}{2\sqrt{\tilde{\omega}(\tilde{\omega}-k^2)(\tilde{\omega}-1)}} \quad (10)$$

комплексна площина відображається на прямокутник [2,3]. У вираженні (10) sn є еліптичною функцією Якобі.

Тепер системна функція може бути представлена через еліптичні функції і еліптичні інтеграли:

$$Q(\xi, \tilde{\omega}) = \frac{\theta(0)\theta(u-j\xi)}{\theta(u)\theta(j\xi)} e^{j\xi K'/K}, \quad (11)$$

де θ - тета-функція; K'/K - повні еліптичні інтеграли першого роду.

Для обчислення повного еліптичного інтегралу першого роду можна скористатися розкладаннями в ряд. В даному випадку величина K приймає наступний вигляд:

$$K = \frac{\pi}{2} \left\{ 1 + \left(\frac{1}{2}\right)^2 k^2 + \left(\frac{1}{2} \cdot \frac{3}{4}\right)^2 k^4 + \dots + \left(\frac{1 \cdot 3 \cdot \dots \cdot (2n-1)}{2 \cdot 4 \cdot \dots \cdot 2n}\right)^2 k^{2n} + \dots \right\}.$$

Якщо значення k близьке до 1, то вигідніше користуватися розкладанням K в ряд за степеню $k' = \sqrt{1-k^2}$:

$$K = a_0 + \left(\frac{1}{2}\right)^2 a_1 k'^2 + \left(\frac{1}{2} \cdot \frac{3}{4}\right)^2 a_2 k'^4 + \left(\frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\right)^2 a_3 k'^6 + \dots$$

Коефіцієнти $a_0, a_1, a_2, \dots, a_n$ визначаються рівностями:

$$a_0 = \ln \frac{4}{k'}, a_1 = a_0 - 1, \dots, a_n = a_{n-1} - \frac{2}{(2n-1)^2}.$$

Якщо виявиться, що $|k| > 1$, то можна скористатися наступним розкладанням:

$$K = \frac{\pi}{2k} \left\{ 1 + \left(\frac{1}{2}\right)^2 \frac{1}{k^2} + \left(\frac{1}{2} \cdot \frac{3}{4}\right)^2 \frac{1}{k^4} + \dots + \left(\frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\right)^2 \frac{1}{k^6} + \dots \right\} +$$

$$+ j \frac{1}{k} \left\{ \ln 4k + \left(\frac{1}{2}\right)^2 (\ln 4k - 1) \frac{1}{k^2} + \left(\frac{1}{2} \cdot \frac{3}{4}\right)^2 \left(\ln 4k - \frac{7}{6}\right) \frac{1}{k^4} + \dots + \left(\frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\right)^2 \left(\ln 4k - \frac{31}{30}\right) \frac{1}{k^6} + \dots \right\}$$

Важливо відзначити, що в (11) для окремого випадку системна функція виражається через тета-функцію. Це пояснюється тим, що кожна з еліптичних функцій Якобі може бути виражена як відношення двох тета-функцій [2].

Висновки

Таким чином, викладений у роботі підхід щодо подання системних функцій через еліптичні функції дає змогу в явному вигляді визначити передавальну функцію багатопробеневого каналу зв'язку. Для розглянутого випадку двочастотних інтервалів кінцевий вираз записано через тета-функції і повний еліптичний інтеграл першого роду.

На закінчення відзначимо, що канал зв'язку з частотно-селективними замираннями і схильний до зосереджених перешкод може мати місце і на радіорелейних лініях зв'язку, а не тільки тропосферних. У роботах [5,6] показана перспективність створення комбінованих засобів зв'язку, наприклад, МЦТрРРС. Тому, математичний опис багатопробневих каналів зв'язку за

допомогою системних функцій, що мають явний вигляд і виражених через еліптичні інтеграли і еліптичні функції актуальний для різноманітних каналів зв'язку з багатопроменевістю.

Література

1. J.G.Proakis. Digital communications. NY:McGraw-Hill, 1995. 800p.
2. Milton Abramowitz and Irene A. Stegun. Handbook of Mathematical Functions With Formulas, Graphs and Mathematical Tables, Wiley-Interscience, 1993.
3. Jahnke E., Emde F., Lösch F. Tables of higher functions. NY:McGraw-Hill, 1960.
4. Почерняев В.М., Повхліб В.С. Мобільна цифрова станція НВЧ діапазону подвійного призначення. *Наукові праці ОНАЗ ім. О.С.Попова*. 2014. № 2. С. 76-82.
5. Почерняев В.М., Повхліб В.С. Стан і напрямки розвитку мобільних цифрових радіорелейних систем. *Системи озброєння і військова техніка*. 2018. №1(53). С.183-188.
6. Почерняев В.М., Повхліб В.С. Стан і напрямки розвитку мобільних цифрових тропосферних систем зв'язку. *Системи озброєння і військова техніка*. 2018. №2(54). С.51-59.

References

1. J.G.Proakis. Digital communications. NY:McGraw-Hill, 1995. 800p.
2. Milton Abramowitz and Irene A. Stegun. Handbook of Mathematical Functions With Formulas, Graphs and Mathematical Tables, Wiley-Interscience, 1993.
3. E. Jahnke, F. Emde, and F. Lösch, Tables of higher functions, McGraw-Hill, New York, 1960.
4. Pochernyaev V.M., Povkhlіb V.S. Mobilna tsifrova stantsiya NVCh diapazonu podviinogo priznachennya. *Naukovi pratsi ONAZ im. O.S.Popova*. 2014. № 2. S. 76-82.
5. Pochernyaev V.M., Povkhlіb V.S. Stan i napryamki rozvitku mobilnikh tsifrovikh radiorelelnikh sistem. *Sistemi ozbroennya i viiskova tekhnika*. 2018. №1(53). S.183-188.
6. Pochernyaev V.M., Povkhlіb V.S. Stan i napryamki rozvitku mobilnikh tsifrovikh troposfernikh sistem zv'yazku. *Sistemi ozbroennya i viiskova tekhnika*. 2018. №2(54). S.51-59.

УДК. 621.372

СУМАТОР ПОТУЖНОСТЕЙ НА БАЗІ КВАЗІОПТИЧНОГО ВІДКРИТОГО РЕЗОНАТОРА ЯК ВИСОКОСТАБІЛЬНИЙ ГЕНЕРАТОР ТЕРАГЕРЦЕВОГО ДІАПАЗОНУ

DOI 10.36994 / 2788-5518-2023-01-05-06

Наритник Т.М., к.т.н., проф. Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна. t.narytnyk@ukr.net

Кузьмичов І.К. Харківський національний університет імені В.Н.Каразіна, Харків, Україна. kuzmichev.igr@i.ua

Авдєєнко Г.Л. Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна. e-mail:t.narytnyk@ukr.net

Музичишин Б.І. Інститут радіофізики та електроніки ім. А. Я. Усікова НАН України, Харків, Україна. feuntabod@gmail.com

Май О.В. Національний технічний університет України Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна. avmay@ukr.net

Алісов А.О. Національний технічний університет України Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна. t.narytnyk@ukr.net

Анотація: В роботі досліджено ефективність збудження аксіально-несиметричного типу коливань TEM_{10q} в квазіоптичному відкритому резонаторі суматора потужностей (ВРСП) за допомогою хвильової хвилі TE_{20} . За допомогою функціоналу, використовуваного в теорії зеркальних антен для визначення коефіцієнта використання поверхні, оцінено ефективність збудження розглянутого типу коливань у напівсферичному резонаторі за допомогою прямокутного хвильоводу перетином $a \times b$, в якому поширюється хвиля TE_{20} . З проведених розрахунків зроблено висновок, що близько 87% потужності, яка підводиться по хвильоводу до резонатора, йде на збудження коливання TEM_{10q} . Отже, така резонансна система повинна мати хороші селектуючі властивості. Для підтвердження цього проведені додаткові чисельні дослідження. Встановлено, що при узгодженому збудженні коливань в ВРСП, останній має кутову і частотну селекцією спектру. Таким чином, при правильному виборі поперечних розмірів прямокутного хвильоводу з хвилею TE_{20} , розташованого в центрі плоского дзеркала ВРСП, останній матиме одночастотний відгук у широкій смузі частот і можна забезпечити високий коефіцієнт складання потужностей в терагерцовому діапазоні. При розміщенні джерел у кожному з двох прямокутних хвильоводів, що переходять у випромінювачі наведених вище розмірів, ми отримаємо суматор потужності на основі напівсферичного відкритого резонатора. Запропонована і досліджена електродинамічна система з двох прямокутних хвильоводів визначених поперечних розмірів, яка дозволяє регулювати величину зв'язку джерел із резонатором. Для цього на плоскому дзеркалі ВРСП розташована E -поляризована дріт'яна дифракційна решітка, змінюючи період якої регулюється зв'язок резонатора з джерелами. Встановлено, що вибираючи параметри одиниць дифракційних решіток (E -поляризація) у розкритих елементах зв'язку, можна забезпечити будь-який наперед заданий зв'язок окремих джерел з резонансним об'ємом. Проведено експериментальні дослідження ВРСП на частоті 130 ГГц. Виміряна потужність на його виході перебувала в межах (10...13 дБм), що відповідає ефективності складання потужностей $k=0,76$. Результати проведених досліджень дозволяють зробити висновок про перспективність використання розробленого ВРСП як високостабільного генератора терагерцового діапазону.

Ключові слова: терагерцовий діапазон, відкритий резонатор суматор потужностей, прямокутний хвильовід, ефективність збудження, високостабільний генератор.

RESEARCH OF A QUASIOPTICAL OPEN RESONATOR FOR POWER SUMMATION IN A HIGHLY STABLE GENERATOR OF THE TERAHERTZ RANGE

Teodor Narytnyk, Ph.D., Prof. National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" Kyiv, Ukraine. t.narytnyk@ukr.net

Igor Kuzmichev. V.N.Karazin Kharkiv National University, Kharkiv, Ukraine. t.narytnyk@ukr.net

Hlib Avdieienko. National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" Kyiv, Ukraine. t.narytnyk@ukr.net

Bohdan Muzychishin. The A. Ya. Usikov's Institute for Radio Physics and Electronics of the NAS of Ukraine, Kharkov, Ukraine. kuzmichev.igr@i.ua

Oleksandr May. National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" Kyiv, Ukraine. avmay@ukr.net

Andrey Alisov. National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" Ukraine, Kyiv, Ukraine. t.narytnyk@ukr.net

Abstract: The efficiency of excitation of the axial-asymmetric vibration TEM_{10q} in a quasioptical open resonator of the power combiner (ORSM) by means of the TE_{20} waveguide wave is investigated. Based on the theory of mirror antennas, the efficiency of excitation of TEM_{10q} oscillations in a hemispherical resonator using a rectangular waveguide with a cross-section in which the TE_{20} wave propagates is estimated. Based on the calculations, it was concluded that about 87% of the power supplied to the resonator via the waveguide is used to excite the oscillation under consideration. Therefore, such a resonant system should have good selective properties. Additional studies were conducted to confirm this. It was established that with coordinated excitation of oscillations in the VRSP, the latter has angular and frequency spectrum selection. With the correct selection of the transverse dimensions of a rectangular waveguide with a TE_{20} wave, located in the center of a VRSP flat mirror, the latter will have a single-frequency response in a wide frequency band and a high power stacking factor in the terahertz range can be provided. When placing the sources in each of the two rectangular waveguides that pass into the emitters of the above dimensions, we will get a power adder based on a hemispherical VRSP. An electrodynamic system of two rectangular waveguides of certain cross-sectional dimensions is proposed and investigated, which allows adjusting the amount of coupling between the sources and the resonator. For this purpose, an E-polarized wire diffraction grating is located on the VRSP flat mirror, changing the period of which regulates the connection of the resonator with the sources. It was found that by choosing the parameters of one-dimensional diffraction gratings (E-polarization) in the openings of the communication elements, it is possible to ensure any predetermined connection of individual sources with a resonant volume. Experimental studies of VRSP at a frequency of 130 GHz were carried out. The measured power at its output was within (10...13 dBm), which corresponds to the efficiency of power assembly $k=0.76$. The results of the conducted research allow us to draw a conclusion about the prospects of using the developed VRSP in highly stable generators of the terahertz range.

Keywords: terahertz range, open resonator power combiner, rectangular waveguide, excitation efficiency, highly stable generator.

Вступ

Однією з тенденцій розвитку сучасної радіоелектроніки є освоєння терагерцового частотного діапазону, що займає проміжне положення між добре вивченими мікрохвильовими і оптичними ділянками спектра електромагнітного випромінювання. Переходом в терагерцовий діапазон вирішується, як питання збільшення швидкості передачі інформації за рахунок використання більш широких спектрів частот, так і питання забезпечення завадозахищеності та електромагнітної сумісності радіоелектронних засобів, а також виключається ймовірність впливу навмисних радіозавад [1,2]. Використання терагерцових хвиль для мобільного зв'язку є однією з ключових технологічних концепцій для впровадження 6G. Теорема Шеннона-Хартлі, яка дає границю для максимально досяжної пропускної здатності каналу (під впливом шуму) $C=B \cdot \log_2(1+S/N)$, де C - пропускна здатність в біт/с, B - смуга пропускання в Гц, а S/N - відношення сигнал/шум, стверджує, що максимально досяжна швидкість передачі даних масштабується пропорційно смузі пропускання B . Це було мотивацією для 5G розширити частотну область до діапазону міліметрового діапазону, оскільки частоти нижче 6 ГГц вже інтенсивно використовуються у мобільному, супутниковому, радіорелейному зв'язку, телерадіомовленні та системах широкосмугового радіодоступу. Для того, щоб ще більше збільшити у порівнянні з 5G швидкість передачі даних для майбутніх стандартів бездротового зв'язку, таких як 6G, ще більш високі частотні діапазони аж до ТГц спектру, де доступні великі суцільні смуги пропускання в декілька ГГц, потенційно можуть бути задіяні. У той час як 5G вперше використовував міліметрові частоти з великою пропускною здатністю до 400 МГц на одну несучу, щоб забезпечити швидкість передачі, необхідну для вимогливих додатків в реальному часі, технологія 6G, що зароджується, націлена на більш високі швидкості передачі даних і менші затримки.

Освоєння терагерцового діапазону становить значні труднощі насамперед через відсутність високостабільних генераторів із достатнім рівнем потужності та низьким рівнем фазового шуму [3,4]. Одним із можливих шляхів вирішення цієї проблеми є застосування суматорів потужності. Планарні лінії передачі в терагерцовому діапазоні, які дозволяють здійснити складання потужностей, втрачають свої переваги на цих частотах через втрати на випромінювання, втрати в підкладці та омичні втрати. У діапазоні надвисоких частот (НВЧ) найбільш широко знайшли своє

застосування мостові схеми та складання потужностей джерел електромагнітних коливань в єдиному резонансному об'ємі [5], а в терагерцевому діапазоні - в квазіоптичних відкритих резонаторах [6-8].

Метою даної роботи є дослідження ефективності збудження аксіально-несиметричного коливання TEM_{10q} у квазіоптичному відкритому резонаторі суматора потужностей (ВРСП) за допомогою хвильоводної хвилі TE_{20} та практична реалізація ВРСП у високостабільному генераторі терагерцевого діапазону.

Виконання досліджень

Оцінимо ефективність збудження коливання TEM_{10q} у напівсферичному резонаторі за допомогою прямокутного хвильоводу перетином $a \times b$, в якому поширюється хвиля TE_{20} [9,10]. Хвильовід розташований у центрі плоского дзеркала. Геометричні параметри резонатора та розподіл збуджуючого поля $\vec{E}_e(x, y)$ та поля коливання ВРСП $\vec{E}_w(x, y)$ наведено на рис. 1. Тут же показано орієнтацію векторів напруженості електричного поля у вхідному хвильоводі та в резонаторі.

Компоненти поперечного електричного поля TE_{mn} -хвиль у прямокутному хвильоводі перетином $a \times b$ у загальному випадку визначаються виразом [11]:

$$\vec{E}_{mn}(x, y, z) = C_0 \vec{e}_{mn}(x, y) e^{\pm i \zeta_{mn} z}. \quad (1)$$

Тут $C_0 = ik_0 W_0 \mu_0$, $k_0 = \omega \sqrt{\varepsilon_0 \mu_0}$, $[m^{-1}]$, $\omega = 2\pi f$ - кругова частота, $[Гц]$, $W_0 = \sqrt{\mu_0 / \varepsilon_0} = 120\pi$, $[Ом]$, $\varepsilon_0 = 10^{-9} / (36\pi)$, $[Ф/м]$, $\mu_0 = 4\pi \times 10^{-7}$, $[Гн/м]$, $\zeta_{mn} = \sqrt{k_0^2 - \chi_{mn}^2}$ - постійна поширення TE_{mn} -хвилі ($\text{Re } \zeta_{mn} \geq 0, \text{Im } \zeta_{mn} \geq 0$), $\chi_{mn}^2 = (m\pi/a)^2 + (n\pi/b)^2$, а поперечна власна функція TE_{mn} -хвилі має вигляд [11]:

$$\vec{e}_{mn}(x, y) = \frac{1}{N_{mn}} \left[-\frac{n\pi}{b} \cos \frac{m\pi x}{a} \sin \frac{n\pi y}{b} \vec{x}_0 + \frac{m\pi}{a} \sin \frac{m\pi x}{a} \cos \frac{n\pi y}{b} \vec{y}_0 \right], \quad (2)$$

де $N_{mn} = \chi_{mn} \sqrt{(1 + \delta_{m0})(1 + \delta_{n0})} \sqrt{ab}/2$ - константа нормування, що визначається з умови ортогональності функцій $\vec{e}_{mn}(x, y)$

$$\int_0^a \int_0^b (\vec{e}_{mn} \cdot \vec{e}_{qp}) dx dy = \delta_{mq} \delta_{np},$$

δ_{mq} та δ_{np} - символ Кронекера.

Розподіл електричної компоненти поля хвилі TE_{20} у прямокутному хвильоводі у площині $z=0$ отримуємо із (1), (2)

$$\vec{E}_e(x, y) = \vec{E}_{20}(x, y, 0) = C_0 \vec{e}_{20}(x, y). \quad (3)$$

Відбиванням хвиль від розкриття хвильоводу нехтуємо і вважаємо апертури дзеркал нескінченними. Запишемо розподіл електричної компоненти поля робочого коливання у площині $z=0$ [12],

$$\vec{E}_w(x, y) = 2A_{10} \frac{\sqrt{2}x}{w_0} \exp\left(-\frac{x^2 + y^2}{w_0^2}\right) \vec{y}_0, \quad (4)$$

де A_{10} - амплітудний коефіцієнт, w_0 - радіус плями поля основного коливання резонатора TEM_{00q} на плоскому дзеркалі.

З теорії дзеркальних антен [13,14] відомо, що для отримання високого коефіцієнта використання поверхні (далі не зустрічається) необхідно узгодити поля у фокальній площині рефлектора і в розкритті опромінювача. З фізичної точки зору це аналогічно узгодженню поля резонатора з полем хвилі, що поширюється хвильоводом, розташованим у центрі плоского

дзеркала відкритого резонатора (ВР). Для визначення ефективності збудження η коливання TEM_{10q} у резонаторі за допомогою хвилі TE_{20} прямокутного хвильоводу скористаємося виразами (3), (4) і співвідношенням [13-15]:

$$\eta = \frac{\left| \int_{-a/2}^{a/2} \int_{-b/2}^{b/2} \vec{E}_e(x, y) \vec{E}_w^*(x, y) dx dy \right|^2}{\|\vec{E}_e(x, y)\|^2 \|\vec{E}_w(x, y)\|^2} \quad (5)$$

Значок “*” вказує на комплексно-спряжену функцію. Співвідношення (5) показує, яка частка енергії, що підводиться хвильоводом до резонатора, йде на збудження робочого коливання. Величини $\|\vec{E}_e(x, y)\|^2$ і $\|\vec{E}_w(x, y)\|^2$ – це квадрати норм функцій збуджуючого і робочого полів, які визначаються виразами [15]

$$\|\vec{E}_e(x, y)\|^2 = \int_{-a/2}^{a/2} \int_{-b/2}^{b/2} \vec{E}_e(x, y) \vec{E}_e^*(x, y) dx dy, \quad (6)$$

$$\|\vec{E}_w(x, y)\|^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \vec{E}_w(x, y) \vec{E}_w^*(x, y) dx dy. \quad (7)$$

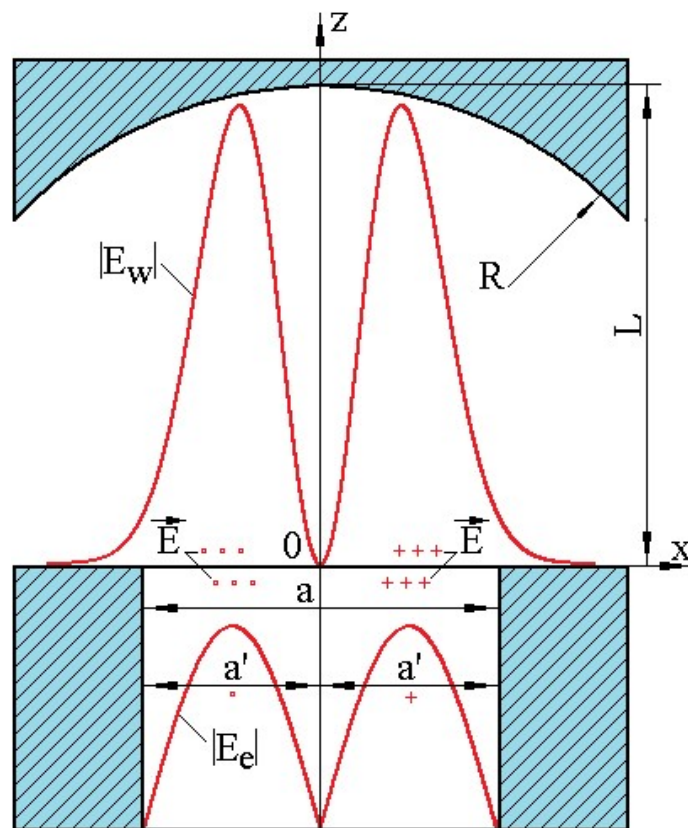


Рис.1. Напівсферичний ВРСР із відрізком прямокутного хвильоводу

Опускаючи проміжні викладки (див. [8]), запишемо вираз, що визначає ефективність збудження коливання TEM_{10q} в ВРСР за допомогою хвильоводної хвилі TE_{20} .

$$\eta = \frac{4\pi^3}{\tilde{a}^3 \tilde{b}} \Phi^2 \left(\frac{\tilde{b}}{2} \right) e^{-2\left(\frac{\pi}{\tilde{a}}\right)^2} \left[2 + e^{\left(\frac{\pi}{\tilde{a}}\right)^2 - \left(\frac{\tilde{a}}{2}\right)^2} (W(g) + W^*(g)) \right]^2, \quad (8)$$

де $\tilde{a} = a/w_0$, $\tilde{b} = b/w_0$, $\Phi(\tilde{b}/2)$ – інтеграл ймовірності, $g = (\pi/\tilde{a}) + j(\tilde{a}/2)$, $W(g)$ – інтеграл ймовірності комплексного аргументу. Залежність $\eta(\tilde{a}, \tilde{b})$ наведено на рис. 2. Як очевидно з малюнка, максимальна величина $\eta = 0,867$ досягається при нормованих розмірах елемента зв'язку $\tilde{a} = 3,338$ і $\tilde{b} = 1,980$.

З проведених розрахунків можна дійти висновку, що близько 87% потужності, яка підводиться по хвильоводу до резонатору, йде на збудження коливання, що нами розглядається. Отже, така резонансна система повинна мати хороші селектуючі властивості. Для підтвердження цього необхідно провести додаткові дослідження.

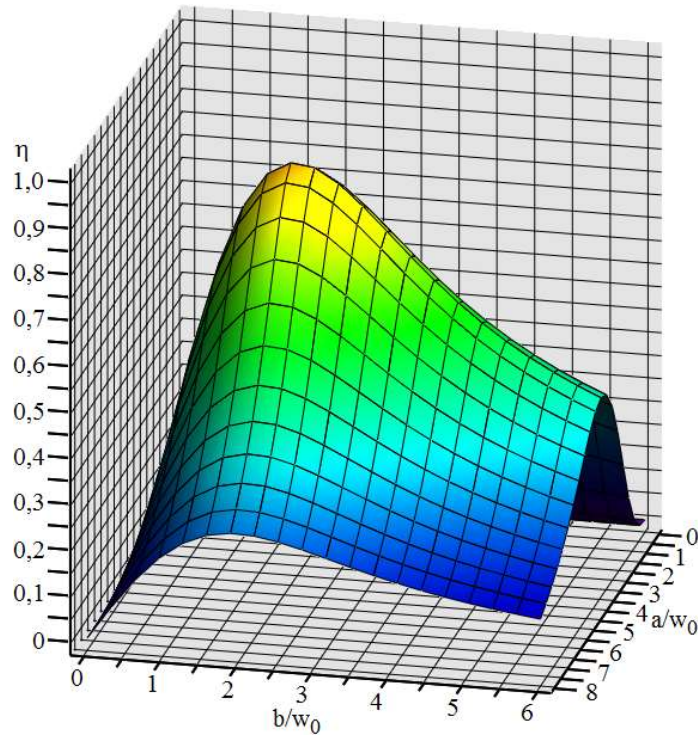


Рис. 2. Ефективність збудження коливання TEM_{10q} в ВРСП за допомогою хвильоводної хвилі TE_{20}

Розглянемо ефективність збудження за допомогою хвильоводної хвилі TE_{20} коливання TEM_{30q} , структура електричного поля якого найбільш близька до коливання, що аналізується. Збудливе поле визначається, як і вище, виразом (3). Розподіл електричної компоненти поля коливання ВРСП TEM_{30q} описується співвідношенням [12]:

$$\vec{E}_w(x, y) = A_{30} \left[8 \left(\frac{\sqrt{2}x}{w_0} \right)^3 - 12 \left(\frac{\sqrt{2}x}{w_0} \right) \right] \exp \left(-\frac{x^2 + y^2}{w_0^2} \right) \vec{y}_0 \quad (9)$$

Використовуючи вирази (5-7) знайдемо, що ефективність збудження коливання TEM_{30q} у напівсферичному ВРСП за допомогою хвильоводної хвилі TE_{20} , має вигляд:

$$\eta = \frac{\tilde{a}}{12\tilde{b}} \Phi^2 \left(\frac{\tilde{b}}{2} \right) \left| \int_{-1/2}^{1/2} \left[8(\sqrt{2}u\tilde{a})^3 - 12(\sqrt{2}u\tilde{a}) \right] \sin(2\pi u) \exp(-u^2\tilde{a}^2) du \right|^2 \quad (10)$$

Результати розрахунку за формулою (10) подано на рис.3. З малюнка видно, що при $\tilde{a}=3,338$ і $\tilde{b}=1,980$ ефективність збудження коливання TEM_{30q} в резонаторі дорівнює нулю. Вище було показано, що при таких розмірах хвильоводу, що збуджує, ефективність збудження коливання TEM_{10q} в ВРСП за допомогою хвилі TE_{20} має максимальне значення. Таким чином, проведені чисельні дослідження показують, що при правильному виборі поперечних розмірів прямокутного хвильоводу з хвилею TE_{20} , розташованого в центрі плоского дзеркала ВРСП, останній матиме одночастотний відгук у широкій смузі частот. Якщо посередині перпендикулярно широким стінкам у хвильоводі, що збуджує, розташувати металеву площину, то ми отримаємо два прямокутні хвильоводи перетином $a \times b$ (див. рис.1). У кожному з таких хвильоводів буде поширюватися основна хвиля TE_{10} . При цьому максимальна ефективність збудження TEM_{10q} в ВРСП за допомогою такої хвилі складе 0,4335 при розмірах прямокутного розкриття $1,669 \times 1,98$.

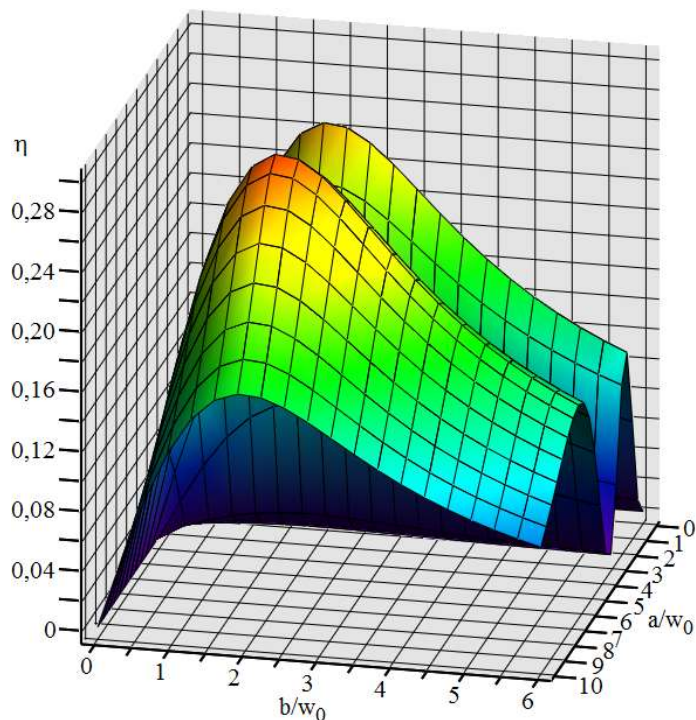


Рис.3. Ефективність збудження коливання TEM_{30q} в ВРСП за допомогою хвильоводної хвилі TE_{20}

Ефективність збудження коливання, що розглядається, за допомогою двох прямокутних хвильоводів зазначених розмірів збігається з величиною, отриманою вище. При розміщенні джерел у кожному з двох прямокутних хвильоводів, що переходять у випромінювачі наведених вище розмірів, ми отримаємо суматор потужності на основі напівсферичного ВРСП. Запропонована система дозволяє регулювати величину зв'язку джерел із резонатором. Для цього на плоскому дзеркалі ВРСП розташовується E - поляризована дротяна дифракційна решітка, змінюючи період якої ми можемо регулювати зв'язок резонатора з джерелами.

Розглядатимемо хвильовід з хвилею TE_{20} , у розкритті якого розташована дротяна E - поляризована дифракційна решітка. Для визначення коефіцієнтів відбиття від ВРСП скористаємося відомим методом підсумовування парціальних коефіцієнтів відбиття від резонансної системи [16,17]. Запишемо вираз, що визначає резонансний коефіцієнт відбиття від ВРСП [18]:

$$\Gamma = |r_1| - \left(|t_1|^2 |r_2| S_e S_p^2 / (1 - |r_1| |r_2| S_p^2) \right), \quad (11)$$

де $|r_1|$ і $|t_1|$ – модулі коефіцієнтів відбивання та передачі дротяних E – поляризованих дифракційних решіток різного періоду у розкритті прямокутного хвильоводу в центрі плоского

дзеркала ВРСП, які пов'язані співвідношенням $|r_1|^2 + |t_1|^2 = 1$ [19]; r_2 – модуль коефіцієнта відбиття від другого дзеркала резонатора, який вважаємо рівним одиниці; S_e і S_p – коефіцієнти передачі по полю за прохід хвилі від одного дзеркала до іншого, які визначаються втратами на збудження $\alpha_e = P_e/P_{рез} = 1 - \eta$, $S_e = \exp(-\alpha_e/2)$ та резонансними втратами $\alpha_p = \alpha_o + \alpha_{ом} = (P_o + P_{ом})/P_{рез}$, $S_p = \exp(-\alpha_p/2)$. Тут P_e – втрати потужності, які складаються, загалом, із втрат на 2 перелив 2 потужності за кромки одного із дзеркал резонатора та втрат потужності на неузгоджене збудження коливання в ВРСП. P_o і $P_{ом}$ – потужності дифракційних та омічних втрат, а $P_{рез}$ – потужність, що надійшла до ВРСП через елемент зв'язки в момент резонансу. При розрахунку ефективності порушення коливання у резонаторі ми поклали, що апертури дзеркал нескінченні. Тому P_e визначатиметься втратами потужності на неузгоджене збудження, а резонансні втрати в ВРСП – потужністю омічних втрат $P_{ом}$.

На рис. 4 наведено залежності коефіцієнта відображення, що визначається виразом (11), від величини резонансних втрат α_p для дріотяних Е-поляризованих дифракційних решіток у розкритті хвилеводу, періоди яких дорівнюють: 0,05 мм (крива 1); 0,10 мм (крива 2); 0,15 мм (крива 3). Всі ґрати намотані дротом діаметром $d = 0,02$ мм. Криві побудовані для довжини хвилі у вільному просторі $\lambda = 2$ мм та значення $S_e = 0,936$ ($\eta = 0,867$), що відповідає максимальній ефективності збудження коливання TEM_{10q} в ВРСП за допомогою хвилі TE_{20} . Модулі коефіцієнтів відображення $|r_1|$ дифракційних ґрат обчислюються за формулою [20],

$$|r_1| = (1 - k_0^2 l_0 l_2) / \sqrt{(1 + k_0^2 l_0^2)(1 + k_0^2 l_2^2)}, \quad (12)$$

де $l_0 = (l/\pi) \ln(l/2\pi d)$, $l_2 = (\pi d^2/l)$, $k_0 = 2\pi/\lambda$.

Як видно з рис.4 загальна поведінка розрахункових залежностей при зменшенні резонансних втрат добре узгоджується з поведінкою виміряного коефіцієнта відбиття від напівсферичного ВР при зменшенні відстані L/R між дзеркалами [21].

У цьому випадку зменшуються розміри "плям" полів на дзеркалах і, як наслідок, стає меншим α_p . При зміні періоду ґрат у розкритті хвилеводу режим узгодження ($\Gamma = 0$) має місце за різних значеннях резонансних втрат: $\alpha_p = 8,72 \times 10^{-5}$ при $l = 0,05$ мм (крива 1); $\alpha_p = 1,39 \times 10^{-3}$ при $l = 0,10$ мм (крива 2); при $\alpha_p = 8,15 \times 10^{-3}$ при $l = 0,15$ мм (крива 3).

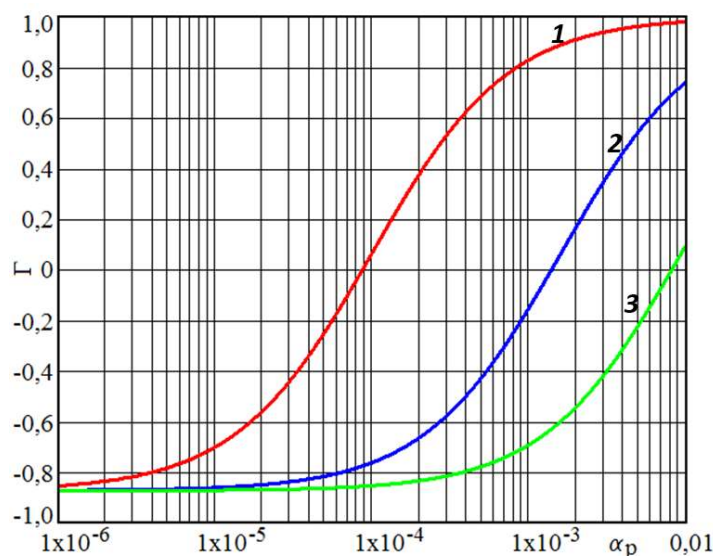


Рис.4. Залежність коефіцієнта відображення від величини резонансних втрат ВРСП для різних періодів решіток: 0,05 мм (крива 1), 0,10 мм (крива 2) , 0,15 мм (крива 3)

Представляє також інтерес розглянути залежність навантаженої добротності резонатора від коефіцієнта передавання решітки. Для цього запишемо вираз, що визначає Q_H [18]:

$$Q_H = \frac{2kL\sqrt{|r_1||r_2|}S_r}{1-|r_1||r_2|S_r^2}. \quad (13)$$

Ці залежності в припущенні, що $|r_2|=1$ наведені на рис. 5 для трьох значень резонансних втрат, відповідних узгодженому збудженню резонатора: $\alpha_r = 8.72 \cdot 10^{-5}$ (крива 1), $\alpha_r = 1.39 \cdot 10^{-5}$ (крива 2) і $\alpha_r = 8.15 \cdot 10^{-3}$ (крива 3). При $|t_1|=0$ добротність зв'язку резонатора з хвильоводом дорівнює нескінченності. Відповідно, Q_H визначатиметься власною добротністю резонатора Q_0 , яка залежить тільки від резонансних втрат у системі. Як видно з рисунка, зі зростанням резонансних втрат і при $|t_1|=0$ значення Q_H/kL зменшиться: 22940 (крива 1), 1439 (крива 2), 245.4 (крива 3). По мірі росту $|t_1|$ відмінність у значеннях навантажених добротностей ВР за різних резонансних втрат проявлятиметься слабкіше. Це пов'язано з тим, що в разі слабого зв'язку ($|t_1| \leq 0.3$), що відповідає застосуванню густих решіток, навантажена добротність ВР, як було сказано вище, визначається добротністю Q_0 , яка, в свою чергу, залежить тільки від величини резонансних втрат. Зі збільшенням коефіцієнта передачі решітки ($|t_1| \geq 0.5$) навантажена добротність резонатора визначатиметься добротністю зв'язку резонатора Q_{con} з хвильоводним трактом, і зміна резонансних втрат практично не позначатиметься на значеннях Q_H .

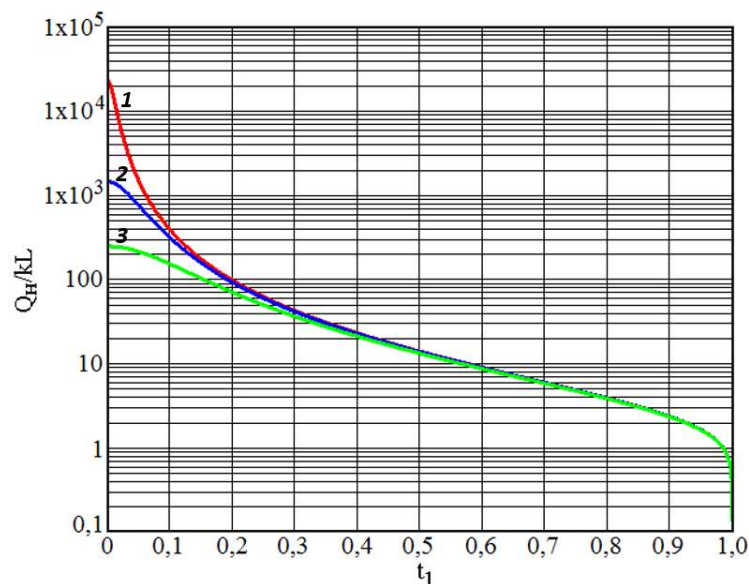


Рис.5. Поведінка навантаженої добротності ВР при зміні коефіцієнта передачі решітки: $\alpha_r = 8.72 \cdot 10^{-5}$ (крива 1), $\alpha_r = 1.39 \cdot 10^{-5}$ (крива 2) і $\alpha_r = 8.15 \cdot 10^{-3}$ (крива 3)

З наведеного малюнка також впливає, що для отримання високих значень навантаженої добротності під час узгодження джерел коливань із резонатором ($\Gamma=0$), необхідно застосовувати густі решітки.

Таким чином, проведені дослідження дозволяють зробити низку важливих практичних висновків.

1. При правильному виборі геометричних розмірів елементів зв'язку окремих джерел з ВРСП можна забезпечити високий коефіцієнт складання потужностей на вкрай високих частотах в і ТГц діапазоні.

2. При узгодженому збудженні коливаль в ВРСП, останній має кутову та частотну селекцію спектра.

3. Вибираючи параметри одновимірних дифракційних решіток (Е-поляризація) у розкривах елементів зв'язку, можна забезпечити будь-який наперед заданий зв'язок окремих джерел з резонансним об'ємом.

Як приклад реалізації високостабільного генератора з використанням квазіоптичних та просторових методів підсумовування потужності розглянемо вихідну частину гетеродинного ланцюга. Вхідну частину гетеродинного ланцюга прийомо-передавача, що складається з високостабільного генератора, що задає частоту 8,1 ГГц і ряду підсилювально-помножувальних ланцюжків, виконаних на монолітних мікросхемах, докладно не розглядаємо. Частота сигналу, яка подається на вхід схеми від попередніх каскадів гетеродинного ланцюга, збільшується до 32,4 ГГц за допомогою двох послідовно включених подвійників частоти, виконаних на мікросхемах CHX 2089 фірми UMS та XX 1000 BD фірми Mimix Broadband (рис. 6).

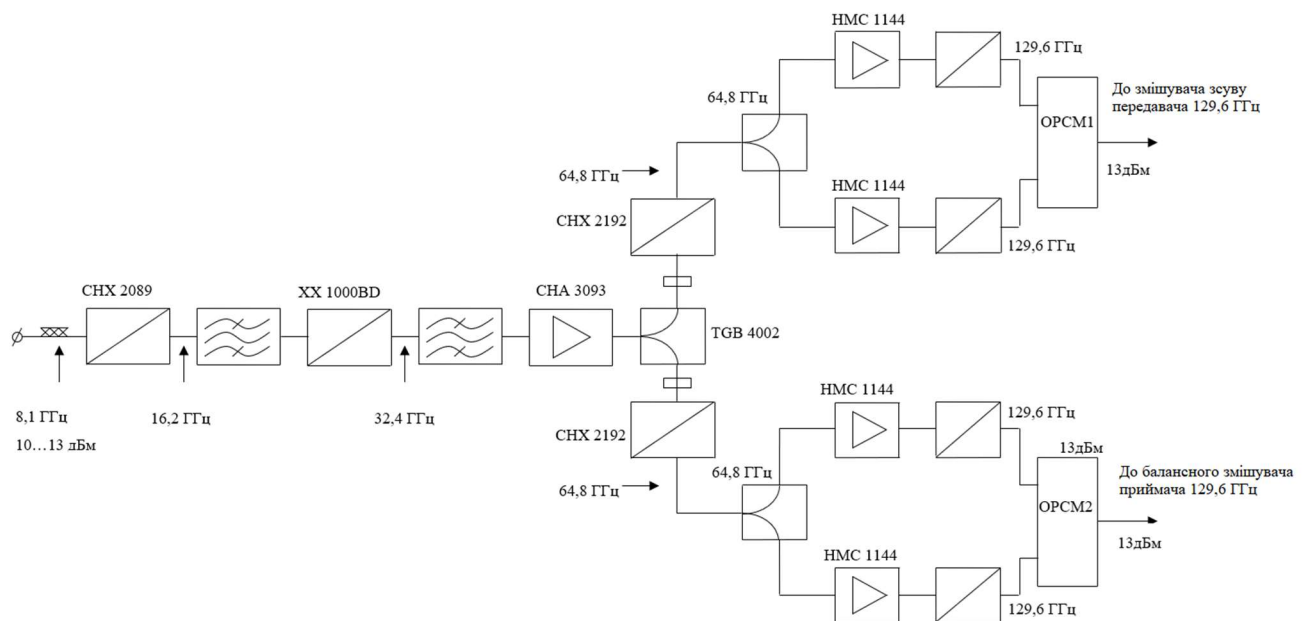


Рис.6. Функціональна схема високостабільного генератора на основі ВР

Ці мікросхеми містять вбудовані підсилювачі, які компенсують втрати сигналу при множенні. Смуго-пропускаючі фільтри на виході помножувальних каскадів запобігають попаданню паразитних гармонік на входи наступних каскадів. Для отримання достатнього рівня вихідних потужностей перед відгалужувачем включений підсилювач потужності, реалізований на мікросхемі CHA 2093 фірми UMS. Посилений сигнал частотою 32,4 ГГц розгалужується на два канали. Вихідні помножувальні каскади виконані на мікросхемах CHX 2192 фірми UMS. Відгалуження здійснюється на мікросхемі TGB 4002 фірми TriQuint. Монолітні мікросхеми помножувачів підсилювача потужності та відгалужувача сигналів включені в мікросмужкові лінії на підкладці з дюріду товщиною 0,127 мм. Сигнали частотою 64,8 ГГц і потужністю 13 дБм надходять на 3 дБ відгалужувачі та діляться за потужністю навпіл. Для отримання достатнього рівня потужності на вході помножувача частоти сигнали надходять на підсилювачі потужності, які виконані на мікросхемах HMC 1144 фірми Analog Devices і посилюються до рівнів 80-100 мВт. Подальше підвищення частоти гетеродину з 64,8 ГГц до 129,6 ГГц здійснюється подвоювачами частоти. Помножувачі частоти побудовані на основі безкорпусних планарних помножувальних діодів Шоткі. Два сигнали частотою 129,6 ГГц і потужністю 10...13 дБм складаються в ВРСП, який має селективні властивості і пригнічує гармоніки помножувача частоти та паразитні частоти на 50 дБ і більше. Виміряні рівні потужності на виході резонаторів перебувають у межах 10...13 дБм.

У процесі роботи для схемотехнічного моделювання (змішувачів, помножувачів, підсилювачів, дільників) використовувався програмний комплекс NI AWR Design Environment; фільтрів та хвильоводних структур – ANSYS HFSS, резонаторів – Altair FEKO.

Запропонована схема складання потужностей має переваги перед іншими схемами на базі ВРСП, оскільки, крім винесення джерел коливань з резонансного обсягу, дозволяє розділити функції узгодження по полю та зв'язку. Вибираючи геометричні розміри кожного з розкривів, ми узгоджуємо електричні компоненти хвиль TE_{10} у хвильоводах із робочим коливанням TEM_{10q} резонатора. Оскільки поперечні розміри кожного з двох випромінювачів у площині дзеркала ВРСП перевищують робочу довжину хвилі, зв'язок активних елементів з резонатором має бути вищим за критичний (коефіцієнт відображення від ВРСП $\Gamma < 0$). Змінюючи період розташовуваної на плоскому дзеркалі резонатора дифракційної решітки, дротяні ламелі якої паралельні вектору напруженості електричного поля збудливої хвилі TE_{10} (випадок Е-поляризації), можна забезпечити будь-який заданий зв'язок джерел з робочим коливанням. Експериментальне дослідження відкритого квазіоптичного резонатора суматора потужностей було проведено на частоті 130 ГГц за допомогою панорамного вимірювача Elmika R4402R (діапазон частот 110...170 ГГц). Умовою узгодження ВРСП на частоті 130 ГГц є мінімальний коефіцієнт стоячої хвилі, що становить величину 1,2. Апертури дзеркал резонатора дорівнюють 40 мм, радіус кривизни сферичного відбивача $R = 39$ мм (рис. 1). Розміри розкривів у площині одного із дзеркал ОР вибиралися з умови порушення коливання TEM_{10q} з максимальною ефективністю. З виконаного вище аналізу вони склали: $a=5,9$ мм, $b=7,0$ мм. У резонаторі збуджується тип коливань TEM_{1019} ($L/R = 0,573$). Період l дротяної Е-поляризованої дифракційної решітки, розташованої на плоскому дзеркалі резонатора, ми вибрали рівним 0,10 мм, що добре узгоджується з проведеними вище розрахунками. Виміряна потужність на виході ВРСП знаходилася в межах (10...13 дБм), що відповідає ефективності підсумовування потужностей $k = 0,76$.

Висновки

1. Велике значення для складання потужностей має отриманий у роботі результат, що свідчить про те, що при узгодженому збудженні аксіально-несиметричних типів коливань резонансна система матиме кутову та частотну селекцію спектру.

2. Вибираючи відповідним чином параметри (період і коефіцієнт заповнення) одновимірної дифракційної решітки (випадок Е-поляризації), що розташовується в розкриві елемента зв'язку, можна забезпечити будь-який заданий зв'язок джерела з резонансним обсягом.

3. Досліджений в роботі суматор потужності на базі квазіоптичного відкритого резонатора на частоту 130 ГГц, дозволив створити зразок високостабільного генератора.

Дані, отримані в результаті дослідження, дозволяють зробити висновок про перспективність використання розробленого ВРСП для побудови високостабільних генераторів субтерагерцевого та терагерцевого діапазонів.

Ця наукова стаття була підтримана Національним фондом досліджень України в рамках виконання грантового проєкту «Мікрохвильові пристрої на основі резонансних структур з метаматеріальними властивостями для захисту життєдіяльності та інформаційної безпеки України» (Реєстраційний номер проєкту: 2021.01/0030).

Література

1. Наритник Т. М., Кравчук С. О. Телекомунікаційні системи терагерцевого діапазону: монографія. Житомир: ФОП «Євенок О.О.», 2015. 394 с.
2. Наритник Т.М., Сайко В.Г. Безпроводні системи зв'язку терагерцевого діапазону. Німеччина: LAP LAMBERT Academic Publishing RU., 2019. 72 с.
3. Popovic Z., Grossman E. N, THz metrology and instrumentation. *IEEE Trans. on Terahertz Science and Technology*, January 2011. vol.1, pp.133-144.
4. Svetlitz A., Slavenko M., Blank T., Brouk I., and et al., "THz Measurements and Calibration Based on a Blackbody Source. *IEEE Trans. on Terahertz Science and Technology*, March 2014. vol.4, pp. 347-359.
5. Карушкин Н.Ф., Касаткин Л.В., Мальцев С.Б. Опыт разработки мощных импульсных источников мощности миллиметрового диапазона волн на ЛПД. *КрыМиКо-2016: труды 16-ой Международной Крымской конференция "СВЧ-техника и телекоммуникационные технологии"*, г.Севастополь, 11-15 сент. 2006 р, Севастополь, 2006. Т.1., С.135-137.
6. Mizuno K, Ajikata T., Hieda M., Nakayama M. Quasi-Optical Resonator for Millimetre and Submillimetre Wave Solid-State Sources. *Electronics Letters*, June 1988, vol.24, №.13, pp.791–793.

7. Judaschke R., Hoft M., Schunemann K., Quasi-Optical 150-GHz Power Combining Oscillator. *IEEE Microwave and Wireless Components Letters*, May 2005. vol.15, No.5, pp.300–302.
8. Kuz'michev I. K., Matching of the quasi-optical open resonators with the waveguide transmission line. *Izv. Vuzov. Radiofizika*, April 2000. vol. 43, pp. 325-334.
9. Машковцев Б.М., Цибизов К.Н., Емелин Б.Ф. Теория волноводов: монография. Москва: Наука, 1966. 351с.
10. Почерняев В.Н., Цибизов К.Н. Теория сложных волноводов: монография Киев: Науковий Світ, 2003. 224 с.
11. Фельдштейн А.Л., Явич Л.Р., Смирнов В.П. Справочник по элементам волноводной техники. Москва: Изд-во Советское радио, 1967. 652 с.
12. Kogelnik, H., 1964. Coupling and conversion coefficients for optical modes. In: Quasi-Optics. *Symposium on Quasi-Optics Proceedings*. Brooklyn, NY: Polytechnic Press. pp. 333–347.
13. Бурштейн Э.Л. О мощности, принимаемой антенной при падении на нее неплоской волны. *Радиотехника и электроника*. Москва, 1958. Т.3. №2. С.186–189.
14. Kay A. Near-Field Gain of Aperture Antennas. *IRE Trans. Antennas Propag.*, November 1960. vol 8. № 6. pp. 586-593.
15. Kuzmichev I.K., Excitation efficiency of quasioptical resonance systems. *Telecommun. Radio Eng.* 2009. vol. 68. №.1. pp. 49-63.
16. Вольман В.И., Пименов Ю.В. Техническая электродинамика / Под ред. Б. З. Айзенберга. Москва: Связь, 1971. 488 с.
17. Звелто О. Принципы лазеров / Пер. с англ. под ред. Т. А. Шмаонова. Москва: Мир, 1990. 558с.
18. Кузьмичев И.К. Апертурное возбуждение открытых резонаторов миллиметрового диапазона: Автореф. дис. канд. физ.-мат. наук: 01.04.03. Ростов-на-Дону: РГУ, 1991. 16 с.
19. Шестопалов В. П., Кириленко А. А., Масалов С. А., Сиренко Ю. К. Резонансное рассеяние волн. Том 1. Дифракционные решетки. Киев: Наукова думка, 1986. 232 с.
20. Вайнштейн Л.А. К электродинамической теории решеток. Часть 1. *Электроника больших мощностей*. Москва, 1963. Вып. 2. С. 26–56.
21. Андросов В.П., Кузьмичев И. К. Влияние на эффективность возбуждения открытого резонатора его параметров и связи с волноводом. Харьков, 1987. 30 с. (*Препринт / Ин-т радиофизики и электроники АН УССР; № 354*).

References

1. Naritnik T. M., Kravchuk S. O Telekomunikatsiini sistemi teragertsovogo diapazonu: monografiya. Zhitomir: FOP «Evenok O.O.», 2015. 394 s.
2. Naritnik T.M., Saiko V.G. Bezprovidni sistemi zv'yazku teragertsovogo diapazonu. Nimechchina: LAP LAMBERT Academic Publishing RU., 2019. 72 s.
3. Popovic Z., Grossman E. N, THz metrology and instrumentation. *IEEE Trans. on Terahertz Science and Technology*, January 2011. vol.1, pp.133-144.
4. Svetlitz A., Slavenko M., Blank T., Brouk I., and et al., "THz Measurements and Calibration Based on a Blackbody Source. *IEEE Trans. on Terahertz Science and Technology*, March 2014. vol.4, pp. 347-359.
5. Karushkin N.F., Kasatkin L.V., Maltsev S.B. Opit razrabotki moshchnikh impulsnikh istochnikov moshchnosti millimetrovogo diapazona voln na LPD. KriMiKo-2016: trudi 16-oi Mezhdunarodnoi Krimskoi konferentsiya "SVCh-tekhnika i telekommunikatsionnie tekhnologii", g.Sevastopol, 11-15 sent. 2006 r, Sevastopol, 2006. T.1., S.135-137.
6. Mizuno K, Ajikata T., Hieda M., Nakayama M. Quasi-Optical Resonator for Millimetre and Submillimetre Wave Solid-State Sources. *Electronics Letters*, June 1988, vol.24, №.13, pp.791–793.
7. Judaschke R., Hoft M., Schunemann K., Quasi-Optical 150-GHz Power Combining Oscillator. *IEEE Microwave and Wireless Components Letters*, May 2005. vol.15, No.5, pp.300–302.
8. Kuz'michev I. K., Matching of the quasi-optical open resonators with the waveguide transmission line. *Izv. Vuzov. Radiofizika*, April 2000. vol. 43, pp. 325-334.
9. Mashkovtsev B.M., Tsibizov K.N., Yemelin B.F. Teoriya volnovodov: monografiya. Moskva: Nauka, 1966. 351s.
10. Pochernyaev V.N., Tsibizov K.N. Teoriya slozhnikh volnovodov: monografiya Kiev: Naukovii Svit, 2003. 224 s.
11. Feldshtein A.L., Yavich L.R., Smirnov V.P. Spravochnik po elementam volnovodnoi tekhniki. Moskva: Izd-vo Sovetskoe radio, 1967. 652 s.
12. Kogelnik, H., 1964. Coupling and conversion coefficients for optical modes. In: Quasi-Optics. *Symposium on Quasi-Optics Proceedings*. Brooklyn, NY: Polytechnic Press. pp. 333–347.
13. Burshtein E.L. O moshchnosti, prinimaemoi antennoi pri padenii na nee neploskoi volni. *Radiotekhnika i elektronika*. Moskva, 1958. T.3. №2. S.186–189.
14. Kay A. Near-Field Gain of Aperture Antennas. *IRE Trans. Antennas Propag.*, November 1960. vol 8. № 6. pp. 586-593.
15. Kuzmichev I.K., Excitation efficiency of quasioptical resonance systems. *Telecommun. Radio Eng.* 2009. vol. 68. №.1. pp. 49-63.
16. Volman V.I., Pimenov Yu.V. Tekhnicheskaya elektrodinamika / Pod red. B. Z. Aizenberga. Moskva: Svyaz, 1971. 488 s.
17. Zvelto O. Printsipi lazerov / Per. s angl. pod red. T. A. Shmaonova. Moskva: Mir, 1990. 558s.
18. Kuzmichev I.K. Aperturnoe vzbuzhdenie otkritikh rezonatorov millimetrovogo diapazona: Avtoref. dis. kand. fiz.-mat. nauk: 01.04.03. Rostov-na-Donu: RGU, 1991. 16 s.
19. Shestopalov V. P., Kirilenko A. A., Masalov S. A., Sirenko Yu. K. Rezonansnoe rasseyanie voln. Tom 1. Difraktsionnie reshetki. Kiev: Naukova dumka, 1986. 232 s.

20. Vainshtein L.A. K elektrodinamicheskoi teorii reshetok. Chast 1. Elektronika bolshikh moshchnostei. Moskva, 1963. Vip. 2. S. 26-5.
21. Androsov V.P., Kuzmichev I. K. Vliyanie na effektivnost возбуждениya otkritogo rezonatora yego parametrov i svyazi s volnovodom. Kharkov, 1987. 30 s. (Preprint / In-t radiofiziki i elektroniki AN USSR; № 354).

УДК 621.396.67

МЕТОД СИНТЕЗУ ВІБРАТОРНИХ АНТЕН З РЕАКТИВНИМИ ВСТАВКАМИ

DOI 10.36994 / 2788-5518-2023-01-05-07

Сайко В.Г. д.т.н., проф. Київський національний університет імені Тараса Шевченка, Київ, Україна. vgsaiko@gmail.com

Ільїнов М.Д. к.т.н., доц. Військовий інститут телекомунікацій та інформатизації імені Героїв Крут, Київ, Україна. ilinov1948@ukr.net

Нестеренко І.К. к.т.н., с.н.с. Військовий інститут телекомунікацій та інформатизації імені Героїв Крут, Київ, Україна. igor.nester@ukr.net

Анотація: Запропонований новий підхід для дослідження електродинамічних властивостей виділеного класу періодично навантажених структур (вібраторних антен з ємнісними вставками), який відрізняється відносною простотою і достатньою точністю і дозволяє отримати рішення в аналітичному вигляді, що дає можливість уникнути труднощів, що виникають при чисельному розв'язанні задачі. В той же час отримані таким чином рішення можуть бути використані в ітераційних методах чисельного аналізу електромагнітних властивостей періодично навантажених структур та антенних пристроїв, що містять ці структури. Наводиться розв'язок задачі поширення електромагнітної енергії в однорідній двохzew'язної лінії передачі з реактивними вклученнями. Її рішення дозволило встановити нові закономірності, а також по-новому підійти до розрахунку вібраторних антен з ємнісними вставками. Проведені дослідження діапазонних властивостей антен з ємнісними вставками дозволили встановити три крайніх випадку розподілу в них струму: експоненціальне, синусоїдальне і лінійне, отримати вирази для розрахунку внутрішніх і зовнішніх характеристик таких антен, виявити залежності їх електричних характеристик від геометричних розмірів і сформулювати рекомендації щодо раціонального вибору величин ємнісних ставок і місць їх вклучення. Розроблено метод розрахунку вібраторних антен з ємнісними вставками, сутність якої полягає в раціональному використанні дисперсійних властивостей таких структур в частотній області для створення широкосмугових антен с вставками. Для забезпечення роботи радіо засобів спеціального призначення виготовлені та досліджені зразки вібраторних антен з двома, трьома ємнісними вставками. Результати випробувань таких антен показали, що їх коефіцієнт посилення на частотах, близьких до критичної, в два та більше разів перевищує коефіцієнт посилення четверть хвильового випромінювача і, отже, коефіцієнт корисної дії розроблених антен також зріс.

Ключові слова: вібраторна антена з ємнісними вставками, дисперсія, однорідна лінія передачі, метод розрахунку вібраторних антен з ємнісними вставками.

SYNTHESIS METHOD OF VIBRATOR ANTENNAS WITH REACTIVE INSERTS

Volodymyr Saiko, Dr.habil., Prof. Taras Shevchenko National University of Kyiv, Kyiv, Ukraine. vgsaiko@gmail.com

Mykhailo Ilyinov, Ph.D., associate professor Military Institute of Telecommunications and Informatization named after Heroes of Krut, Kyiv, Ukraine. ilinov1948@ukr.net.....

Igor Nesterenko, Ph.D., SSR.. Military Institute of Telecommunications and Informatization named after Heroes of Krut, Kyiv, Ukraine. igor.nester@ukr.net.....

Abstract: A new approach is proposed for investigating the electromagnetic properties of a specific class of periodically loaded structures (vibrator antennas with capacitive inserts), which is characterized by relative simplicity and sufficient accuracy. It allows obtaining analytical solutions, thus avoiding the difficulties associated with numerical solutions. At the same time, the obtained solutions can be used in iterative methods for numerical analysis of electromagnetic properties of periodically loaded structures and antenna devices containing these structures. The solution of the problem of electromagnetic energy propagation in a uniform two-conductor transmission line with reactive inclusions is presented. Its solution has revealed new regularities and provided a fresh perspective on the calculation of vibrator antennas with capacitive inserts. Studies on the frequency characteristics of antennas with capacitive inserts have identified three extreme cases of current distribution: exponential, sinusoidal, and linear. Expressions for calculating internal and external characteristics of such antennas have been derived, and dependencies of their electrical characteristics on geometric dimensions have been established. Recommendations for the rational selection of capacitance values and their locations have been formulated. A calculation method for vibrator antennas with capacitive inserts has been developed, which involves the rational use of dispersion properties of such structures in the frequency domain to create wideband

antennas with inserts. For the operation of specialized radio equipment, samples of vibrator antennas with two or three capacitive inserts have been manufactured and studied. The results of testing such antennas have shown that their gain at frequencies close to the critical value exceeds the gain of a quarter-wavelength radiator by two or more times, thus increasing the overall efficiency of the developed antennas.

Keywords: vibrator antenna with capacitive inserts, dispersion, uniform transmission line, calculation method for vibrator antennas with capacitive inserts.

Вступ

Аналіз математичних підходів до розрахунку широкосмугових вібраторних антен з ємнісними вклученнями (далі ВАВ) показує, що строгі методи розрахунку, через важкість рішення оптимізаційної задачі за вхідним опором і діаграмою спрямованості, не дозволяють окреслити границі пошуку варіюваних параметрів антен, від яких залежить ефективність процесу оптимізації [1-3]. Крім того, чисельні рішення задач синтезу імпедансних структур, як правило, призводять до необхідності досліджувати погано зумовлені системи лінійних рівнянь алгебраїчних високого порядку, вирішення яких, як і раніше, становить одну зі складних проблем чисельних методів в електродинаміці [4-6]. Методи, засновані на аналогії з довгою лінією, мають низьку точність отриманих результатів, що позбавляє сенсу чисельної оптимізації. Через зазначені причини такі підходи не дозволяють наочно виявляти характер впливу всіх параметрів ВАВ: величин і координат вклучення вставок, довжини і радіусу провідника антени. Відсутність розуміння фізичного явища, що призводить до зміни розподілу струму уздовж лінійного вібратора, призводить до того, що в більшості відомих публікацій даються в основному експериментальні дані або ж визначаються закони, що забезпечують синфазний розподіл струму уздовж антени в необхідному діапазоні частот. З цієї причини представляє інтерес розрахунок параметрів вібраторної антени з ємнісними вставками на основі фізичного осмислення задачі [7-10].

Виконання досліджень

В якості вихідного приймається допущення, що ємнісні вставки вклучені безперервно по всій довжині антени. Таке припущення вважається справедливим при досить великій кількості ємнісних вставок на ділянці лінії, порівнянню з довжиною хвилі. Для уточнення меж застосовності допущення, досліджуємо властивості зазначених електродинамічних систем [7].

Уявімо з'єднання відрізків лінії передачі з вклученими в них реактивними у вигляді каскадного з'єднання 4-полюсниками (рис. 1).

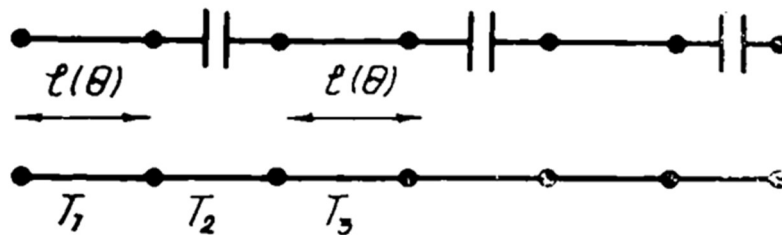


Рис.1 З'єднання відрізків лінії передачі з вклученими в них реактивними вставками

Перший 4-полюсник – відрізок однорідної двохзв'язної лінії довжиною l , характеризується матрицею передачі:

$$T_1 = \begin{Bmatrix} e^{j\theta} & 0 \\ 0 & e^{-j\theta} \end{Bmatrix}, \quad j\theta = j\omega\sqrt{[L_0 C_0] * l}.$$

Наступна за ним послідовно вклучена ємність характеризується матрицею:

$$T_2 = \begin{Bmatrix} 1 + z^\wedge/2 & -z^\wedge/2 \\ z^\wedge/2 & 1 - z^\wedge/2 \end{Bmatrix}.$$

де $z^\wedge = z/p = 1/j\omega C_{b_0}\sqrt{[L_0/C_0]}$; C_b – ємність включення.

Потім включений відрізок однорідної лінії, який характеризується матрицею $T_3 = T_1$, далі включено кілька аналогічних з'єднань 4-полісників.

Матриця передачі T двох каскадно-включених 4-полісників, кожен з яких характеризується матрицею $T_4 = T_1T_2T_3$:

$$T = T_4T_4 =$$

$$= \begin{vmatrix} (1 + z^\wedge/2)e^{j4\theta} - z^\wedge/4 & -z^\wedge/2[(1 + z^\wedge/2)e^{j2\theta} + (1 - z^\wedge/2)e^{-j2\theta}] \\ z^\wedge/2[(1 + z^\wedge/2)e^{j2\theta} + (1 - z^\wedge/2)e^{-j2\theta}] & (1 + z^\wedge/2)^2 e^{-j4\theta} = z^\wedge/4 \end{vmatrix}$$

Умови рівності одиниці коефіцієнта передачі:

$$t_{2ej\theta} = |z^\wedge/2[(1 + z^\wedge/2)e^{j2\theta} + (1 - z^\wedge/2)e^{-j2\theta}]|^2 = 0.$$

Практичний інтерес для дослідження представляє рівність нулю другої співмножники. Після його перетворення отримаємо

$$2\theta = \arctg 2\omega C_b \sqrt{[L_0 C_0]} \quad (1)$$

Залежність (1) показує: якщо ємнісні вставки встановлені на відстанях, рівних або більших чверті довжини робочої хвилі ($l \geq \lambda/4$), коефіцієнт передачі t_{21} носитиме резонансний характер. Такі структури з нулями на частотах, що визначаються виведеною рівністю, добре відомі.

Якщо прийняти довжину відрізків лінії, що з'єднують реактивності, на порядок меншими чверті довжини хвилі, то значення тангенса можна замінити значенням його аргументу. Отримаємо:

$$C_0 l = C_b \quad (2)$$

З виразу (2) видно, що в разі, якщо ємнісні реактивності встановлені на відстанях, менших, ніж довжина робочої хвилі, і кожна ємність по своїй величині дорівнює сумарній розподіленій ємності відрізка сполучної лінії l , коефіцієнт передачі такої структури незалежно від робочої частоти дорівнює одиниці.

Залежності (1), (2) дозволяють зробити висновок: якщо включити рівномірно по всій довжині ємності дискретно або ємності розподіленого характеру на відстанях, рівних або менше однієї восьмої довжини хвилі, то така структура буде представляти однорідну лінію передачі. Розглянемо для неї рішення задачі поширення електромагнітної енергії.

З точки зору теорії довгих ліній кожна ділянка такої однорідної лінії з ємнісними включеннями володіє індуктивністю L_0 , ємністю C_0 , активним опором R_0 , провідністю G_0 і ємністю C_ϵ , включену в лінію послідовно. Еквівалентна схема такої лінії зображена на рис. 2.

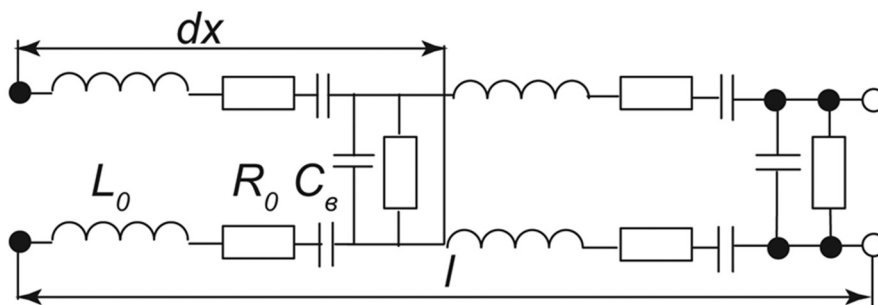


Рис. 2 Еквівалентна схема з'єднання відрізків лінії передачі з включеними в них реактивними вставками

Шляхом вирішення телеграфних рівнянь можна показати, що коефіцієнт поширення і коефіцієнт фази будуть визначатися виразами:

$$\gamma = j\beta = j\omega\sqrt{L_0C_0}\sqrt{[1 - (\omega_{кр}/\omega)^2]}; \quad \beta = \beta_0\sqrt{[1 - (\omega_{кр}/\omega)^2]},$$

де $\omega_{кр} = 1/\sqrt{L_0C_0}$; β_0 – коефіцієнт фази двухпровідної лінії без втрат і без ємнісних вставок.

З отриманих виразів видно, що в даній структурі двухзв'язної лінії з послідовно включеними ємнісними вставками фазова швидкість залежить від частоти, тобто в ній проявляється явище дисперсії, подібне дисперсії в хвилеводі. Доречно зауважити, що відомі радіочастотні лінії передачі електромагнітної енергії, що містять два струмонесучі провідники, розташованих паралельно один одному, хоча б в один з яких послідовно включені однакові ємності (1), мають властивості смуго-пропускаючих фільтрів, тобто не володіють дисперсійними властивостями.

Діапазон робочих частот, в якому будуть активно проявлятися дисперсійні властивості, складе до $4\omega_{min}$ т.е. $\omega_{min}/\omega_{max} = C_0/C_B = 1 \dots 0.25$.

Наведені результати дозволили по-новому пояснити багато раніше відомих результатів (2, 3) і підійти до розрахунку вібраторних антен з нових позицій.

Прийmemo, що несиметричний вібратор отриманий з відрізка розглянутої однорідної двухзв'язної лінії з ємнісними включеннями і складається з окремих відрізків циліндричного провідника $\Delta l \ll \lambda/4$ суміжні торці яких утворюють ємнісні вставки.

Традиційно розглядатимемо несиметричний вібратор з циліндричних проводів загальною довжиною l і діаметром $2a$. Радіус проводу настільки малий у порівнянні з його довжиною, що виконується умова $2 \ln 2l/a \gg 1$.

Розподіл струму в несиметричному вібраторі з ємнісними вставками матиме залежність:

$$I(z) = I_m \sin \beta(l - |z|), \quad (3)$$

де β – коефіцієнт фази хвилі струму (напруги) в лінії з ємнісними включеннями.

З формули (3) випливає, що постійна поширення γ в залежності від співвідношення робочої довжини хвилі до критичної довжини хвилі лінії $\lambda/\lambda_{кр}$ може бути дійсною, уявною і рівною нулю.

На частотах, вище критичної, розподіл струму вздовж вібратора підпорядковується синусоїдальному закону. Однак якщо в вібраторі без реактивних включень ця залежність однозначно визначалася тільки співвідношенням робочої довжини хвилі і власної довжини вібратора, то для вібратора з ємнісними вставками вона ще залежить від співвідношення робочої і критичної довжини хвилі. На частотах, нижче критичної, хвилі струму не можуть поширюватись вздовж вібратора. Це пояснюється тим, що амплітуда струму зменшується від початку до кінців вібратора по експоненціальному закону з постійною фазою і відповідно енергія не може поширюватися уздовж вібратора.

Для практики представляє інтерес дослідження розподілу струму з ємнісними вставками на частотах, близьких до критичної. На частотах $\omega_{кр} \lesssim \omega_{раб}$ величина постійної поширення стає малою і формула (3) набуде вигляду $I(z) = I_A(1 - |z|/l)$.

Для області $\omega_{кр} \gtrsim \omega_{раб}$, аналогічно міркуючи, приходимо також до лінійного розподілу струму $I(z) = I_A \operatorname{sh} \beta(1 - |z|)/\operatorname{sh} \beta l = I_A(1 - |z|/l)$.

Для уточнення результату в критичних умовах слід підійти до тих самих висновків, але з іншого боку.

Для цього визначимо діючу довжину несиметричного вібратора з ємнісними вставками в трьох крайніх випадках.

Коли співвідношення $\lambda/\lambda_{кр} < 1$, діюча довжина вібратора буде визначатися

$$l_g = 1/\sin \beta \int_0^l \sin \beta(l - |z|) dz = 1/\beta (\operatorname{tg} \beta l/2)$$

На частотах, близьких до критичної $\lambda/\lambda_{кр} \lesssim 1$, $l_g = l/2$, тобто діюча довжина вібратора буде рівною половині його геометричної довжини. Діюча довжина вібратора при $\lambda/\lambda_{кр} < 1$ буде змінюватися зі зростанням частоти значно повільніше, ніж у звичайного вібратора. Цим пояснюються широкосмугові властивості таких вібраторів.

На частотах, нижче критичної, діюча довжина:

$$l_g = 1/\text{sh } \beta l \int_0^l \text{sh } \beta(l - |z|) dz = 1/\beta \text{sh } \beta l (\text{ch } \beta l - 1).$$

При $\lambda/\lambda_{кр} \gtrsim 1$ $l_g = l/2$.

Таким чином, апроксимація струму вздовж вібратора з ємнісними включеннями синусоїдальним та експоненціальним законами дає в критичних умовах однакові кількісні і якісні результати. Звідси можна зробити висновок, що струм вздовж вібратора з включеннями, працюючого на частоті, близькій до критичної, за фазою не змінюється, а за амплітудою змінюється практично по лінійному закону. Фізично це явище пояснюється тим, що довжина хвилі в вібраторі з включеннями при роботі на частоті, близькій до критичної, прямує до нескінченності, і вібратор буде працювати в режимі великого подовження.

Відповідно можливі три крайні випадки розподілу струму вздовж вібратора:

γ – дійсна, якщо $\lambda/\lambda_{кр} > 1$ – експоненціальний розподіл струму;

γ – уявна, якщо $\lambda/\lambda_{кр} > 1$ – синусоїдальний розподіл струму;

γ – дорівнює нулю, якщо $\lambda/\lambda_{кр} = 0$ – лінійний розподіл струму.

Виходячи з того, що для антен типу, що розглядаються, характерні три крайні випадки розподілу струму, формули для розрахунку зміни діаграми спрямованості $f(\theta)$, опору випромінювання $R_{\Sigma n}$, вхідного опору R_A і X_A і коефіцієнту напруженої дії (КНД) D в цих областях наведені [7-10].

Результати обчислення діаграми спрямованості, опору випромінювання і КНД при синусоїдальному та експоненціальному розподілу струму представлені на рис. 3 а,б та рис. 4 а,б відповідно.

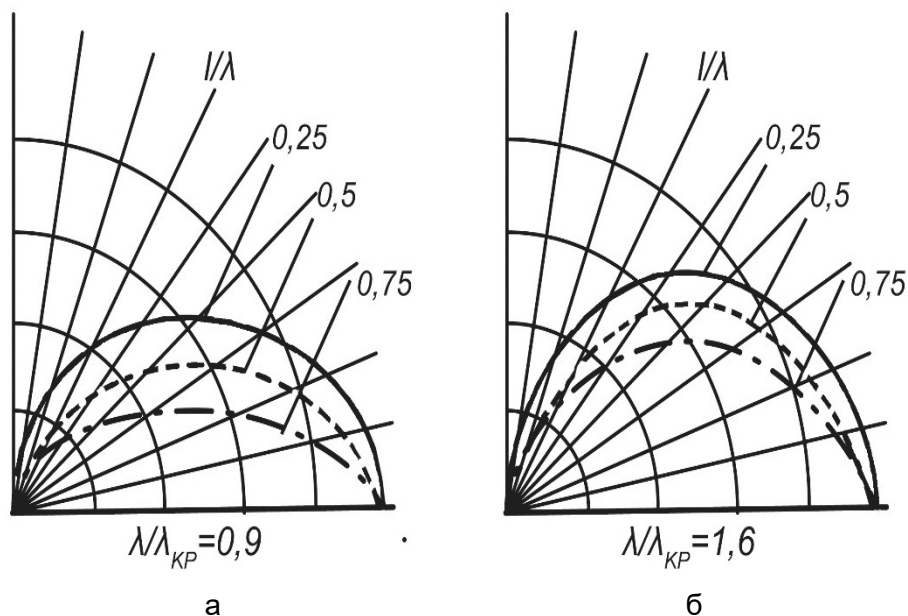


Рис. 3 Діаграми спрямованості при синусоїдальному та експоненціальному розподілу струму

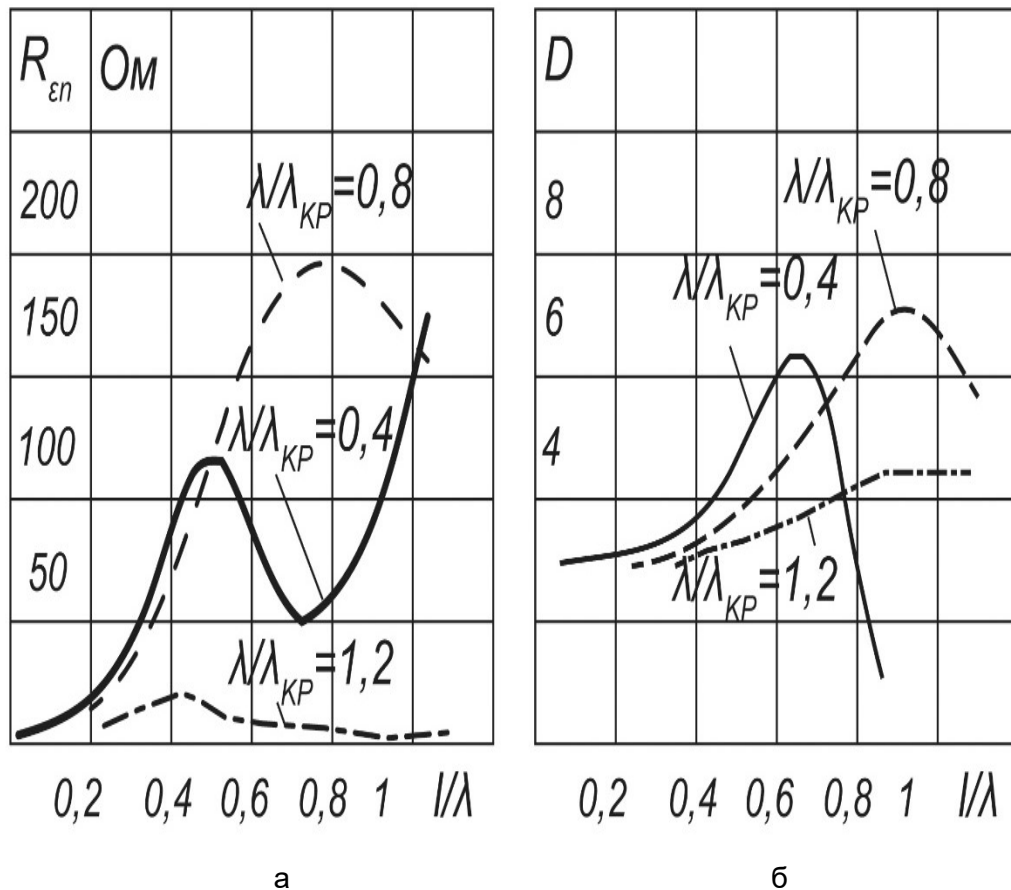


Рис. 4 Залежності опору випромінювання і КНД при синусоїдальному та експоненціальному розподілі струму

З графіків слідує, що направлені властивості вібраторної антени по мірі зростання $\lambda/\lambda_{\text{кр}} (\omega_{\text{раб}} > \omega_{\text{кр}})$ збільшуються. Але при $\lambda/\lambda_{\text{кр}} \gg 1$ опір випромінювання стає малим. Низьке значення опору випромінювання антени ускладнює боротьбу за її високий КПД, так як в порівнянні з опором $R_{\Sigma A}$ стають відчутні будь-які втрати в проводах, ізоляторах, оточуючих антену предметах і на землі. При цьому втрати і випромінювання оточуючих предметів можуть спотворити діаграму спрямованості. Низький КПД антени з ємнісними вставками при експоненціальному розподілі струму являється тенденцією, негативною для проектування таких антен. Однак якщо робоча частота більше, але близька до критичної ($\omega_{\text{раб}} \gtrsim \omega_{\text{кр}}$), то відбувається зростання $R_{\Sigma A}$, КНД, розширюються діапазонні властивості і збільшується ефективність антени. Це являється позитивною тенденцією при проектуванні штирових антен для рухомих об'єктів зв'язку.

Таким чином, можна зробити висновок про наявність деякої оптимальної області значень $\lambda/\lambda_{\text{кр}}$, яку доцільно використовувати при побудові вібраторних антен з ємнісними вставками.

Результати, отримані в попередньому параграфі, дозволяють зробити висновок про можливість створення вібраторної антени, що працює в широкій смузі частот (3-4 октави). Така антена є вібратор з ємнісними вставками, кожне з плечей, якій містить безліч послідовно розташованих секцій і з'єднують їх ємності включення, розташовані між ними.

Конструктивні параметри її повинні мати такі наступні обмеження [8,9]:

1. Вібратор повинен бути системою дисперсійною, тобто в ньому включені ємності вставки $C_{\text{в}}$, рівні або більше розподіленої ємності кожної секції C_0 , тобто $C_{\text{в}} \geq C_0$. Це обмеження забезпечує отримання квазісинусоїдального розподілу струму уздовж вібратора, тобто забезпечення максимально можливої діючої довжини антени.

2. Довжина кожної секції lc повинна бути коротше однієї восьмої довжини робочої хвилі, тобто:

$$lc \leq \lambda_{max}/8.$$

Умова визначає вибір необхідної кількості секцій та ємнісних вставок. Виконання його призводить до зникнення в квазірозподіленій системі резонансних властивостей, тобто така антена буде частотно-незалежною. Іншими словами, така антена буде мати аналогічні властивості однорідної лінії передачі електромагнітної енергії з включеними в неї послідовно зосереджені ємностями.

3. Погонні параметри секції (L_o , C_o) і ємності вставки (C_e) обрані так, що вібратор працює на частоті, близькій або рівній критичній частоті вібратора, яка визначається залежністю:

$$\omega_{кр} = 1/\sqrt{L_o C_B}$$

4. Загальна довжина антени l повинна бути обрано рівною:

$$l = 3/4 * \lambda_{min},$$

де λ_{min} - мінімальна робоча довжина хвилі, $\lambda_{кр}$ - критична довжина хвилі:

$$\Lambda_{min} = \frac{\lambda_{min}}{\sqrt{1 - \left(\frac{\lambda_{min}}{\lambda_{кр}}\right)^2}};$$

Зупинимось на двох останніх обмеженнях, що визначають умови забезпечення збереження однопелюскової діаграми спрямованості вібраторної антени з ємнісними вставками. Прийнято, що довжина плеча антени вибирається з міркувань збереження діаграми в площині вібратора у вигляді одного пелюстка (тобто бажано отримати діаграму спрямованості неізрезану). Для антен з ємнісними включеннями електрична довжина змінюється відповідно до зміни частоти з урахуванням дисперсійних властивостей системи. З цієї точки зору для збереження однопелюскової діаграми спрямованості необхідно виконати наступні дві умови. Перша умова може бути представлено, як умова пропорційної залежності мінімальної частоти від величини індуктивності секції вібратора і ємності включення:

$$\omega_{min} = \omega_{кр} = 1/\sqrt{L_o C_B}$$

Виконання цього співвідношення забезпечить отримання на нижніх частотах коефіцієнта посилення антен не гірше, ніж коефіцієнт посилення на нижніх частотах вібраторної антени без ємнісних включень. Друга умова, з урахуванням дисперсійних властивостей системи можна записати, як умова рівності довжини антени з ємнісними включеннями трьом чвертям довжини мінімальної робочої хвилі. При цьому замість довжини хвилі у вільному просторі слід брати фазову довжину хвилі в дисперсійній лінії з ємнісними включеннями:

$$l = 3/4 \Lambda_{min} = 3/4 * \lambda_{min} / \sqrt{1 - (\lambda_{min}/\lambda_{кр})^2}$$

Дане співвідношення визначає мінімальну довжину хвилі, при якій діаграма спрямованості залишається не порізаною. На більш коротких хвилях в діаграмі спрямованості з'являться провали. Кількість ємнісних вставок визначають два обмеження [8,9]:

- довжина кожної секції антени повинна бути коротше $\lambda_{max} / 8$;
- критична частота дорівнює:

$$\omega_{кр} = 1 / \sqrt{L_o C_B} \quad \text{при } C_B = C_o.$$

Перше обмеження визначає необхідну мінімальну кількість ємнісних вставок, при якому антена буде мати властивості дисперсійної системи передачі електромагнітної енергії. Така структура випромінювача має широкосмугові властивості. При подальшому зменшенні кількості ємнісних вставок антена буде представляти систему з резонансними властивості.

Друге обмеження визначає максимальну кількість ємнісних вставок для отримання необхідного значення критичної частоти в заданому діапазоні частот.

Іншими словами, перша умова накладає обмеження знизу, а друге - зверху на необхідну кількість ємнісних вставок для реалізації діапазонних властивостей антен.

Виходячи з вищевказаних умов відкриваються два принципово різних шляхи конструювання антен з ємнісними вставками, вибір яких повинен визначатися її конкретним призначенням. Можна, не змінюючи геометричної довжини антени, домогтися (за допомогою зміни $\lambda_{кр}$) поліпшення діапазонних властивостей антени. Або, не змінюючи критичну довжину хвилі $\lambda_{кр}$, змінювати геометричну довжину антени, домагатися розширення її діапазонних властивостей і поліпшення деяких електричних характеристик. Можливий також розрахунок антен з ємнісними вставками, заснований на комбінації цих шляхів.

При конструюванні першим шляхом, з метою збільшення діапазонних властивостей вібраторних антен з ємнісними вставками, доцільно зрушувати вправо критичну частоту у порівнянні з необхідної мінімальної частотою робочого діапазону f_{min} . Тоді мінімальна частота f_{min} буде визначатися нерівністю:

$$f_{min} < 1 / \sqrt{L_o C_B}.$$

Вхідний реактивний опір ємнісного характеру в частотній області $f_{min} \dots f_{max}$ можна компенсувати узгоджувальною індуктивністю, підключеної до основи антени:

$$\omega_{кр} / \omega_{max} = C_o / C_B = 1 \dots 0,25 \text{ -- } \omega_{max} \leq 4 \omega_{кр}.$$

Такий шлях розробки антен для збільшення їх діапазонних властивостей дозволяє збільшити верхню межу діапазону антени при заданій іншій її довжині. Це впливає з того, що дисперсійні властивості вібратора проявляються активно при співвідношенні ємності секції та ємності включення 0,25 ... 1, тобто діапазон робочих частот, в якому буде виконуватися вищевикладене, складе до $4\omega_{кр}$.

Отже, змінюючи $\omega_{кр}$, можна збільшити максимальну частоту, на якій зберігається однопелюсткова ДН. Вищевикладені міркування є новими і дозволяють підійти до проектування антенних пристроїв з нових позицій.

Мета експерименту полягала в дослідженні діапазонних властивостей вібраторних антен з ємнісними вставками УКХ діапазону, що використовуються для рухомих комплексів радіозв'язку нового покоління спеціального призначення. Відповідно до цільового використання вібраторних антен з ємнісними вставками необхідно було отримати необхідні параметри вібраторів і перевірити працездатність їх у цілому згідно з технічним завданням.

Експериментальне дослідження виконувалося у Запорізькому науково-дослідному інституті радіозв'язку (ЗНДІРЗ) з вивчення електричних характеристик таких антен.

Відповідно до поставленої мети були розроблені і досліджені декілька варіантів штирьових антен з ємнісними вставками, кожне з плечей яких складалося з окремих секцій, з'єднані між собою послідовно через ємнісні зв'язку.

Вимірювання електричних характеристик проводилося на полігоні ЗНДІРЗ при установці досліджуваних зразків в центрі металевого листа розміром 2,5×2,5 м на висоті 1,7 м від поверхні землі та на поверхні даху кузова рухомого комплексу радіозв'язку спеціального призначення.

Дослідницька робота проводилася в польових умовах, де у діапазоні частот вимірювалися основні характеристики штирьових антен АШ-3М з ємнісними вставками: вхідний опір та

коефіцієнт біжучої хвилі на вході антени. Вимірювання зводилися за допомогою вимірювача комплексних опорів.

Вимірювання відносного коефіцієнта підсилення досліджуваних антен проводилося шляхом порівняння їх випромінювання з випромінюванням чверть хвильового штиря при однакових поданих на них потужностях.

При визначенні параметрів використовувалися методи вимірювань, наведені в інструкціях до відповідних приладу і відомі з науково-технічної літератури.

Виходячи з отриманих закономірностей, були теоретично розраховані в залежності від довжини секції антени і товщини вібратора та досліджені шість типів антен з ємнісними вставками. При експериментальних дослідженнях розглянуто вплив зміни відстані між ємнісними вставками на діапазонні властивості в смузі робочих частот 30...100 МГц для двох макетів вібраторних антен з двома ($L = 3$ м), трьома ($L = 3$ м) ємнісними вставками. Зміни дисперсійних властивостей представлені на рис. 5 [11-14].

В якості вихідних даних при розробці використані не тільки вимоги до електричних характеристик і параметрів антен, але в окремих випадках і технологічні вимоги (наприклад, вибір типу випромінювача і ємнісних вставок, умови з'єднання антен з збудливим фідером).

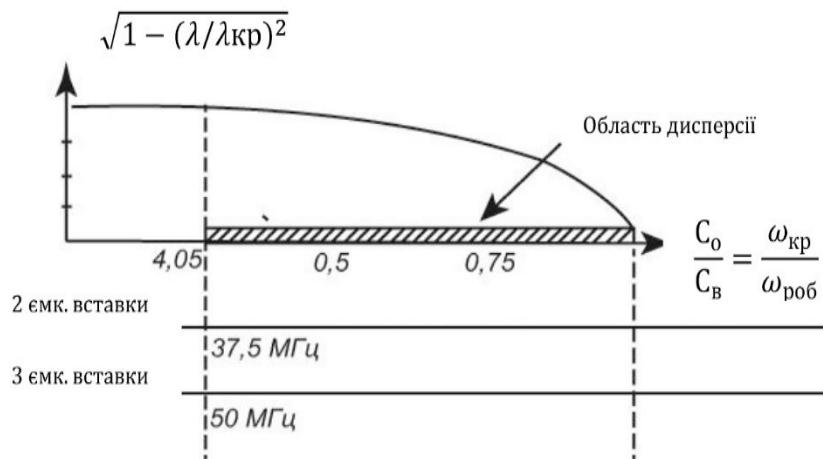


Рис. 5. Характер дисперсійних властивостей антен з ємнісними вставками

Ємнісні вставки представляли собою циліндричні конденсатори, розміщені рівномірно по всій довжині антени, замість з'єднання сталевих колін. В якості діелектрика використовувався текстоліт. Вібратори з двома ємнісними вставками ($C_1 = 7...9$ пФ, $C_2 = 10...18$ пФ) і трьома ємнісними вставками ($C_1 = 7...8$ пФ, $C_2 = 10...12$ пФ, $C_3 = 11...13$ пФ) встановлювалися в центрі металевих листів розміром $2,5 \times 2,5$ м на висоті 1,7 м від поверхні землі і підключалися до вимірювального приладу комплексних опорів через індуктивність 0,23 мГн.

Результати експериментальних досліджень: КСХ; відносного коефіцієнта підсилення (G_A) штирьових антен двох типів представлені на рис. 6. (а, б).

У процесі проведення вимірювань розроблений і четвертьхвильовий випромінювачі почерзі встановлювалися в центрі металевих листів. Коефіцієнт посилення піддослідних антен з урахуванням ступені узгодження (КСВ) показаний на рис. 6.б.

Графічні залежності результатів порівняльного аналізу експериментальних і теоретичних залежностей, отриманих методами чисельного інтегрування на основі розробленої методики розрахунку параметрів несиметричної (симетричної) антени з ємнісними вставками та оцінка її електричних характеристик, призводить до незначних розбіжностей і допустимі для практики [11-14].

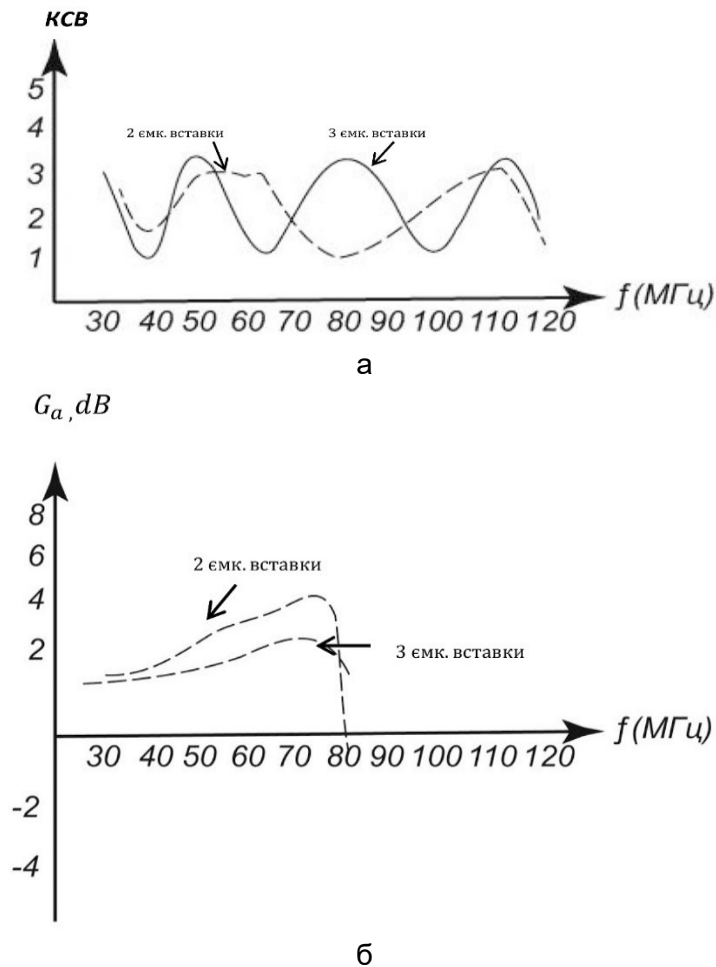


Рис. 6. Результати експериментального дослідження діапазонних властивостей електричних параметрів розроблених антенних пристроїв при установці випромінювачів на даху мобільного вузла радіозв'язку: а) коефіцієнта стоячої хвилі; б) коефіцієнта підсилення

Висновки

1. Запропоновано новий підхід до розуміння фізичних процесів, що сприяють поліпшенню діапазонних властивостей штирьових антен з ємнісними вставками. Наводиться розв'язок задачі поширення електромагнітної енергії в однорідній двохзв'язної лінії передачі з реактивними включеннями. Її рішення дозволило встановити нові закономірності, а також по-новому підійти до розрахунку вібраторних антен з ємнісними вставками. Рішення проведено з використанням апарату теорії лінії передачі і дозволило виявити наявність критичної частоти (довжини хвилі) і дисперсії електромагнітних хвиль в двохзв'язних лініях передачі з включеннями.

2. Проведені дослідження діапазонних властивостей антен з ємнісними вставками дозволили встановити три крайніх випадку розподілу в них струму: експоненціальне, синусоїдальне і лінійне, отримати вирази для розрахунку внутрішніх і зовнішніх характеристик таких антен, виявити залежності їх електричних характеристик від геометричних розмірів і сформулювати рекомендації щодо раціонального вибору величин ємнісних ставок і місць їх включення.

3. Розроблено метод розрахунку вібраторних антен з ємнісними вставками, сутність якої полягає в раціональному використанні дисперсійних властивостей таких структур в частотній області для створення широкосмугових антен з вставками.

4. Запропоновано нові конструкції антен з ємнісними вставками, ряд з яких впроваджений у виробництво і експлуатацію при установці на спеціальних приземних об'єктах, що забезпечують їх радіозв'язок з мобільними комплексами радіозв'язку нового покоління спеціального призначення. Результати випробувань таких антен показали, що їх коефіцієнт посилення на частотах, близьких до критичної, в два та більше разів перевищує коефіцієнт посилення четверть

хвильового випромінювача і, отже, коефіцієнт корисної дії розроблених антен також зріс. Випробування вібраторних антен з ємнісними вставками проводилися на випробувальному полігоні Запорізькому науково-дослідному інституті радіозв'язку.

Література

1. Маттей Г.Л., Янг Л., Джонс Е.М.Т. Фильтры СВЧ, согласующие цепи и цепи связи. М.: Связь, 1972. Т.2. 496с.
2. Овсяников В.В. Вибраторные антенны с реактивными нагрузками. М.: Радио и связь, 1985. 120с.
3. Левин Б.М., Яковлев А.Д. Антенна с нагрузками как импедансный вибратор с применимым импедансом. *Радиотехника и электроника*. Москва, 1985. Вып. 1. С. 25-33.
4. Сайко В.Г., Артемьев В.А., Ильинов М.Д., Федяев В.Е. Вибраторные антенны с реактивными включениями. *Зарубежная радиоэлектроника*. Москва, 1991. №11. С.105-108.
5. Сайко В.Г., Артемьев А.В., Федяев В.Е. Состояние и перспективы развития вибраторных антенн с реактивными вставками. *Изв. вузов Радиоэлектроника*. Киев, 1993. № 5. С. 17-28.
6. Сайко В.Г. Аналитический обзор состояния развития штыревых антенн с реактивными вставками. *Зв'язок*. Киев, 1996. № 2. С.42-36.
7. Сайко В.Г. Расчет параметров вибраторной антенны с емкостными вставками. *Изв. вузов Радиоэлектроника*. Киев, 1991. № 1. С. 17-22.
8. Сайко В.Г. Методика расчета вибраторной антенны с емкостными вставками. *Изв. вузов Радиоэлектроника*. Киев, 1992. № 7. С. 19-25.
9. Сайко В.Г., Мазор С.Ю. Методика расчета проволочных антенн с реактивными вставками. *Научно-методический сборник*. Киев, НИЦ при КВВИУС. 1991. №2. С. 14-21.
10. Сайко В.Г., Мазор С.Ю. Диапазонные свойства штыревых антенн с реактивными вставками. *Научно-методический сборник*. Киев, НИЦ при КВВИУС. 1991. №2. С. 5-13.
11. Сайко В.Г. Широкополосные УКВ антенны с включениями для подвижных объектов. *Техника средств связи. сер. ТРС*. Воронеж, 1990. Вып.7. С. 146-150.
12. Сайко В.Г. Низькопрофільні антени для мереж мобільного зв'язку. *Сучасні тенденції розвитку технологій в інфокомунікаціях та освіті*: Матеріали докладів 6-ї наукової конференції, м. Київ, ДУТ, 05-06 листопада 2009. Київ. 2009. с. С.263.
13. Сайко В.Г., Ильинов М.Д. Широкополосные малогабаритные антенны УКВ диапазона. *Зарубежная радиоэлектроника*. Москва, 1990. № 2. С. 54-60.
14. Сайко В.Г., Ильинов М.Д. Современное состояние исследований малогабаритных антенн. *Зарубежная радиоэлектроника*. Москва, 1990. № 5. С. 82-87.

References

1. Matvey G.L., Yang L., Jones E.M.T. Filters SVCh, soglashuyushchie tsepí i tsepí svyazi. M.: Svyaz', 1972. T. 2. 496s.
2. Ovsyanikov V.V. Vibratornye anteny s reaktivnymi nagruzkami. M.: Radio i svyaz', 1985. 120s.
3. Levin B.M., Yakovlev A.D. Antenna s nagruzkami kak impedansnyy vibrator s primenimym impedansom. *Radiotekhnika i elektronika*. Moskva, 1985. Vyp. 1. S. 25-33.
4. Sayko V.G., Artemyev V.A., Ilyinov M.D., Fedayev V.E. Vibratornye anteny s reaktivnymi vklucheniyami. *Zarubezhnaya radioelektronika*. Moskva, 1991. №11. S.105-108.
5. Sayko V.G., Artemyev A.V., Fedayev V.E. Sostoyanie i perspektivy razvitiya vibratornykh anten s reaktivnymi vstavkami. *Izv. vuzov Radioelektronika*. Kiev, 1993. № 5. S. 17-28.
6. Sayko V.G. Analiticheskiy obzor sostoyaniya razvitiya shtyrevykh anten s reaktivnymi vstavkami. *Zv'yazok*. Kiev, 1996. № 2. S.42-36.
7. Sayko V.G. Raschet parametrov vibratornoy anteny s emkostnymi vstavkami. *Izv. vuzov Radioelektronika*. Kiev, 1991. № 1. S. 17-22.
8. Sayko V.G. Metodika rascheta vibratornoy anteny s emkostnymi vstavkami. *Izv. vuzov Radioelektronika*. Kiev, 1992. № 7. S. 19-25.
9. Sayko V.G., Mazor S.Yu. Metodika rascheta provolochnykh anten s reaktivnymi vstavkami. *Nauchno-metodicheskiy sbornik*. Kiev, NITS pri KVVIUS. 1991. №2. S. 14-21.
10. Sayko V.G., Mazor S.Yu. Diapazonnye svoystva shtyrevykh anten s reaktivnymi vstavkami. *Nauchno-metodicheskiy sbornik*. Kiev, NITS pri KVVIUS. 1991. №2. S. 5-13.
11. Sayko V.G. Shirokopolosnye UKV anteny s vklucheniyami dlya podvizhnykh ob"ektov. *Tekhnika sredstv svyazi. ser. TRS*. Voronezh, 1990. Vyp.7. S. 146-150.
12. Sayko V.G. Nizkoprofilni anteni dlya merezh mobilnogo zv'yazku. *Suchasni tendentsii rozvitku tekhnologii v infokomunikatsiyakh ta osviti*: Materiali dokladiv 6-i naukovoi konferentsii, m. Kiiv, DUT, 05-06 listopada 2009. Kiiv. 2009. S.263
13. Sayko V.G., Ilyinov M.D. Shirokopolosnye malogabaritnye anteny UKV diapazona. *Zarubezhnaya radioelektronika*. Moskva, 1990. № 2. S. 54-60.
14. Sayko V.G., Ilyinov M.D. Sovremennoe sostoyanie issledovaniy malogabaritnykh anten. *Zarubezhnaya radioelektronika*. Moskva, 1990. № 5. S. 82-87.

ІНФОКОМУНІКАЦІЙНІ ТЕХНОЛОГІЇ

УДК 621.396.94

РЕГІОНАЛЬНА НАВІГАЦІЙНА СУПУТНИКОВА СИСТЕМА ЯК ДЖЕРЕЛО PNT ПІД ЧАС НАВМИСНИХ ТРИВАЛИХ ЗАГРОЗ СИГНАЛАМ GNSS

DOI 10.36994 / 2788-5518-2023-01-05-08

Гайдаманчук В.А., к.т.н. WIRCOM, Київ, Україна. gva@wircom.ua

Матеїєнко М.В., WIRCOM, Київ, Україна.

Пастушенко І.М., WIRCOM, Київ, Україна.

Семенко А.І., д.т.н., проф. Відкритий міжнародний університет розвитку людини «Україна», Київ, Україна. setel@ukr.net

Анотація: Статтю присвячено розгляду проблем щодо забезпечення споживачів PNT (Position, Navigation, Time) інформацією в умовах зростання загроз стабільному функціонуванню систем GPS та інших систем GNSS. Розглянута можливість створення і використання, як альтернативного, більш захищеного джерела сигналів PNT для споживачів, задіяних в забезпеченні критичної інфраструктури, регіональної супутникової системи навігації ((Regional Navigation Satellite System - RNSS), яка була б більш ніж GNSS захищена від навмисних загроз та могла би бути доступною на всій території України.

Розглянуто можливості та основні обмеження використання вже існуючих супутників на геосинхронних орбітах (GEO, IGSO) в якості ретрансляторів PNT інформації у разі, якщо такі сигнали будуть створюватись і передаватись через них спеціалізованими наземними станціями.

Розглянуто особливості адаптації і вдосконалення таких технічних рішень для умов України, можливості застосування в складі RNSS модернізованих станцій VSAT зв'язку та супутників на геостационарних орбітах з використанням широкосмугових сигналів (ШСС) з розширенням спектру методом прямої послідовності (Direct Sequence Spread Spectrum – DSSS) та завадостійкого кодування, що суттєво зменшить можливість їх виявлення та ускладнить протидію їх роботі.

Ключові слова: Навігація, Jamming, Spoofing, GPS, GNSS, RNSS, PNT, NAVSOP, GEO, IGSO, VSAT.

REGIONAL NAVIGATION SATELLITE SYSTEM AS PNT SOURCE DURING INTENTIONAL LONG-TERM THREATS TO GNSS SIGNALS

Viktor A. Gaidamanchuk, Ph.D., WIRCOM, Kyiv, Ukraine. gva@wircom.ua

Mykola V. Matvienko, WIRCOM, Kyiv, Ukraine.

Ihor M. Pastushenko, WIRCOM, Kyiv, Ukraine.

Anatoliy Semenko, Dr.habil., Prof. Open International University of Human Development "Ukraine", Kyiv, Ukraine. setel@ukr.net

Abstract: The article is devoted to consideration of the possibility, in the conditions of growing threats to the stable functioning of GPS and other GNSS systems due to long-term deliberate interference, to use as an alternative, more protected source of PNT (Position, Navigation, Time) signals, a local satellite navigation system, which will be capable to function, providing coverage of all or a significant part of the territory of Ukraine.

The possibilities and limitations of using already existing satellites in geosynchronous orbits (GEO, IGSO) as relayers of PNT information in the event that such NAVSOP signals will be created and transmitted through them by specialized ground stations.

The peculiarities of adaptation and improvement of such technical solutions for the conditions of Ukraine are considered along with the possibility of using VSAT stations adapted to such a system, which are part of satellite communication systems using satellites in geostationary orbits with the use of interference-resistant coding and DSSS (Direct Sequence Spread Spectrum) signals that will significantly reduce the possibility of their detection and make it more difficult to counter their reception during long-term threats to GNSS signals.

In addition, the implementation of authentication algorithms in such systems will make it possible to ensure their protection against detection and alteration of data transmission, and to increase the accuracy of PNT determination and expand the scope of application in the future, the regional navigation satellite system (RNSS) can be transformed

into a hybrid one by supplementing it with additional sources of PNT and NAVSOP.

Key words: *Navigation, Jamming, Spoofing, GPS, GNSS, RNSS, PNT, NAVSOP, GEO, IGSO, VSAT.*

Вступ

Системи глобальної навігації GNSS при їх створенні мали стати основним та надійним джерелом інформації PNT для різноманітних, як у промисловості, наукових дослідженнях та обороні, так і для широкого кола споживачів. І якщо не брати до уваги чутливості таких систем до ненавмисних і навмисних перешкод, загалом їм вдається забезпечувати необхідної точності глобальну навігацію [1]. Однак, оскільки сигнали GNSS відносно слабкі, досить незначного рівня перешкод для виникнення перебоїв при прийомі сигналів споживачами, що, відповідно, знижує якість позиціонування та синхронізації часу та частоти. У зв'язку з цим наявність навмисних та/або ненавмисних перешкод не дозволяє повною мірою покладатися на GNSS як на завжди доступне джерело PNT, особливо для деяких критично важливих застосувань, таких як критична інфраструктура та оборона. З можливих технічних рішень, які дозволять при цьому покращити доступ до PNT інформації, можна вирізнити ті, що спрямовані на покращення прийому GNSS в умовах вразливості, а саме рішення для виявлення та придушення перешкод на рівні антени, а також на рівні обробки сигналу, а також такі технічні рішення, що використовують при розробці нові або модернізують існуючі джерела PNT, які можуть забезпечити альтернативу GNSS з формування сигналів достатньої якості для забезпечення споживачів інформацією PNT. Такі системи відомі як системи навігації NAVSOP (Navigation via Signals of Opportunity) мають різноманітні конфігурації та використовують радіосигнали від різних джерел та технологій для обчислення місцезнаходження споживача за методом триангуляції. Відповідно, використання в системах NAVSOP існуючих супутників які не включені до складу систем GNSS дозволяє не тільки забезпечувати інформацією PNT регіони із великою площею а також може забезпечувати підвищений рівень захисту в умовах застосування довготривалих навмисних перешкод.

Виконання досліджень

Створення регіональної навігаційної супутникової системи RNSS, сигнали якої можуть бути використані як сигнали NAVSOP в якості альтернативи для глобальних систем навігації GNSS, хоча й істотно менш витратне рішення, ніж створення будь-якої системи навігації з сімейства GNSS, проте передбачає використання групи супутників та наземних станцій, а також спеціалізованих приймачів, що потребує як значних інвестицій у розробку, так і значних адміністративних зусиль при її впровадженні та експлуатації. У зв'язку з цим необхідно зазначити, що не для всіх випадків появи перебоїв з прийомом сигналів GNSS, пов'язаних з навмисними перешкодами, подібне рішення є єдино можливим, і деякі проблеми такого типу можуть бути вирішені значно простіше й з меншими витратами. Розглядаючи більш детально причини та тривалість критичного погіршення прийому сигналів GNSS, при яких можна обійтися без їх заміни на сигнали RNSS, можна виділити з них кілька, коли цілком достатньо застосування більш простих і менш витратних рішень.

Так у разі нетривалих за часом проблем із прийомом сигналів GNSS досить ефективно дозволить мінімізувати проблему сукупність наступних факторів, врахованих при побудові комплексної системи навігації: наявність антени та приймача сигналів GNSS з опцією захисту від «Jamming&Spoofing» [2], наявність інерційної системи у складі системи навігації і високоточного внутрішнього опорного генератора, – що в сукупності дозволить захистити систему навігації від короткочасної відсутності прийому сигналів GNSS і в тому числі завдяки можливості забезпечити точність часу не гірше 100 нс (на час відсутності сигналів GNSS) при використанні кварцових генераторів HP OCXO (до 1 години) або атомних генераторів “Chip Scale Atomic Clock” (до 2 годин) [4].

Також, наприклад, у разі такого розташування джерел навмисних радіоперешкод, при якому напрямок поширення сигналу істотно відрізняється від напрямку поширення сигналів GNSS (наприклад при їх наземному розташуванні). Для таких випадків існують ефективні технології, які дозволяють за рахунок використання сучасних антенних систем мінімізувати негативний вплив на прийом GNSS навмисних перешкод у більшості випадків. Для формування діаграми спрямованості антени використовуються адаптивні багатоелементні комбіновані антени, так звані антени з регульованою діаграмою спрямованості, де створюється такий промінь діаграми спрямованості антени, який має нулі в напрямку небажаних сигналів або максимальне посилення сигналу в напрямку потрібних сигналів супутників, завдяки чому мінімізується вплив передавача перешкод на приймачі GNSS [5;6] або більш універсальна альтернатива – місцева NAVSOP з використанням наземних сигналів мереж радіо та мобільного зв'язку, Wi-Fi, телебачення та низькоорбітальних супутників (LEO).

Але у ситуаціях, коли користувачі PNT знаходяться в умовах, коли прийому сигналів GNSS протидіють сучасні систем РЕБ, і вони виявляються повністю ізольованими від сигналів GPS і територія, де це відбувається є значною за площею, потрібно забезпечити повну заміну джерела надання такої інформації на систему, куди повинні входити як більш приховані і завадостійкі джерела сигналів NAVSOP, так і спеціальні приймачі, на які в таких ситуаціях повинні перемикатися споживачі, що дозволить приймати і

обробляти інформацію PNT, що надходить [7]. Отже, необхідність забезпечення споживачів сигналами системи NAVSOP рівня RNSS доцільна в тих випадках, коли протидія прийому сигналів GNSS є довготривалою, впливає на велику кількість критично важливих споживачів, розташованих на великій території.

Якщо розглянути наявні супутники різного призначення, то одним з найбільш придатних для адаптації та подальшого використання в якості RNSS є системи супутникового зв'язку, які складаються з геосинхронних супутників зв'язку разом з відповідною наземною апаратурою та VSAT станціями споживачів [8].

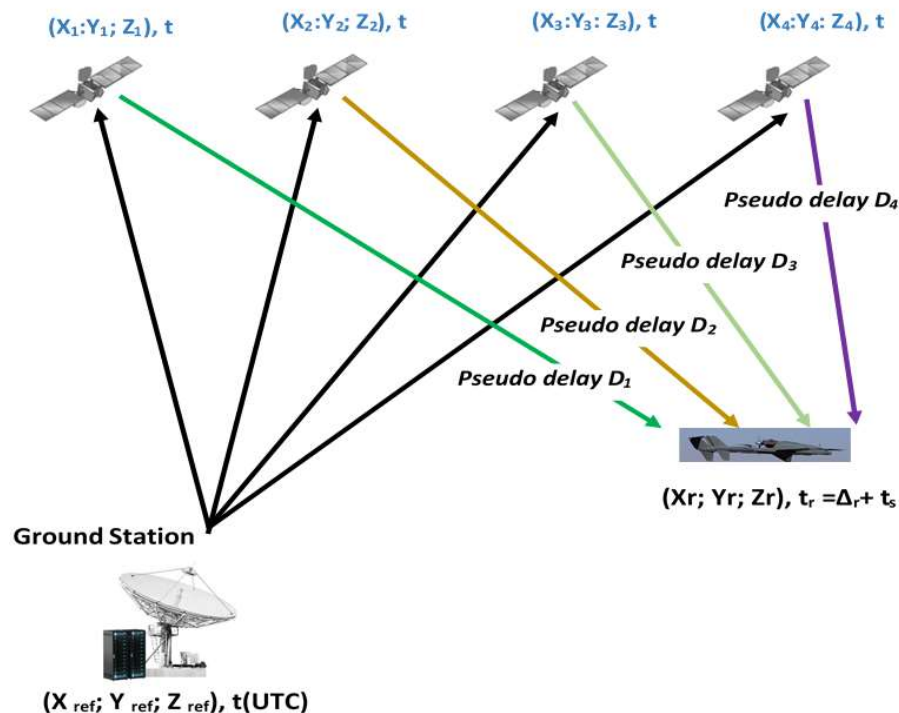


Рис.1. Схема передачі PNT інформації з використанням супутникової системи зв'язку

При використанні супутників зв'язку GEO для розрахунків розповсюдження PNT у складі RNSS (рис. 1) як і у випадку GNSS в загальному вигляді можна застосовувати систему рівнянь (1) [8], де D_i – псевдодальність; x_i, y_i, z_i – позиція супутника; x_r, y_r, z_r – позиція споживача; c – швидкість світла; Δ_r – зсув годинника приймача відносно UTC:

$$\begin{aligned} cD_1 &= [(x_1 - x_r)^2 + (y_1 - y_r)^2 + (z_1 - z_r)^2]^{1/2} + c\Delta_r \\ cD_2 &= [(x_2 - x_r)^2 + (y_2 - y_r)^2 + (z_2 - z_r)^2]^{1/2} + c\Delta_r \\ cD_3 &= [(x_3 - x_r)^2 + (y_3 - y_r)^2 + (z_3 - z_r)^2]^{1/2} + c\Delta_r \\ cD_4 &= [(x_4 - x_r)^2 + (y_4 - y_r)^2 + (z_4 - z_r)^2]^{1/2} + c\Delta_r \end{aligned} \quad (1).$$

Основна системна відмінність в функціонуванні RNSS полягає в тому, що на відміну від систем GNSS, в яких інформація PNT, що надходить споживачу, формується безпосередньо на супутнику, в RNSS інформація PNT формується в наземній базовій станції і передається споживачу через супутник зв'язку. Ці системи не потребують при їх створенні використання нових спеціалізованих супутників, а використовують вже існуючі телекомунікаційні супутники, що не потребує аж занадто великих інвестицій порівняно з глобальними навігаційними супутниковими системами GNSS. Крім того, локальна система навігації такого типу може використовувати існуючі супутники в режимі, який не заважає виконанню їх основних функцій. Щодо наявності таких супутників, то згідно з даними UCS Satellite Database на геосинхронних орбітах зараз сумарно знаходиться більше 500 супутників. Разом з тим, треба зазначити, що RNSS в забезпеченні точності PNT, як правило, поступаються системам GNSS. І тому дуже важливо, щоб при їх створенні вибір супутників, з яких складається RNSS, відбувався з врахуванням концепції DOP (Dilution of Precision), згідно з якою на погіршення точності супутникової навігації суттєво впливає не оптимальна геометрія взаємного розташування супутників. Досягти такої геометрії розташувань, як в спеціально створених супутникових системах навігації, таких як GNSS, і відповідно такого ж рівня точності PNT в системах RNSS дуже складно.

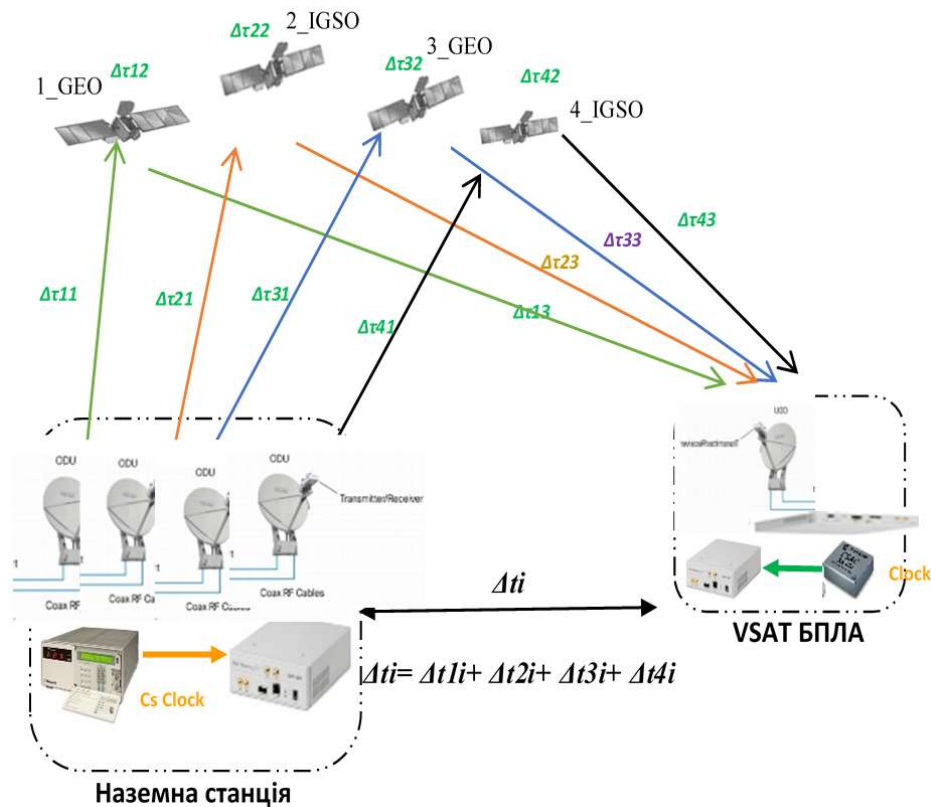


Рис. 2. Навігаційна система RNSS з використанням супутників GEO та IGSO.
Зв'язок між показниками DOP та точністю супутникових систем навігації видно з наступного виразу:

$$\text{Помилка PNT} = \text{DOP} * \text{UERE} \quad (2)$$

де UERE (User Equivalent Range Error) – загальна помилка визначення псевдодальності, яка визначає похибку відстані між приймачем споживача PNT і конкретним супутником, який входить до складу супутникової системи навігації. Для забезпечення точності позиціонування згідно з розрахунками показників DOP необхідний оптимальний вибір геометрії їх взаємного розташування при поєднанні разом з супутниками GEO супутників IGSO. Такий принцип, наприклад, використовується в китайській регіональній супутниковій навігаційній системі CAPS (Chinese Area Positioning System) [9]. Космічний сегмент CAPS складається не тільки з декількох супутників GEO, а також використовує декілька супутників IGSO на нахиленій геосинхронній орбіті. При цьому, як видно з Таблиці 1, найкращих показників DOP і, відповідно, точності RNSS можливо досягти лише при спільному використанні в RNSS 6 супутників та забезпеченні максимального показника доступності RNSS, який залежить від географічних координат країни чи регіону покриття сигналами [10].

Таблиця 1.
Значення DOP в залежності від складу RNSS

Склад RNSS	Доступність RNSS	DOP ^G	DOP ^P	DOP ^H
4*1 GSO 2* GEO	100%	4.93	4.3	2.64
2*1 GSO 4* GEO	100%	6.1	5.71	4.09

Таким чином, космічний сегмент RNSS може бути створений у складі супутників зв'язку GEO та IGSO у кількості та геометрії взаємного розташування виходячи з мінімально можливих показників DOP з врахуванням того, що для території України доступність GEO/IGSO RNSS буде складати ~ 92%.

Разом з тим дуже важливою особливістю розробки системи RNSS є можливість при її створенні реалізувати спеціальні заходи захисту від виявлення сигналів (прихованості) системи RNSS та від

різноманітних перешкод. Сигнали ШСС DSSS [11], які використовуються в GPS, є надзвичайно пристосованими для протидії перешкодам за своєю природою, але в умовах функціонування GNSS ці їх властивості не можуть бути використані в повному обсязі, як це можливо при створенні систем RNSS. Якщо взяти до уваги такі параметри як прихованість та завадостійкість і звернути увагу на рівняння, якими їх можна описати, то побачимо, що завадостійкість DSSS визначається формулою (3), що зв'язує відношення сигнал-перешкода на виході приймача споживача q^2 з відношенням сигнал-перешкода на вході приймача ρ^2 :

$$3q^2 = 2B\rho^2 \quad (3)$$

а прихованість або час виявлення T визначається формулою:

$$T \approx F \cdot (\rho^2)^{-2} \cdot 2 \cdot q^4 \quad (4)$$

$$\text{для } F\Delta t > 1 \quad \Delta t \approx 2d^2/Rc$$

де R – відстань між передавачем і приймачем, F – ширина спектру, B – база DSSS.

Таким чином, чим ширший спектр сигналів PCCH, тим вища завадостійкість та прихованість як сигналів, так і системи навігації в цілому. Виходячи з того, що однією з основних причин створення систем RNSS є неможливість сталого прийому GNSS в умовах довготривалих навмисних перешкод, цілком зрозуміло, що особлива увага при створенні RNSS повинна бути приділена забезпеченню стійкості від перешкод та мінімізації можливості виявлення функціонування системи RNSS тобто її прихованості.

Крім цих питань, при створенні RNSS дуже важливою є модернізація VSAT станцій, які сумісні з відібраними для RNSS супутниками GEO/IGSO для забезпечення роботи наземної станції RNSS з відповідним рівнем системи шкали часу [12] та приймачів споживачів RNSS з використанням технологій адаптивних ФАР [13] (рис. 3).

Також враховуючи специфіку розташування та дрейфу на орбіті супутників GEO/IGSO при створенні RNSS дуже важливо забезпечити калібрування та тестування основних системних параметрів на основі відповідних методик, одна із яких схематично представлена на рис. 3.

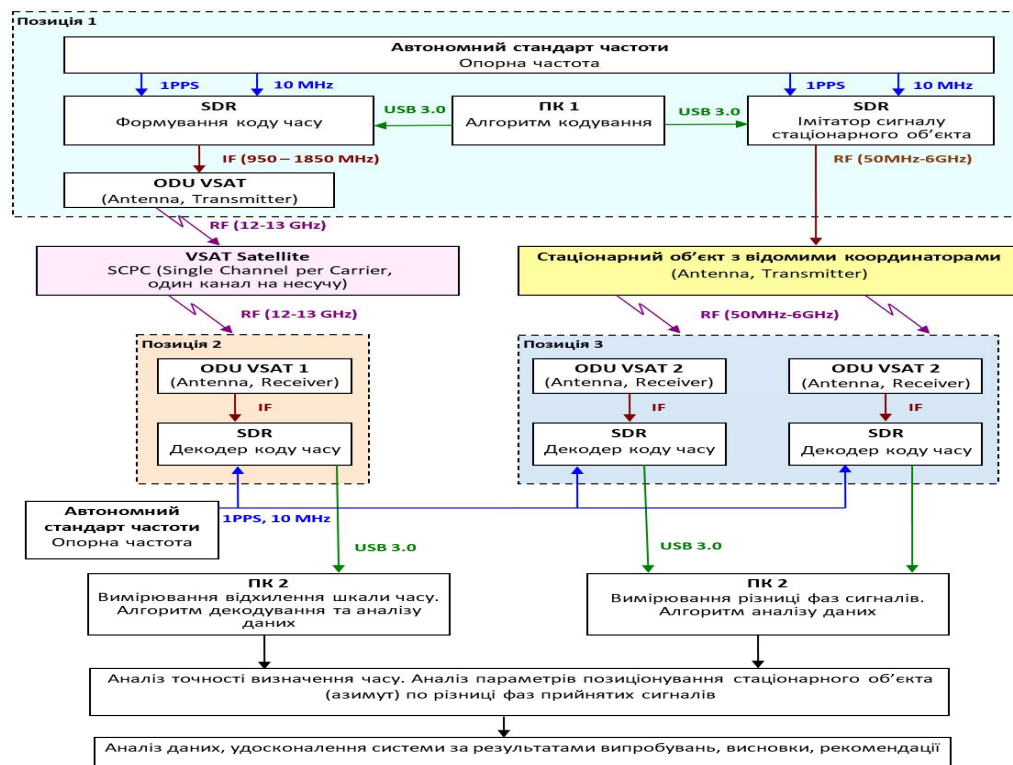


Рис.3. Схема проведення досліджень точності параметрів позиціонування RNSS

Треба також зазначити, що в Україні накопичений успішний досвід [14] при розробці та виробництві НВО «Сатурн» (м. Київ) VSAT «Епіграма» на основі ШСС DSSS (рис. 4), який може бути продуктивно використаний при створенні системи навігації RNSS для потреб України та дозволить суттєво зменшити витрати та терміни розробки.

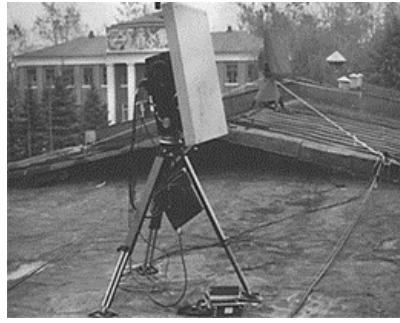


Рис. 4. Наземна станція (VSAT) супутникового зв'язку GEO «Епіграма» (Україна) з використанням ШСС DSSS сигналів та адаптивної ФАР

Висновки

Зважаючи на те, що глобальні навігаційні системи GPS (GNSS) є вразливими до навмисних перешкод, в випадках коли ці перешкоди довготривалі, поширені на великих територіях та регулярно створюють загрози прийому споживачам PNT інформації, доцільним є створення і використання регіональних навігаційних систем з використанням супутників GEO/IGSO. Такі системи працюють більш ефективно на широтах, де доступність до супутників з яких вони складаються є не менше 80% та забезпечують посилену захищеність від довготривалих навмисних перешкод Jamming, Spoofing та інших перешкод РЕБ за рахунок застосування ШСС DSSS, алгоритмів зміни псевдовипадкових кодів та завадостійкого кодування. Крім того, введення алгоритмів автентифікації в таких системах дозволить забезпечити їх захищеність від виявлення та підміни передачі даних. Для підвищення точності в подальшому система RNSS може буде перетворена в гібридну RNSS за рахунок доповнення додатковими джерелами PNT

Література

1. M.S. Grewal, A.P. Andrews, C.G. Bartone. Global Navigation Satellite Systems, Inertial Navigation, and Integration. – Wiley, 2020. – 568 p.
2. J.J. Spilker. Digital Communications by Satellite. – Prentice-Hall Inc., 1977. – 592 p.
3. J. Zhang, X. Cui, H. Xu, M. Lu. A Two-Stage Interference Suppression Scheme Based on Antenna Array for GNSS Jamming and Spoofing, Journal, Sensors, 2019, 19 (18).
4. Low-Noise Chip-Scale Atomic Clock. Datasheet, rev H. – Microchip Technology Inc., 2019.
5. K. Hünerbein, W. Lange Timing and Navigation in UAVs: Synchronization of UAV Swarms and Testing GPS Error Effects on GNSS Reception. - Embedded Selforganizing Systems, Vol 5. No 1. 2018, pp. 17-25.
6. S. Albrektsen, T. Johansen User-Configurable Timing and Navigation for UAVs. - Sensors, 2018, 18, 2468
7. G. Y. Ma, Q.T. Wan, T. Gan. Communication-based positioning systems: past, present and prospects. – Research in Astron. Astrophys. 2012 Vol. 12 No. 6, pp. 601–624.
8. J.W Betz. Engineering Satellite-Based Navigation And Timing. – Wiley, 2016. – 640 p
9. B. Li, A. Dempster. China's Regional navigation Satellite System – CAPS// Inside GNSS – June 2010, pp. 59-63.
10. F. Keyvani, S.Torabi. Design and Simulation of Regional Navigation Constellation with Optimized Mean DOP based on Hybrid GEO and IGSO Satellites. International Journal of Aviation, Aeronautics, and Aerospace, 6(2). – 2019, pp.1-18.
11. Л. Е. Варакин Системы связи с шумоподобными сигналами. – Радио и связь, 1985. – 384 с.
12. V. Gaidamanchuk, A. Semenko et al. Peculiarities of TWSTFT modem development for national time dissemination system, – IEEE International Conference Radio Electronics & Info Communications (UkrMiCo), Kyiv – 2016.
13. R.A. Monzingo, T. W. Miller. Introduction to Adaptive Arrays. – John Wiley&Sons, 1980. – 240 p.
14. V. Gaidamanchuk, M. Matvienko. Main principles of the satellite system providing PNT information for moving objects under GNSS vulnerability, Journal of Information and Telecommunication Sciences/ Igor Sikorsky Kyiv Polytechnic Institute, 2017, - Vol. 9, No. 2, pp. 27-31

References

1. M.S. Grewal, A.P. Andrews, C.G. Bartone. Global Navigation Satellite Systems, Inertial Navigation, and Integration. Wiley, 2020. – 568 p.
2. J.J. Spilker. Digital Communications by Satellite. – Prentice-Hall Inc., 1977. – 592 p.
3. J. Zhang, X. Cui, H. Xu, M. Lu. A Two-Stage Interference Suppression Scheme Based on Antenna Array for GNSS Jamming and Spoofing, Journal, Sensors, 2019, 19.

4. Low-Noise Chip-Scale Atomic Clock. Datasheet, rev H. – Microchip Technology Inc., 2019.
5. K. Hünerbein, W. Lange Timing and Navigation in UAVs: Synchronization of UAV Swarms and Testing GPS Error Effects on GNSS Reception, *Embedded Selforganizing Systems*, 2018, Vol 5. No 1. - pp. 17-25.
6. S. Albrektsen, T. Johansen User-Configurable Timing and Navigation for UAVs. – *Sensors*, 2018, 18.
7. G. Y. Ma, Q.T. Wan, T. Gan. Communication-based positioning systems: past, present and prospects. – *Research in Astron. Astrophys.* 2012 Vol. 12 No. 6, pp. 601–624.
8. J.W Betz. *Engineering Satellite-Based Navigation And Timing*. – Wiley, 2016. – 640 p.
9. B. Li, A. Dempster. China's Regional navigation Satellite System – CAPS// *Inside GNSS* – June 2010, pp. 59-63.
10. F. Keyvani, S.Torabi. Design and Simulation of Regional Navigation Constellation with Optimized Mean DOP based on Hybrid GEO and IGSO Satellites. *International Journal of Aviation, Aeronautics, and Aerospace*, 6(2). – 2019, pp.1-18.4.
11. L. E. Varakyn Системи зв'язу з шумоподобними сигналами. – *Радіо у зв'язі*, 1985. – 384 с.
12. V. Gaidamanchuk, A. Semenko et al. Peculiarities of TWSTFT modem development for national time dissemination system, – *IEEE International Conference Radio Electronics & Info Communications (UkrMiCo)*, Kyiv – 2016.
13. R.A. Monzingo, T. W. Miller *Introduction to Adaptive Arrays*. – John Wiley&Sons, 1980. – 240 p.
14. V. Gaidamanchuk, M. Matvienko. Main principles of the satellite system providing PNT information for moving objects under GNSS vulnerability, *Journal of Information and Telecommunication Sciences/ Igor Sikorsky Kyiv Polytechnic Institute*, 2017, - Vol. 9, No. 2, pp. 27-31.

УДК 519.6, 539.3

МАТЕМАТИЧНА МОДЕЛЬ ВЗАЄМОДІЇ СТАЦІОНАРНИХ SH-ХВИЛЬ З СИСТЕМОЮ КРИВОЛІНІЙНИХ ТРІЩИН У НАПІВПРОСТОРІ

DOI 10.36994 / 2788-5518-2023-01-05-09

Панченко Б.Є. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна.
pr-bob@ukr.net

Буката Л.М. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна.
ygrikhuda@gmail.com

Багачук Д.Г. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна.
bagachukdg@gmail.com

Анотація: Запропоновано метод розв'язання задач математичної фізики для напівнескінчених середовищ, що містять системи криволінійних тріщин-розрізів. Більшість відомих робіт стосується проблем дифракції пружних хвиль на прямих і кругових розрізах. Однак насправді тріщина зазвичай не є прямою чи круглою. Дослідження показали, що кривизна тріщини істотно впливає на значення коефіцієнтів інтенсивності динамічних напружень. Значення цього параметра також залежить від близькості дефектів один до одного, оскільки вони завжди потрапляють в дію відбитої хвилі.

Напружено-деформований стан середовищ зі складними властивостями можна ефективно моделювати за допомогою обчислювальних комплексів у поєднанні з програмними системами. Більшість досліджень присвячено розвитку методу скінченних елементів. Проте метод інтегральних рівнянь дуже ефективний для розв'язання антиплоских задач теорії дифракції. Перевагою цього методу є зменшення кількості просторових змінних, висока швидкість збіжності та можливість використання різноманітних ефективних схем чисельної реалізації. Метод також надає можливість будувати ефективні схеми паралельних обчислень. Розроблено єдиний підхід до розв'язання задачі на основі сингулярних інтегральних рівнянь (CIP). Досліджено відповідні динамічні крайові задачі для защемленої та вільної від сил півплощини. Досліджено вплив кривизни дефекту, їх взаємодії та близькості межі на величину та характер коефіцієнтів інтенсивності динамічних напружень. Паралельні алгоритми дозволяють значно скоротити час обчислень і більш детально проаналізувати характеристики хвильового поля. Поєднання методу SIE призводить до значного підвищення ефективності запропонованого алгоритму. Метод може бути використаний для оцінки впливу різних механічних або геометричних факторів на міцність тіл з дефектами

Ключові слова: крайова задача, хвилі зсуву, сингулярні інтегральні рівняння, паралельні обчислення, система тріщин-розрізів.

MATHEMATICAL MODEL OF THE INTERACTION OF STATIONARY SH-WAVES WITH A SYSTEM OF CURVILINEAR CRACKS IN A HALF-SPACE

Borys Panchenko. State University intellectual technologies and communication, Odesa, Ukraine. pr-bob@ukr.net
Liudmyla Bukata. State University intellectual technologies and communication, Odesa, Ukraine. ygrikhuda@gmail.com
Denys Bahachuk. State University intellectual technologies and communication, Odesa, Ukraine.
bagachukdg@gmail.com

Abstract: A method for solving mathematical physics problems is proposed for semi-infinite media containing systems of curved crack-slits. Most of the studies known from the literature relate to the problems of diffraction of elastic waves on straight and circular cuts. However, in reality, the crack is usually not straight or circular. Studies have shown that the curvature of a crack significantly affects the value of the dynamic stress intensity coefficients. The value of this parameter also depends on the proximity of the defects to each other, since they always fall within the range of the reflected wave.

The stress-strain state of media with complex properties can be effectively modeled by computing complexes in combination with software systems. Most studies are devoted to the development of the finite element method. However, the method of integral equations is very effective for solving anti-planar problems of diffraction theory. The advantage of this method is the reduction in the number of spatial variables, the high speed of convergence, and the possibility of using various efficient numerical solution methods. The method also has the ability to build efficient parallel computing schemes.

A unified approach to solving the problem is developed on the basis of singular integral equations (SIEs). The corresponding dynamic boundary value problems for a clamped and force-free half-plane are investigated. The influence of the defect curvature, their interaction, and the proximity of the boundary on the magnitude and

nature of the dynamic stress intensity coefficients is studied. Parallel algorithms allow to significantly reduce the computation time and analyse the characteristics of the wave field in more detail. The combination of the SIE method, which reduces the dimensionality of the problem by one, and provides significant savings in computing time due to the parallelization of computational procedures, leads to a significant increase in the efficiency of the proposed algorithm. The method can be used to assess the influence of various mechanical or geometric factors on the strength of bodies with defects.

Keywords: boundary value problem, shear waves, singular integral equations, parallel computations, system of cracks-sections.

Вступ

Сучасні конструкції та споруди працюють в умовах не тільки статичних, а й динамічних навантажень. Відомо, що при динамічному навантаженні середовищ з тріщинами ймовірність розвитку цих дефектів суттєво підвищується. Тому з метою оцінки граничного стану таких середовищ важливо з'ясувати вплив інерційного ефекту поширення дефектів.

Дослідженню проблеми динамічного руйнування конструкцій має передувати аналіз хвильових полів у модельних задачах. У зв'язку з цим актуальною є розробка методів вирішення динамічних задач математичної фізики для нескінченних та напівнескінченних середовищ із тріщинами.

З теоретичної точки зору тріщина є математичним розрізом, при переході через який зміщення зазнають розривів. Якщо розриви зсувів заздалегідь невідомі, точний математичний опис хвильового поля, що виникає через наявність тріщини, виявляється дуже складним.

Більшість відомих з літератури досліджень належить до задач дифракції пружних хвиль на прямих та кругових тріщинах-розрізах. Проте насправді тріщина, зазвичай, немає прямолінійної чи кругової форми. І, як показали дослідження, кривизна дефекту істотно впливає на величину динамічних коефіцієнтів інтенсивності напружень. Значення цього параметра залежить також від близькості дефектів один до одного, оскільки вони завжди потрапляють в зону дії відбитої хвилі.

Ця робота присвячена розробці методу вирішення динамічних задач математичної фізики для ізотропного напівпростору, послабленого тунельними криволінійними розрізами. Розглядається стаціонарний хвильовий процес.

Напружено-деформований стан середовищ з ускладненими властивостями може ефективно моделюватись обчислювальними комплексами у поєднанні з програмними системами. Більшість досліджень присвячено розвитку методу скінченних елементів [1]. Проте для вирішення антиплоских задач теорії дифракції [5, 10] суттєву ефективність має метод інтегральних рівнянь [2, 6-9, 11-14]. Перевага цього методу полягає у скороченні числа просторових змінних, досить високій швидкості збіжності та можливості застосування різних ефективних чисельних методів розв'язання [8, 12]. Також метод має можливість побудови ефективних паралельних обчислювальних схем [7, 9].

Виконання досліджень

Розглянемо пружний ізотропний напівпростір, послаблений системою тунельних уздовж осі OZ криволінійних розрізів L_j ($j = \overline{1, K}$) (рис. 1.), де L_j - проста розімкнена дуга Ляпунова з початком у точці a_j і кінцем у точці b_j (припускаємо, що $\cap L_j = \emptyset$).

Вважаємо, що береги розрізів вільні від сил, а переміщення при переході через $L = \cup L_j$ зазнають розриву.

Нехай із нескінченності випромінюється монохроматична хвиля зсуву, нормаль до фронту якої становить кут ψ з віссю OX (залежність від часу виражається множителем $e^{-i\omega t}$):

$$W_0 = \tau e^{-i\gamma_2(x \cos \psi + y \sin \psi)} \quad \tau = \text{const} \quad \gamma_2 = \frac{\omega}{c_2} \quad (1)$$

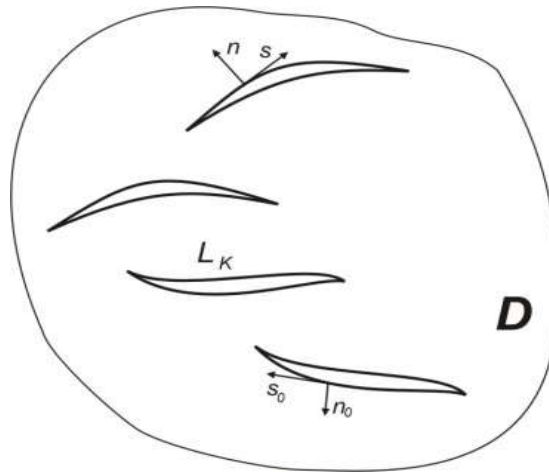


Рис. 1. Схема довільної системи розрізів у середовищі

Тут ω - частота коливань, c_2 - швидкість поширення поперечної хвилі.

В результаті взаємодії падаючої W_0 хвилі з розрізами виникає розсіяна хвиля переміщень W_1 , яка задовольняє рівнянню антиплоскої деформації [6,9]:

$$\Delta W_1 + \gamma_2^2 W_1 = 0, \quad \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \quad (2)$$

Для механіки руйнування вирішальне значення має асимптотичний розподіл напружень в околиці вершин дефектів [12]. Ненульові компоненти тензора напружень σ_{31}, σ_{32} є дотичні напруження у площині поперечного перерізу тріщин. Вони пов'язані з переміщенням $W = W_0 + W_1$ формулами (тут μ - модуль зсуву):

$$\sigma_{31} = \mu \frac{\partial W}{\partial x}, \quad \sigma_{32} = \mu \frac{\partial W}{\partial y}, \quad \sigma_{31} - i\sigma_{32} = 2\mu \frac{\partial W}{\partial z}, \quad z = x + iy. \quad (3)$$

Дотичне напруження σ_n , що діє на L у точці $\zeta = \xi + i\eta \in L$ з боку позитивної нормалі, дорівнює:

$$\sigma_n = \sigma_{31} \sin \varphi - \sigma_{32} \cos \varphi = \operatorname{Im} \left\{ e^{i\varphi} (\sigma_{31} - i\sigma_{32}) \right\}, \quad (4)$$

де φ - кут між позитивною дотичною до L у точці ζ і віссю OX.

При переході через L воно забезпечує стрибок переміщень $([W_1] = W_1^+ - W_1^- = f(s))$ і безперервність напружень $([\sigma_n] = \sigma_n^+ - \sigma_n^- = 0)$. Залишається виконати граничну умову на L :

$$\sigma_{n_0} = \mu \frac{\partial}{\partial n_0} (W_0 - W_1) = 0, \quad (5)$$

яке перепишемо у вигляді:

$$\left(e^{i\varphi_0} \frac{\partial W}{\partial z} - e^{-i\varphi_0} \frac{\partial \overline{W}}{\partial \overline{z}} \right)^{\pm} = 0. \quad (6)$$

Розв'язання антиплоскої динамічної задачі математичної фізики зводиться до визначення функції $W_1(x, y)$ – рішення рівняння Гельмгольца (2) у площині із системою тріщин-розрізів при виконанні умов типу Зоммерфельда випромінювання на нескінченності [8].

Розглядається три варіанти нескінченного середовища: необмежений простір ($A=0$), напівпростір з границею $y=0$, вільною від сил ($A=-1$) та напівпростір із заземленою границею $y=0$ ($A=1$):

$$\begin{aligned} A=-1: \quad \sigma_y|_{y=0} &= \mu \frac{\partial W}{\partial y}|_{y=0} = 0; \\ A=1: \quad W|_{y=0} &= 0. \end{aligned}$$

Як і в [11], запишемо функцію $W_1(x, y)$, що характеризує розсіяну розрізами хвилю переміщень в області D , наступним чином:

$$\begin{aligned} W_1(x, y) = \frac{1}{4} \int_L f(s) & \left(\left[\frac{\partial}{\partial z} H_0^{(1)}(\gamma_2 r) d\zeta - A \frac{\partial}{\partial z} H_0^{(1)}(\gamma_1 r_1) \right] - \right. \\ & \left. - \left[\frac{\partial}{\partial \overline{z}} H_0^{(1)}(\gamma_2 r) d\overline{\zeta} - A \frac{\partial}{\partial \overline{z}} H_0^{(1)}(\gamma_1 r_1) \right] \right) + (\nu_0 - A\nu_0^-), \\ r = |z - \zeta|, \quad r_1 = |z - \overline{\zeta}| \end{aligned} \quad (7)$$

Тут $H_n^{(1)}(x)$ - функція Ханкеля першого роду n -го порядку; $f(s)$ - невідома функція, що задовольняє на L умові Гельдера.

Інтегральне подання (7) автоматично задовольняє рівнянню Гельмгольца (2) в області D та умовам випромінювання на нескінченності.

Припустимо, що стрибки переміщень на кінцях розрізів L_j дорівнюють нулю, тобто:

$$f(a_j) = f(b_j) = 0, \quad j = \overline{1, k}. \quad (8)$$

В інтегральній формі (8) можна записати у вигляді:

$$\int_L df(s) = \int_L f'(s) ds = 0. \quad (9)$$

Скористаємося відомими співвідношеннями [9]:

$$\frac{\partial^2}{\partial z^2} H_0^{(1)}(\gamma r) = \frac{\gamma^2}{4} e^{-2i\alpha} H_2^{(1)}(\gamma r),$$

$$\frac{\partial^2}{\partial \bar{z}^2} H_0^{(1)}(\gamma r) = \frac{\gamma^2}{4} e^{2i\alpha} H_2^{(1)}(\gamma r), \quad (10)$$

$$\frac{\partial^2}{\partial z \partial \bar{z}} H_0^{(1)}(\gamma r) = -\frac{\gamma^2}{4} H_2^{(1)}(\gamma r), \quad z - \zeta = r e^{i\alpha}.$$

Поведінка в нулі ($x \rightarrow 0$) функцій Ханкеля нульового та другого порядків характеризується асимптотичними формулами:

$$H_0^{(1)}(x) = \frac{2i}{\pi} \ln x + H_0(x), \quad H_2^{(1)}(x) = \frac{4}{i\pi x^2} + H_2(x), \quad (11)$$

де $H_0(x)$ и $H_2(x)$ - неперервні в точці $x = 0$.

Підстановка інтегрального подання (7) у вираз, що стоїть у лівій частині граничної умови (6), з урахуванням (10), (11) дає:

$$\begin{aligned} e^{i\varphi} \frac{\partial v_*}{\partial z} - e^{-i\varphi} \frac{\partial v_*}{\partial \bar{z}} &= \frac{1}{4} \int_L f(s) \left\{ \left[\left(e^{i(\varphi_0+\varphi)} \frac{\partial^2 H^{(1)}(\gamma r)}{\partial z^2} - e^{i(\varphi_0-\varphi)} \frac{\partial^2 H^{(1)}(\gamma r)}{\partial z \partial \bar{z}} \right) - \right. \right. \\ &\quad \left. \left. - \left(e^{-i(\varphi_0-\varphi)} \frac{\partial^2 H^{(1)}(\gamma r)}{\partial z \partial \bar{z}} - e^{-i(\varphi_0+\varphi)} \frac{\partial^2 H^{(1)}(\gamma r)}{\partial \bar{z}^2} \right) \right] - \right. \\ &\quad \left. - A \left[\left(e^{i(\varphi_0+\varphi)} \frac{\partial^2 H^{(1)}(\gamma r_1)}{\partial z \partial \bar{z}} - e^{i(\varphi_0-\varphi)} \frac{\partial^2 H^{(1)}(\gamma r_1)}{\partial z^2} \right) - \right. \right. \\ &\quad \left. \left. - \left(e^{-i(\varphi_0-\varphi)} \frac{\partial^2 H^{(1)}(\gamma r_1)}{\partial \bar{z}^2} - e^{-i(\varphi_0+\varphi)} \frac{\partial^2 H^{(1)}(\gamma r_1)}{\partial z \partial \bar{z}} \right) \right] \right\} ds = \\ &= \frac{\gamma^2}{16} \int_L f(s) \left\{ \left[\left(e^{i(\varphi_0+\varphi-2\alpha)} + e^{i(\varphi_0-\varphi-2\alpha)} \right) H_2^{(1)}(\gamma r) + \right. \right. \\ &\quad \left. \left. + \left(e^{i(\varphi_0-\varphi)} + e^{-i(\varphi_0-\varphi)} \right) H_0^{(1)}(\gamma r) \right] + A \left[\left(e^{i(\varphi_0-\varphi-2\alpha_1)} + e^{-i(\varphi_0-\varphi-2\alpha_1)} \right) H_2^{(1)}(\gamma r_1) + \right. \right. \\ &\quad \left. \left. + \left(e^{i(\varphi_0+\varphi)} + e^{-i(\varphi_0+\varphi)} \right) H_0^{(1)}(\gamma r_1) \right] \right\} ds = \end{aligned} \quad (12)$$

Тут використано формули інтегрування частинами для гіперсингулярних інтегралів [2] за додаткових умов (9).

Залучення формул Сохоцького-Племеля [6] для обчислення граничних значень інтегралів типу Коші, що виникають при задоволенні граничної умови (6) з урахуванням співвідношень (1), (12), зводить розглянуту крайову задачу до сингулярного інтегро-диференціального рівняння $f(s)$ [11]:

$$\begin{aligned} &\frac{1}{2i\pi} \int_L \operatorname{Re} \left\{ \frac{e^{i\varphi_0}}{\xi - \xi_0} \right\} df - \frac{\gamma^2}{4i\pi} \int_L f(s) \cos(\varphi_0 - \varphi) \ln r_0 ds + \\ &+ \frac{\gamma^2}{8} \int_L \left\{ f(s) \left[\cos(\varphi_0 + \varphi - 2\alpha_0) \left(H_2^{(1)}(\gamma r) - \frac{4}{i\pi \gamma^2 r_0^2} \right) + \cos(\varphi_0 - \varphi) \left(H_0^{(1)}(\gamma r_0) - \frac{2i}{\pi} \ln r_0 \right) \right] + \right. \\ &\quad \left. + A \left[\cos(\varphi_0 - \varphi - 2\alpha_{10}) H_2^{(1)}(\gamma r_{10}) + \cos(\varphi_0 + \varphi) H_0^{(1)}(\gamma r_{10}) \right] \right\} ds = \\ &= \tau \gamma_2 \sin(\varphi_0 - \psi) e^{-i\gamma_2 (\xi_0 \cos \psi + \eta_0 \sin \psi)} \end{aligned} \quad (13)$$

Для однозначної розв'язності інтегро-диференціального рівняння (13) до нього слід додати додаткову умову (9).

Представимо контур L у параметричній формі. Для цього на кожному контурі $L_j (j = \overline{1, K})$ вибираємо локальну систему координат з урахуванням паралельного перенесення та повороту осей координат. З урахуванням цього на контурі L_j будемо мати:

$$\begin{aligned} \frac{\zeta_0'}{\zeta - \zeta_0} &= \frac{1}{\beta - \beta_0} + \left(\frac{\zeta_0'}{\zeta - \zeta_0} - \frac{1}{\beta - \beta_0} \right), \\ s' \ln |\zeta - \zeta_0| &= s'_0 \ln |\beta - \beta_0| + (s' \ln |\zeta - \zeta_0| - s'_0 \ln |\beta - \beta_0|), \\ \zeta &= \zeta(\beta), \quad \zeta_0 = \zeta(\beta_0), \quad -1 \leq \beta, \beta_0 \leq 1, \quad \zeta(-1) = a_j, \quad \zeta(+1) = b_j. \end{aligned} \quad (14)$$

Помножимо рівняння (13) на $s'(\beta_0)$ і виділимо в ньому ядро типу Коши та логарифмічне ядро наступним чином:

$$\begin{aligned} \frac{\zeta_0'}{\zeta - \zeta_0} &= \frac{1}{\beta - \beta_0} + \left(\frac{\zeta_0'}{\zeta - \zeta_0} - \frac{1}{\beta - \beta_0} \right), \\ s' \ln |\zeta - \zeta_0| &= s'_0 \ln |\beta - \beta_0| + (s' \ln |\zeta - \zeta_0| - s'_0 \ln |\beta - \beta_0|). \end{aligned} \quad (15)$$

Тут, використовуючи правило Лопіталя, знаходимо:

$$\begin{aligned} \lim_{\beta \rightarrow \beta_0} \left(\frac{\zeta_0'}{\zeta - \zeta_0} - \frac{1}{\beta - \beta_0} \right) &= -\frac{\zeta''(\beta_0)}{2\zeta'(\beta_0)}, \\ \lim_{\beta \rightarrow \beta_0} (s' \ln |\zeta - \zeta_0| - s'_0 \ln |\beta - \beta_0|) &= s'_0 \ln |\zeta'_0|. \end{aligned} \quad (16)$$

Тепер інтеграл з логарифмічним ядром інтегруємо частинами з урахуванням додаткової умови (9):

$$\int_{-1}^1 f(\beta) \ln |\beta - \beta_0| d\beta = - \int_{-1}^1 (\beta - \beta_0) (\ln |\beta - \beta_0| - 1) df(\beta). \quad (17)$$

Параметрична форма інтегрального рівняння набуває вигляду:

$$\begin{aligned} &\frac{1}{2i\pi} \int_{-1}^1 \frac{f'(\beta)}{\beta - \beta_0} d\beta + \frac{\gamma^2}{4i\pi} \operatorname{Re}(\zeta'_0 \zeta') \int_{-1}^1 (\beta - \beta_0) (\ln |\beta - \beta_0| - 1) f'(\beta) d\beta + \\ &+ \frac{\gamma^2}{8} \int_{-1}^1 f(\beta) \left\{ \left[\operatorname{Re} \left(\zeta'_0 \zeta' \frac{\overline{\zeta_0 - \zeta}}{\zeta_0 - \zeta} \right) H_2^{(1)}(\gamma r_0) - \frac{4}{i\pi \gamma^2 (\beta_0 - \beta)^2} \right] + \right. \\ &\quad \left. + \left[\operatorname{Re}(\zeta'_0 \overline{\zeta'}) H_0^{(1)}(\gamma r_0) - \frac{2i}{\pi} \operatorname{Re}(\zeta'_0 \overline{\zeta'}) \ln |\beta - \beta_0| \right] + \right. \\ &\quad \left. + A \left[\operatorname{Re} \left(\zeta'_0 \overline{\zeta'} \frac{\overline{\zeta_0 - \zeta}}{\zeta_0 - \zeta} \right) H_2^{(1)}(\gamma r_{10}) - \operatorname{Re}(\zeta'_0 \overline{\zeta'}) H_0^{(1)}(\gamma r_{10}) \right] \right\} d\beta = \\ &= \tau \gamma \zeta'_0 (e^{-i\gamma \eta_0} + A e^{i\gamma \eta_0}) \end{aligned} \quad (18)$$

Запропонована процедура регуляризації інтеграла з логарифмічним ядром дозволяє звести розв'язання задачі до сингулярного інтегрального рівняння (СІУ) щодо функції $f'(\beta)$. Ядра інтегралів, що відповідають функції $f(\beta)$, неперервні. Єдине рішення отриманого сингулярного

інтегро-диференціального рівняння (18) з урахуванням додаткової умови (9) будемо отримувати в класі функцій, що мають кореневу особливість на кінцях розрізів [10]. Таким чином, маємо:

$$f'(\beta) = \frac{\Omega(\beta)}{\sqrt{1-\beta^2}}, \quad f(\beta) = \int_{-1}^{\beta} f'(\beta) d\beta, \quad \beta = \cos \theta.$$

Представимо невідому функцію $\Omega(\beta)$ інтегрального рівняння (18) як сукупність функцій, $\Omega_j(\beta^j)$ визначених на контурах $L_j, j = \overline{1, K}$. Чисельна реалізація інтегрального рівняння (18) проводиться методом механічних квадратур [8]. Рівняння, що відповідає контуру L_j , задовольняється у вузлах Чебишева другого роду $\theta_m = \frac{\pi m}{n_j} \quad (m = \overline{1, n_j - 1})$ і зводиться до системи лінійних алгебраїчних рівнянь (СЛАР) [13] щодо значень функції $\Omega_j(\beta)$ у вузлах Чебишева першого роду $\theta_k = \frac{2k-1}{n_j} \pi \quad (k = \overline{1, n_j})$, де n_j - число точок дискретизації контуру L_j .

Для інтеграла типу Коші використовуємо квадратурну формулу:

$$\int_{-1}^1 \frac{\Omega_j(\beta)}{\sqrt{1-\beta^2}(\beta-\beta_m)} d\beta \approx \frac{\pi}{n_j} \sum_{k=1}^{n_j} \frac{\Omega_j(\beta_k)}{\beta_k - \beta_m}, \quad \beta_m = \cos \theta_m, \quad \beta_k = \cos \theta_k. \quad (19)$$

До інтегралу, що містить регулярне ядро $D(\beta_0, \beta)$ і має кореневу особливість, застосовуємо квадратурну формулу Гауса:

$$\int_{-1}^1 \frac{\Omega_j(\beta)}{\sqrt{1-\beta^2}} D(\beta_m, \beta) d\beta \approx \frac{\pi}{n_j} \sum_{k=1}^{n_j} D(\beta_m, \beta_k) \Omega_j(\beta_k). \quad (20)$$

Щодо додаткової умови (9) маємо:

$$\sum_{k=1}^{n_j} \Omega_j(\beta_k) = 0. \quad (21)$$

Інтерполяційний багаточлен Лагранжа для функції $f_j(\beta)$ має вигляд [8]:

$$f_j(\beta) = \int_{-1}^{\beta} f'(\beta) d\beta \approx -\frac{2}{n_j} \sum_{k=1}^{n_j} \Omega_j(\beta_k) \sum_{l=1}^{n_j-1} \frac{\cos l \theta_k \sin l \theta}{l}. \quad (22)$$

СЛАР щодо функцій $\Omega_j, j = \overline{1, K}$ набуває вигляду:

$$\sum_{k=1}^{n_j} A_{mk}^j \Omega_j(\beta_k) = N_m^j, \quad \sum_{k=1}^{n_j} \Omega_j(\beta_k) = 0, \quad (23)$$

$$A_{mk}^j = \frac{\pi}{n_j} \left\{ \frac{1}{2i\pi} \frac{1}{\beta_k - \beta_m} + \frac{\gamma_2^2}{4i\pi} \operatorname{Re}(\zeta'_m \overline{\zeta'_m}) (\beta_k - \beta_m) (\ln|\beta_k - \beta_m| - 1) + \right. \\ \left. + \frac{\gamma_2^2}{8} \sum_{v=1}^{n_l} B_{mv} \sin \theta_v \left(-\frac{2}{n_j} \sum_{j=1}^{n_j-1} \frac{\cos l \theta_k \sin l \theta_v}{l} \right) \right\},$$

$$B_{mv} = \operatorname{Re} \left(\zeta'_m \zeta'_v \frac{\overline{\zeta_m} - \overline{\zeta_v}}{\zeta_m - \zeta_v} \right) H_2^{(1)}(\gamma |\zeta_m - \zeta_v|) - \frac{4}{i\pi \gamma^2 (\beta_v - \beta_m)^2} + \\ + \operatorname{Re} \left(\zeta'_m \overline{\zeta'_v} \right) H_0^{(1)}(\gamma |\zeta_m - \zeta_v|) - \frac{2i}{\pi} \operatorname{Re}(\zeta'_m \overline{\zeta'_m}) \ln|\beta_v - \beta_m| + \\ + A \left[\operatorname{Re} \left(\zeta'_m \overline{\zeta'_v} \frac{\overline{\zeta_m} - \overline{\zeta_v}}{\zeta_m - \zeta_v} \right) H_2^{(1)}(\gamma |\zeta_m - \overline{\zeta_v}|) + \operatorname{Re}(\zeta'_m \zeta'_v) H_0^{(1)}(\gamma |\zeta_m - \overline{\zeta_v}|) \right]$$

Таким чином, за чисельної реалізації СИУ (18), (9) задача зводиться до рішення СЛАР (23) з $N=n_1+n_2+\dots+n_k$ невідомими.

Проведено чисельне дослідження описаної задачі. З метою дослідження збіжності побудованого алгоритму розглянуто випадок нормального падіння хвилі зсуву на систему [7], що складається з параболічних тріщин, які по черзі розташовані в пружному напівпросторі на однаковій відстані один від одного і симетрично орієнтовані вздовж осі X (рис. 2). Однак число тріщин праворуч і ліворуч не є обов'язково рівним ($L_{-T} \neq L_K$)

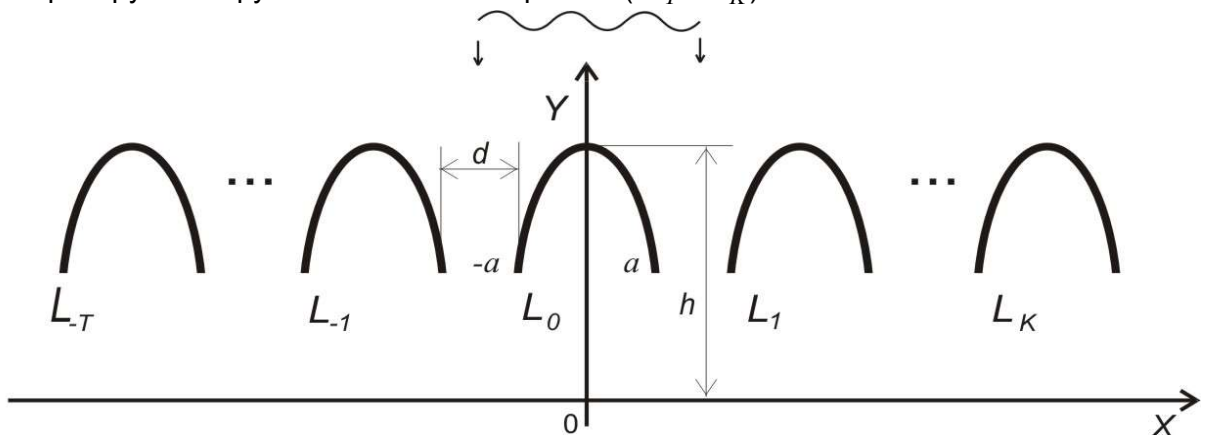


Рис. 2 Схема навантаження системи параболічних розрізів для обчислень

Елементи матриці СЛАР, до якої, зрештою, зводяться СИУ, є результатом дискретизації контурів. Очевидно, що розмір матриці пропорційний числу тріщин. Застосуємо розпаралелювання алгоритму, у якому кожен елемент матриці визначається координатами вузлів дискретизації. Як показано в [7], даний метод у обчислювальному сенсі зводиться до обходу кожного контуру по точках колокації позаінтегральної змінної ζ_{k0} і одночасному обходу кожного контуру за аналогічним чи іншим вузлом змінної інтегрування ζ_k .

Паралельно-конвеєрна схема обчислень побудована аналогічно [9]. Тут також обчислення мають такі етапи: синтез масивів вихідних даних, синтез матриці СЛАР, рішення СЛАР методом Гауса, синтез масивів підсумкових рішень. Перший, другий та четвертий етапи макроконвеєра не вимагають пересилання даних, що означає незалежність обчислень. На третьому етапі для реалізації СЛАР існує оптимальна кількість процесів, що визначається специфікою матриці. Це означає, що для 1, 2 та 4 етапів алгоритму оптимальним є число процесів, що відповідає числу коефіцієнтів СЛАР.

У цій методиці при розв'язанні крайової задачі основною операцією є визначення поточної відстані між точками колокації та інтегрування, заданої на безлічі значень параметричних координат неоднорідностей. Зазначена відстань є аргументом функції Гріна. І оскільки комбінації самих функцій Гріна та коефіцієнтів при них є елементами матриці СЛАР, зазначена процедура може бути базовою при розробці програми. Як показано в [7, 9], алгоритм добре масштабується по обчислювальних вузлах.

Обчислювальний процес рішення СЛАР розпаралелюється згідно [7,15]. Паралельне обчислення підсумкових шуканих характеристик здійснюється шляхом встановлення масивів значень невідомих функцій $f_k(\beta_p)$ в інтегральні подання рішень аналогічно процедурам формування матриці СЛАР. Для реалізації СЛАР ефективніше використовувати постстрокове розпаралелювання, коли пересилання та обчислення перебувають у балансі.

Уявимо контур L у параметричній формі:

$$\zeta = \zeta(\beta), \quad \zeta_0 = \zeta(\beta_0), \quad -1 \leq \beta, \quad \beta_0 \leq 1, \quad \zeta(-1) = a_j, \quad \zeta(+1) = b_j. \quad (24)$$

Параметричне рівняння центральної параболи ставилося у вигляді:

$$x = p_1 \beta, \quad y = p_2 \beta^2 \quad (25)$$

Отже, на рис. 2 наведена конфігурація, де коефіцієнт кривизни параболи $p_2 = -0,5$

У роботі визначалися залежності величини коефіцієнту інтенсивності напружень (максимального напруження на продовженні L за вершину) від параметра q [11,13]:

$$q = 2l\gamma/\pi \quad (26)$$

$$K_3 = \lim_{\rho_0 \rightarrow 0} \sqrt{2\pi\rho_0} \sigma_n^c = \frac{\mu}{2} \sqrt{\frac{\pi}{s'(\mp 1)}} |f_0(\mp 1)|. \quad (27)$$

Отже, проводились розрахунки безрозмірної величини:

$$\delta = \frac{\mu}{\sigma_{32}^{\max}} \frac{|\Omega(\mp 1)|}{2l}, \quad (28)$$

де $\sigma_{32}^{\max} = \mu\tau$ максимальне напруження в хвилі, що падає. Тут τ - амплітуда навантаження, γ - хвильове число, l - напівдовжина контуру, тобто, $L = 2l$.

Для перевірки достовірності алгоритму проведено два тести: віддалення трьох неоднорідностей на відстань 10^6 їх довжини один від одного та насичення системи [7,9] зі збільшенням числа геометрично рівних відбивачів у системі до 13-15. При віддаленні характеристики хвильового поля від кожного відбивача набуває значення відповідного поля від одного. Тому проводилося порівняння отриманих результатів із результатами, наведеними в [3,13]. Співпадіння результатів показало хорошу достовірність алгоритму.

У разі збільшення числа дефектів, розташованих паралельно фронту хвилі, що набігає, для більшості відбивачів характеристики співпадають із характеристиками результату розв'язання періодичної задачі. Відхилення значень виявляються лише для крайніх відбивачів.

Точність обчислень перевірялася шляхом порівняння результатів при різних значеннях N . Стабілізація значень до 10^{-10} досягається при 50 точках колокації кожного контуру. Збіжність алгоритму так само, як і в [7], не залежить від числа тріщин. Обумовленість матриць перевірялася згідно алгоритму, описаному в [15].

Рис. 3 ілюструє результати першого тесту. На рис. 3а наведено залежність коефіцієнта інтенсивності (28) - величини δ від $\mathcal{N}^2/4$ для будь-якої з трьох параболічних тріщин, відстань

між якими 10^6 їх довжини. Тут $p_1 = 1$, $p_2 = \frac{\pi}{\gamma}$. Як і в роботі [13], криві 1, 2, 3 послідовно відповідають випадкам напівпросторів: 1 - із защемленою ($A=1$), 2 - вільною ($A=-1$) границею, а також 3 - нескінченного простору ($A=0$). Результати повністю співпадають. На рис. 3б наведено залежність, що співпадає з аналогічним результатом роботи [3] для випадку прямолінійної тріщини при $A=0$.

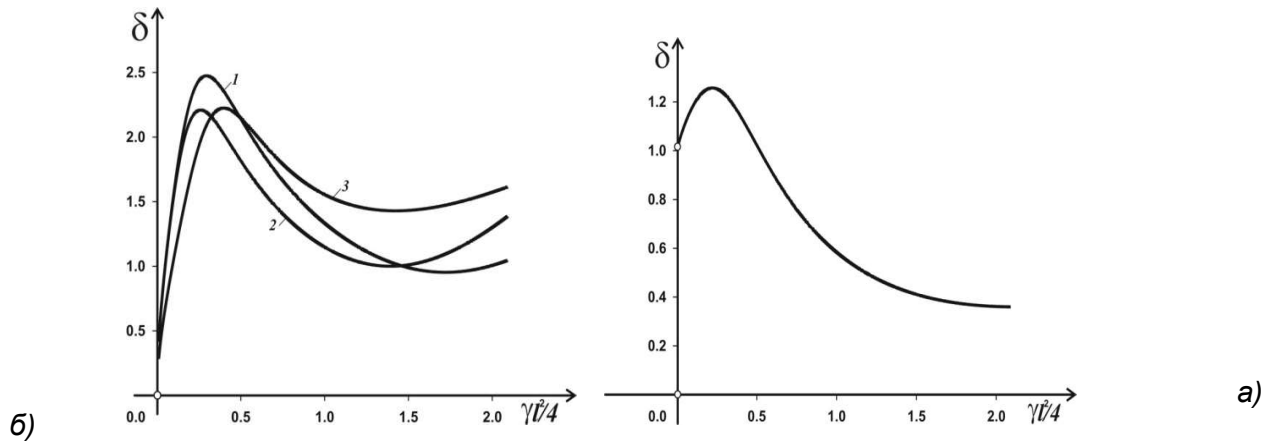
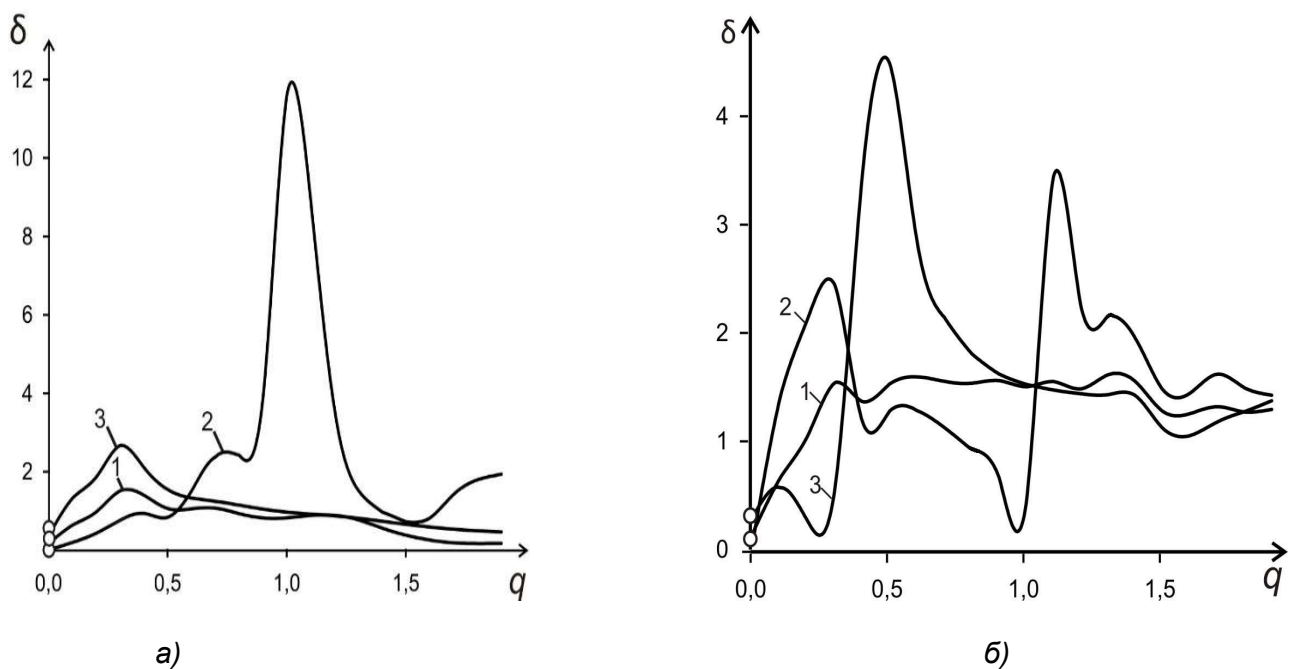


Рис. 3. Результати тестових обчислень

У цій роботі всі залежності наведено для центрального об'єкта. На рис. 4а наведено залежність коефіцієнта інтенсивності від параметра q згідно з формулою (26) [11,13]. Криві 1,2,3 послідовно відповідають випадкам: крива 1 - $A=0$, 2 - $A=-1$, 3 - $A=1$, коефіцієнт кривизни параболи $p_2 = 0,5$. У розрахунках використовувалося 9 об'єктів, параметри (рис. 2) $h=0,05$, $d=1$. На рис. 4б криві 1,2,3 також послідовно відповідають випадкам $A=0$, $A=-1$, $A=1$, $p_2 = 0,5$; 5 об'єктів, $h=2$, $d=1$.

Рис. 4. Розподіл коефіцієнта інтенсивності від параметра q

На рис. 5а криві 1,2,3 послідовно відповідають випадкам $A=0$, $A=-1$, $A=1$, $p_2 = 0,5$, 5 об'єктів, $h=6$, $d=1$. На рис. 5б для кривих 1,2,3 відповідно $A=0$, $A=-1$, $A=1$, $p_2 = 0,5$; 5 об'єктів, $h=10$, $d=1$.

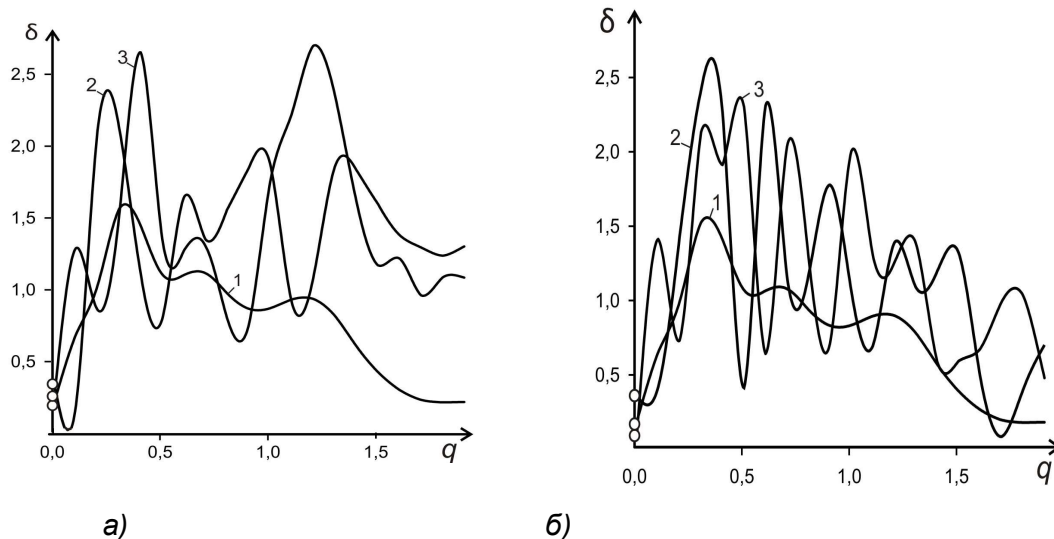


Рис.5. Розподіл коефіцієнта інтенсивності від параметра q

На рис. 5а криві 1,2,3 послідовно відповідають випадкам $A=0$, $A=-1$, $A=1$, $p_2 = 0,5$, 5 об'єктів, $h=6$, $d=1$. На рис. 5б для кривих 1,2,3 відповідно $A=0$, $A=-1$, $A=1$, $p_2 = 0,5$; 5 об'єктів, $h=10$, $d=1$.

Висновки

У задачі дифракції SH-хвилі на системі тріщин в ізотропному напівпросторі паралельні алгоритми дозволяють значно скоротити час обчислень та детальніше проаналізувати характеристики хвильового поля. Поєднання методу CIY, який знижує на одиницю розмірність задачі, та значна економія часу обчислень за рахунок розпаралелювання обчислювальних процедур призводить до суттєвого збільшення ефективності запропонованого алгоритму.

Література

1. Агапов В.П. Метод кінцевих елементів у статичі, динаміці та стійкості просторових тонкостінних підкріплених конструкцій / М: Вид. АСВ (Асоціації будівельних ВНЗ), 2000 - 152 с.
2. Білоцерківський С.М., Ліфанов І.К. Численні методи в сингулярних інтегральних рівняннях. - М: Наука, 1985. - 256 с.
3. Бородачов Н.М., Динамічна задача про тріщину у разі деформації поздовжнього зсуву// Проблеми міцності, № 4, 1973, с. 23-25
4. Вертгейм І.І., Терпугов В.М. Паралельні технології обчислень у механіці суцільних середовищ та МДТТ: Навчальний посібник / Перм: ПГУ, 2007. - 84 с.
5. Гузь О.М., Кубенко В.Д., Черевко М.О. Дифракція пружних хвиль/К: Наук. думка, 1978. - 307 с.
6. Мухелішвілі Н.І. Сингулярні інтегральні рівняння/М. Наука, 1968. - 512 с.
7. Назаренко О.М., Панченко Б.Є. Схема паралельних обчислень у задачах дифракції хвиль зсуву на системі отворів у нескінченному пружному середовищі // Проблеми програмування. - 2010. - № 2-3. - С. 604-610.
8. Панасюк В.В., Саврук М.П., Назарчук З.Т. Метод сингулярних інтегральних рівнянь у двовимірних задачах дифракції / К: Наук. думка, 1984. - 344 с.
9. Панченко Б.Є., Назаренко О.М. Каркасний аналіз предметної галузі: стаціонарні динамічні завдання теорії пружності для ізотропних середовищ із довільними неоднорідностями // Кібернетика та системний аналіз, 2013, - № 1, С. 172-187.
10. Селезов І.Т., Кривоніс Ю.Г., Яковлев В.В. Розсіяння хвиль локальними неоднорідностями в суцільних середовищах / К: Наук. думка, 1985. - 136 с.
11. Фільштинський Л.А. Динамічна завдання теорії пружності області з криволінійними розрізами (деформація поздовжнього зсуву) // Докл. АН СРСР, 1977-№6 - С. 1327-1330.

12. Фільштинський Л.А. Дифракція пружних хвиль на тріщинах, отворах, включення в ізотропному середовищі // Изв. АН СРСР. Механіка твердого тіла, 1991. - №4. - С. 119 -127.
13. Фільштинський Л.А., Взаємодія хвиль напруг з криволінійними тунельними розрізами поздовжнього зсуву в напівпросторі // Прикладна математика та механіка, Т. 46 № 3, 1982, с. 482-487
14. Фільштинський Л.А., Назаренко А.М., Взаємодія хвиль напруги з періодичною системою тріщин поздовжнього зсуву в напівпросторі// Акустичний журнал, Т. 29, вип. 3, 1983, с. 382-386
15. Хімич О.М., Поляню В.В. Ефективність двовимірних блочно-циклічних паралельних алгоритмів // Проблеми програмування. - 2008. - №3. - С. 145-149.

References

1. Agapov V.P. Metod kincevih elementiv u statiki, dinamici ta stijkosti prostоровih tonkostinnih pidkriplenih konstrukcij / M: Vid. ASV (Asociaciyi budivelnih VNZ), 2000 - 152 s.
2. Bilocerktivskij S.M., Lifanov I.K. Chislenni metodi v singulyarnih integralnih rivnyannayah. - M: Nauka, 1985. - 256 s.
3. Borodachov N.M., Dinamichna zadacha pro trishinu u razi deformaciyi pozdovzhnogo zsuvu// Problemi micnosti, № 4, 1973, s. 23-25
4. Vertgejm I.I., Terpugov V.M. Paralelni tehnologiyi obchislen u mehanici sucilnih seredovish ta MDTT: Navchalnij posibnik / Perm: PGU, 2007. - 84 s.
5. Guz O.M., Kubenko V.D., Cherevko M.O. Difrakciya pruzhnih hvil/K: Nauk. dumka, 1978. - 307 s.
6. Mushelishvili N.I. Singulyarni integralni rivnyannya/M. Nauka, 1968. - 512 s.
7. Nazarenko O.M., Panchenko B.Ye. Shema paralelnih obchislen u zadachah difrakciyi hvil zsuvu na sistemi otvoriv u neskinchenomu pruzhnomu seredovishi // Problemi programuvannya. - 2010. - № 2-3. - S. 604-610.
8. Panasyuk V.V., Savruk M.P., Nazarchuk Z.T. Metod singulyarnih integralnih rivnyan u dvovimirnih zadachah difrakciyi/K: Nauk. dumka, 1984. - 344 s.
9. Panchenko B.Ye., Nazarenko O.M. Karkasnij analiz predmetnoyi galuzi: stacionarni dinamichni zavdannya teorii pruzhnosti dlya izotropnih seredovish iz dovilnimi neodnorodnostyami // Kibernetika ta sistemnij analiz, 2013, - № 1, С. 172-187.
10. Selezov I.T., Krivonis Yu.G., Yakovlyev V.V. Rozsiyannya hvil lokalnimi neodnorodnostyami v sucilnih seredovishah/K: Nauk. dumka, 1985. - 136 s.
11. Filshtinskij L.A. Dinamichna zavdannya teorii pruzhnosti oblasti z krivolinijnimi rozrizami (deformaciya pozdovzhnogo zsuvu) // Dokl. AN SRSR, 1977-№6 - С. 1327-1330.
12. Filshtinskij L.A. Difrakciya pruzhnih hvil na trishinah, otvorah, vklyuchennya v izotropnomu seredovishi // Izv. AN SRSR. Mехanika tverdogo tila, 1991. - №4. - S. 119 -127.
13. Filshtinskij L.A., Vzayemodiya hvil naprug z krivolinijnimi tunelnimi rozrizami pozdovzhnogo zsuvu v napivprostori // Prikladna matematika ta mehanika, Т. 46 № 3, 1982, s. 482-487
14. Filshtinskij L.A., Nazarenko A.M., Vzayemodiya hvil naprug z periodichnoyu sistemoyu trishin pozdovzhnogo zsuvu v napivprostori// Akustichnij zhurnal, Т. 29, vip. 3, 1983, s. 382-386
15. Himich O.M., Polyanko V.V. Efektivnist dvovimirnih blochno-ciklichnih paralelnih algoritmiv // Problemi programuvannya. - 2008. - №3

УДК 621.391.8

КОМПЛЕКСНИЙ АНАЛІЗ ВЛАСТИВОСТЕЙ КОДУ ХЕММІНГА В СУПЕРПОЗИЦІЇ ТЕОРІЙ КОДУВАННЯ, ЗАВАДОСТІЙКОСТІ ТА ІНФОРМАЦІЇ

DOI 10.36994 / 2788-5518-2023-01-05-10

Уривський Л.О., д.т.н., проф., Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна. leonid_uic@ukr.net

Косогор А.В. Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна. kosogoranastasia@gmail.com

Анотація: В статті досліджуються властивості кодів Хеммінга, які активно використовуються на практиці вже більш ніж 70 років, в контексті визначення їх найбільш ефективної області використання в різноманітних сферах. Специфічний інтерес до кодів Хеммінга викликаний також й тим, що вони мають унікальну властивість: вони – за класичним визначенням – є досконалими кодами. Складність дослідження полягає в тому, що поняття «найбільш ефективна область використання» в кожній із згаданих теорій є особливим, тому комплексний аналіз властивостей коду Хеммінга в суперпозиції теорій кодування, завадостійкості та інформації є актуальною і цікавою задачею дослідження.

Ключові слова. Коди Хеммінга, виправна здатність, надлишковість, завадостійкість, інформаційна ефективність.

COMPREHENSIVE ANALYSIS OF THE HAMMING CODE PROPERTIES BASED ON SUPERPOSITION OF THE CODING, NOISE IMMUNITY AND INFORMATION THEORIES

Leonid Uryvsky, Dr. habil, Prof. National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine. leonid_uic@ukr.net

Anastasia Kosogor. National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine. kosogoranastasia@gmail.com

Abstract: In the modern era, the task of transferring and preserving information is extremely important. Errors can often occur when data is transmitted over channels with interference or when it is stored on an unreliable medium. Scientific studies of the properties of electrical signals, which are information carriers in electronic communication systems, traditionally have three directions: the theory of immunity, the theory of information, and the theory of probability. Each of these theories has its own historical lineage, outstanding achievements and untapped potential. But the peculiarity of these studies is that more often they are implemented independently of each other, using their own set of initial parameters, criteria and performance indicators, quantitative characteristics. The article examines the properties of Hamming codes, which have been actively used in practice for more than 70 years, in the context of determining their most effective area of use in various fields. The specific interest in Hamming codes is also caused by the fact that they have a unique property: they are - according to the classical definition - perfect codes.

The difficulty lies in the fact that the concept of "the most effective area of use" in each of the mentioned theories is special, therefore, a comprehensive analysis of the properties of the Hamming code in the superposition of the theories of coding, immunity and information is a relevant and interesting research task.

From the standpoint of coding theory, Hamming codes have the same correcting ability - they correct only one error in a block. After all, as the length of the block n increases, the redundancy decreases. At the same time, the degree of closeness of the Hamming codes to the Plotkin limit increases. Therefore, preference according to these criteria should be given to Hamming codes with the largest block length n .

From the standpoint of the theory of immunity, the best is the code that provides the best correction indicators at the smallest value of the data transmission channel parameter h^2 (signal/noise ratio) under the conditions of ensuring the required reliability. Compared to the case of no coding, the shortest code – the Hamming code (7,4) – gives the greatest gain.

From the point of view of information theory, the transmission of signals using Hamming codes requires speed correction (acceleration) in the communication channel in order to prevent the loss of source symbols due to the redundancy of the correction code. At the same time, the advantage of the Hamming code (7,4) over other, longer codes remains.

According to the integral indicator - the energy costs to achieve the required reliability without losing the

speed of information transmission from the source - preference should be given to the Hamming code with parameters (7, 4) with a gain in the signal-to-noise ratio $\Delta h_2 = 2.72$ dB compared to the case without coding.

Keywords. *Hamming codes, error-correction ability, redundancy, noise immunity, informational efficiency*

Вступ

У сучасну епоху завдання передачі та збереження інформації є надзвичайно важливими. Часто можуть виникати помилки, коли дані передаються по каналах із перешкодами або коли вони зберігаються на ненадійному носії.

Наукові дослідження властивостей електричних сигналів, які є носіями інформації в системах електронних комунікацій, традиційно мають три напрямки: теорія завадостійкості, теорія інформації, та теорія ймовірності.

Кожна з цих теорій має свою історичну лінію, видатні досягнення і невичерпний потенціал. Але особливістю цих досліджень є те, що частіше вони реалізуються незалежно одна від одної, задіюючи свій набір вихідних параметрів, критеріїв та показників ефективності, кількісні характеристики.

Коди Хеммінга є одними з найпоширеніших та ефективних методів виявлення та виправлення помилок в передачі даних. Розроблені Річардом Хеммінгом в 1950 році, ці коди забезпечують надійність передачі інформації, дозволяючи виявити та виправити певну кількість помилок.

В статті досліджуються властивості кодів Хеммінга, які активно використовуються на практиці вже більш ніж 70 років [1], в контексті визначення їх найбільш ефективної області використання в різноманітних сферах:

- у деяких прикладних програмах у сфері зберігання даних;
- при побудові дискових масивів RAID 2;
- у пам'яті типу ECC (англ. error-correcting code memory, пам'ять з корекцією помилок), яка автоматично розпізнає і виправляє зміни (помилки) бітів пам'яті, що спонтанно виникли.

Специфічний інтерес до кодів Хеммінга викликаний також й тим, що вони мають унікальну властивість: вони є – за класичним визначенням – *досконалими кодами*.

Складність полягає в тому, що поняття «найбільш ефективна область використання» в кожній із згаданих теорій є особливим, тому комплексний аналіз властивостей коду Хеммінга в суперпозиції теорій кодування, завадостійкості та інформації є актуальною і цікавою задачею дослідження.

Виконання досліджень

Першим етапом дослідження кодів Хеммінга є визначення тих критеріїв ефективності їх застосування, які традиційно використовуються у теорії кодування, та відповідна оцінка різновидів коду Хеммінга.

У теорії кодування ефективність кодів визначається кількістю помилок, які той може виправити, кількістю необхідної надлишкової інформації, а також складністю реалізації кодування та декодування.

Коди Хеммінга належать до класу надлишкових лінійних блокових кодів, тобто таких, які мають довжину блоку n біт, з яких k біт є інформаційними, а решта s є надлишковими і використовуються для реалізації виправної здатності коду: $n = k + s$.

Коди Хеммінга мають наступну закономірність у формуванні параметрів (n, k) -блоку:

- $n = 2^p - 1, p \geq 3$;
 - $s = p$;
 - $k = n - s$.
- (1)

Отже, мінімальну довжину n має код Хеммінга $(7, 4)$. Прикладами (n, k) -блоків Хеммінга є коди з параметрами $(15, 11)$, $(31, 26)$, $(63, 57)$ і так далі.

Незалежно від довжини коду n виправні можливості кодів (кількість t помилок, які може виправити блоковий код із довжиною блоку n) визначає так звана кодова відстань Хеммінга d , яка визначається за виразом:

$$d \geq 2t + 1. \quad (2)$$

Як правило, із збільшенням довжини блоку n кодова відстань Хеммінга d (і виправна здатність коду) зростає. Але коди Хеммінга мають помірні можливості щодо виправлених помилок, а саме: для будь яких значень довжин блоку n в кодах Хеммінга кодова відстань $d = 3$, відповідно завжди $t = 1$.

Отже, за критерієм *виправної здатності* всі коди Хеммінга *однакові*.

Але кодам Хеммінга належить особлива властивість: вони належать до так званих досконалих кодів, тобто кодів із максимальною щільністю запакування – з мінімальною кількістю надлишкових розрядів.

Нехай є двійковий блоковий код (n, k) з коригувальною здатністю t . Тоді справедлива нерівність (звана межею Хеммінга) [2]:

$$\sum_{i=0}^t C_n^i \leq 2^{n-k}, \quad (3)$$

де C_n^i – поєднання, тобто кількість наборів із i елементів, відібраних із n -елементної множини.

В теорії кодування межа Хеммінга відома як межа сферичного упакування. Коди, що досягають межі Хеммінга, називають досконалими або щільно запакованими.

Коди Хеммінга є *досконалими* (хоча часто практично застосовувані коди з великою коригувальною здатністю не є досконалими).

У відповідності кодів Хеммінга умові (3) із знаком рівності, тобто умові досконалості, можна упевнитися, користуючись системою рівнянь (1) і співвідношенням (3), а саме:

$$\sum_{i=0}^t C_n^i = C_n^0 + C_n^1 = 1 + n = 2^{n-k},$$

або

$$1 + n = 2^{n-k} = 2^s = 2^p.$$

Отже, $n = 2^p - 1$, $p \geq 3$ відповідно до умови (1).

Приналежність кодів Хеммінга до класу *досконалих* привертає до них специфічну увагу.

Мірою надлишковості коду є відношення: $r_c = k / n < 1$ – так звана *швидкість кодування*. Бажано досягати найбільшої виправної здатності коду при наближенні швидкості кодування до 1.

Виходячи із принципів формування блоків коду Хеммінга, можна стверджувати, що із зростання довжини блоку n зростає швидкість кодування r_c .

Дійсно, для коду Хеммінга (7,4) швидкість кодування $r_c = 0,57$, для коду Хеммінга (15,11) швидкість кодування $r_c = 0,73$, для (31,26) – $r_c = 0,84$, для (511,502) – $r_c = 0,98$, і так далі.

Це наводить на думку, що, при приблизно однакової складності процедур кодування/декодування, слід віддати перевагу кодам Хеммінга із найбільшими значеннями n .

В контексті оцінки швидкості кодування слід звернутися до критерію, який пов'язаний саме з цим показником надлишкових кодів. Для цього можна скористатись межею завадостійкого кодування Плоткіна, яка визначає теоретично досяжну межу співвідношень між швидкістю кодування r_c , довжиною блоку n і виправною здатністю коду через параметр d [2].

Отже, границя Плоткіна визначає максимально можливу виправну здатність, або інформаційну ефективність (у випадку, коли вихідними даними є виправна здатність) і визначається наступним співвідношенням [2, 3]:

$$r_c \leq 1 - \frac{\frac{md-1}{m-1} - 1 - \log_2 m d}{n}, \quad (4)$$

де m – кратність маніпуляції (кількість біт у символі, що передається по каналу зв'язку),

d – кодова відстань Хеммінга, яка визначається за виразом: (2).

Додатковим критерієм ефективності кодів є границя Варшимова-Гільберта, яка визначає гарантовано існування коду із заданими показниками виправної здатності. Але цей критерій

використовується для кодів із значною довжиною ($n > 200$). Стосовно кодів Хеммінга з малими значеннями n це є недоцільною оцінкою.

На рис. 1 зображено границі Плоткіна та Варшамова-Гільберта в координатах $(r_c, d/2n)$ для випадку, коли $n = 1023$.

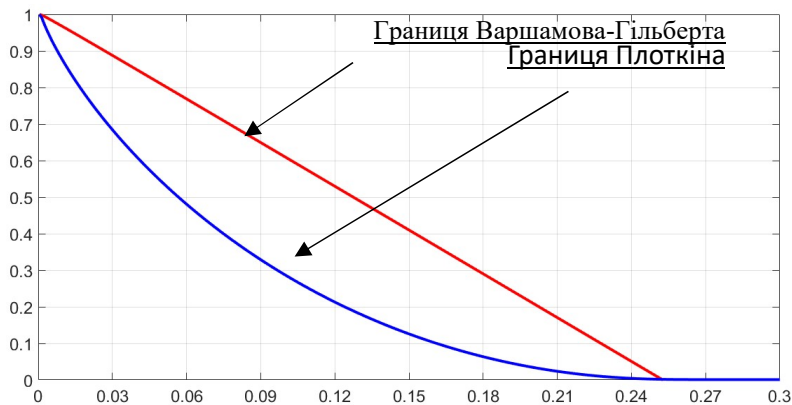


Рис.1. Границі Плоткіна та Варшамова-Гільберта у координатах завадостійкості для $n=1023$

Слід зауважити, що в деяких випадках дослідження виправних властивостей кодів, відображення границі Плоткіна використовується саме в такому вигляді, як на рисунку 1 незалежно від фактичної довжини блоку n . Некоректність такого підходу, який часто веде до хибних висновків щодо фактичних властивостей кодів, ґрунтується на високій чутливості поведінки границі Плоткіна щодо значення n для конкретного коду.

На рис. 2 проілюстровано декілька прикладів розміщення границі Плоткіна для значень $n = 10, 50, 100$.

На рисунку відображено точки, які відповідні кодам Хеммінга $(7, 4)$, $(15, 11)$, $(31, 26)$.

Можна бачити, що жодна координата кодів Хеммінга не перетинає границю Плоткіна за умов відповідності довжини блоку коду n і аргументу n в формулі для границі Плоткіна (4). Порушення саме цієї умови й веде до хибних висновків щодо властивостей коригувальних кодів.

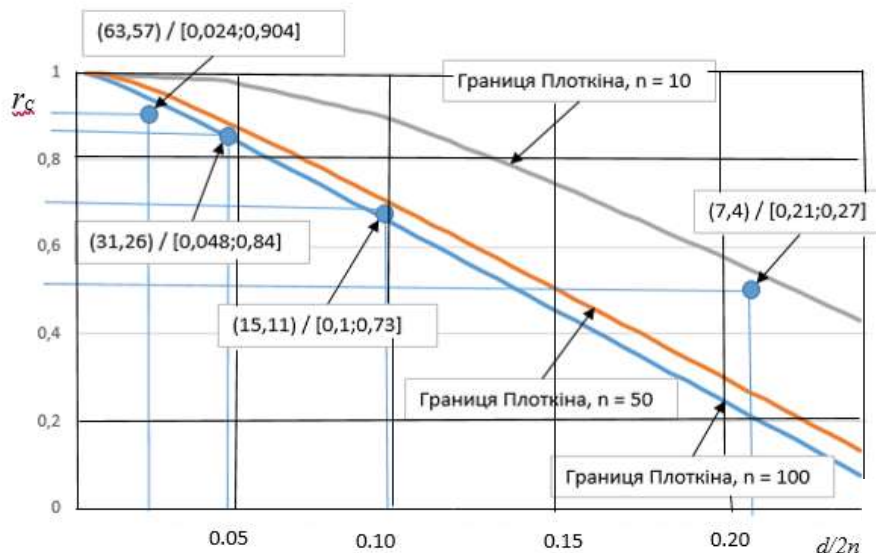


Рис.2. Границі Плоткіна та Варшамова-Гільберта у координатах завадостійкості для $n=1023$

В той же час, можна бачити, що, з одного боку, кожен із кодів достатньо близький до границі Плоткіна, з іншого боку, із збільшенням довжини блоку коду n міра наближеності до границі Плоткіна зростає.

Про це свідчать точні розрахунки міри наближеності кожного із кодів до границі Плоткіна в залежності від довжини блока коду n : зміні n від 7 до 63 міра наближеності монотонно збільшується з 87,24% до 92,24%.

Таблиця 1.
Міра наближення кодів Хеммінга до границі Плоткіна.

Код (n, k)	(7,4)	(15,11)	(31,26)	(63,57)
$d/2n$	0,214286	0,1	0,048387	0,02381
r_c	0,571429	0,733333	0,83871	0,904762
Границя Плоткіна	0,654995	0,838998	0,922096	0,980833
Міра наближеності до границі Плоткіна	87,24%	87,41%	90,96%	92,24%

Отже, виходячи з критеріїв теорії кодування, перевага на стороні кодів Хеммінга із *найбільшим значенням* довжини блока коду n .

Наступним етапом дослідження кодів Хеммінга є визначення критеріїв ефективності їх застосування, які використовуються у теорії завадостійкості, та відповідна оцінка різновидів коду Хеммінга.

Відповідно до цієї групи критеріїв найкращим є код, який надає найкращі показники коригування при найменшому значенні параметру каналу передавання даних h^2 , відомого як відношення сигнал/ шум за умов забезпечення потрібної достовірності $p_{\text{потр}}$.

Визначимо для практичного випадку дослідження ймовірність помилки в будь-якому символі нескінченної інформаційної послідовності $p_{\text{потр}} = 10^{-6}$.

Дійсно, за визначенням :

$$h^2 = \frac{E_S}{N_0}, \quad (5)$$

де E_S – енергія корисного сигналу; N_0 – спектральна щільність шуму в каналі.

Поряд із цим:

$$h^2 = \frac{E_S}{N_0} = \frac{P_S \cdot T_S}{N_0} = \frac{P_S}{N_0 \cdot V_C} = \frac{P_S}{N_0 \cdot \Delta F} = \frac{P_S}{P_{\text{ш}}} , \quad (6)$$

де P_S – потужність корисного сигналу; T_S – протяжність елементу сигналу, обернено пропорційна швидкості передавання $V_C = 1/T$ за умов $V_C = \Delta F$; ΔF – смуга частот в каналі передавання; $P_{\text{ш}}$ – потужність шуму.

В разі, якщо інформаційна послідовність передається без кодування, за обраним критерієм достовірності $p_{\text{потр}} = 10^{-6}$ можна визначити значення параметру h^2 для заданого виду дискретної модуляції [4,5].

Наприклад, при модуляції ФМ-2 (бінарної фазової маніпуляції) розрахункове значення за формулою з використанням функції Крампа $\Phi(x)$ [4]:

$$P_{\text{пом ФМ-2}} = \frac{1}{2} \left[1 - \Phi(\sqrt{2Q^2}) \right]. \quad (7)$$

дає значення $h^2 = 11,8$ (10,7 дБ).

Використання надлишкових (коригувальних) кодів (зокрема, кодів Хеммінга) має покращити енергетичні умови використання каналу передавання даних.

Оскільки виправна здатність всіх кодів Хеммінга однакова, тобто виправлення лише **однієї** помилки в блоці із n елементів, - то перевага має бути віддана коду, який досягає достовірності $p_{\text{потр}} = 10^{-6}$ при найменшому значенні параметра h^2 (потужність шуму вважаємо незмінною, оскільки швидкість потоку даних вважаємо сталою незалежно від способу її розбиття на блоки).

Визначимо умови виникнення саме однієї помилки в блоці із n елементів, при якому може бути доцільним використання саме кодів Хеммінга.

Узагальнена формула має вигляд:

$$P(m, n) = C_n^m \cdot p^m \cdot (1 - p)^{n-m}, \quad (8)$$

де p – ймовірність поодинокій помилки елемента сигналу в неперервному потоці в каналі із енергетичним параметром h^2 .

Співвідношення (7) відповідає умові нормування:

$$\sum_0^n P(m, n) = 1, \quad (9)$$

тобто існує ймовірність появи в переданому блоці будь-якої кількості помилок (від 0 до n).

Область доцільного використання кодів Хеммінга обмежується умовою:

$$P_{\text{Хем}} = 1 - [P(0, n) + P(1, n)] \leq 10^{-6}, \quad (10)$$

тобто результативна достовірність в каналі передавання у випадку появи більш ніж однієї помилки (її виправить код), решта помилок не погіршить поріг достовірності.

Для випадку $m=0$ співвідношення (8) має вигляд:

$$P(0, n) = (1 - p)^n, \quad (11)$$

Для випадку $m=1$ співвідношення (8) має вигляд:

$$P(1, n) = n \cdot p \cdot (1 - p)^{n-1}, \quad (12)$$

В табл.і 2 надані результати розрахунку значень ймовірності безпомилкового передавання та передавання з однією помилкою блоків коду Хеммінга довжиною 7, 15 та 31 біт з використання співвідношень (7), (10), (11), (12).

Таблиця 2
Енергетичні пороги використання кодів Хеммінга.

Довжи на блоку	(7)		(15)		(31)		Без кодування
$P_{\text{пом.}}$ ФМ-2	$P_{\text{Хем}}$	h^2	$P_{\text{Хем}}$	h^2	$P_{\text{Хем}}$	h^2 , дБ	h^2
0,001	2 $,09 \cdot 10^{-5}$	$,8$	1 $,04 \cdot 10^{-4}$	$,8$	4 $,56 \cdot 10^{-4}$	$,8$	,82
0,00022	1 $,01 \cdot 10^{-6}$	,3	5 $,07 \cdot 10^{-6}$	$,3$	2 $,24 \cdot 10^{-5}$	$,3$	,0
0,0001	2 $,01 \cdot 10^{-7}$	$,0$	1 $,05 \cdot 10^{-6}$	,0	4 $,64 \cdot 10^{-6}$	$,0$	,47
0,00005	5 $,25 \cdot 10^{-8}$	$,6$	2 $,62 \cdot 10^{-7}$	$,6$	1 $,16 \cdot 10^{-6}$	,6	,81
0,000001							11,8 / 10,72 дБ

Графічно результати, надані в Таблиці 2, відображені на рисунку 3, відношення сигнал/шум подано в децибелах (дБ).

Дані Таблиці 2 та графіки на рис.3 свідчать проте, що для коду Хеммінга (7,4) умови забезпечення необхідної достовірності досягаються при відношенні сигнал/шум $h^2 = 6,3$ (8 дБ).

При збільшенні довжини блоку ймовірність помилки на його довжині зростає, тому для коду (15,11) критичним є $h^2 = 7,0$ (8,5 дБ), а для коду (31,26) – відповідно $h^2 = 7,6$ (8,8 дБ).

Слід зауважити, що при використанні кожного з кодів, які аналізовані, має місце виграш у забезпеченні потрібної достовірності у порівнянні із випадком відсутності кодування, а саме: код (7,4) забезпечує виграш у відношенні сигнал/шум $\Delta h^2 = 2,72$ дБ у порівнянні із випадком без кодування, код (15,11) забезпечує виграш $\Delta h^2 = 2,25$ дБ, код (31,26) $\Delta h^2 = 1,94$ дБ.

Можна бачити, що зі збільшенням довжини блоку коду енергетичний виграш спадає.

Отже, за критеріями завадостійкості перевагу має код із найменшою довжиною, тобто код Хеммінга (7,4).

Заключним етапом дослідження кодів Хеммінга є оцінка кодів у просторі показників ефективності за критеріями теорії інформації.

З позицій теорії інформації суттєвим є той факт, що кожен із кодів Хеммінга є надлишковим, тобто при передаванні сигналів по каналу зв'язку частина інформації має втрачатися, якщо швидкість передавання по каналу зв'язку дорівнює швидкості джерела інформації. Втрати зменшуються, коли зростає швидкість кодування I_c по мірі зростання довжини блоку коду.

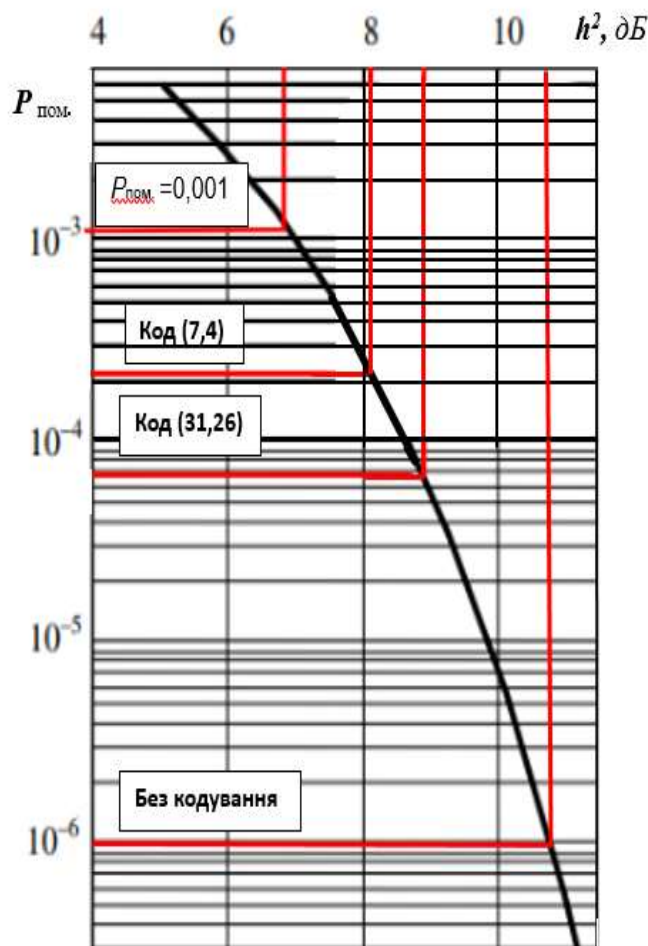


Рис.3. Показники енергетичної межі досягнення потрібної достовірності кодами Хеммінга (7, 4) та (31, 26) у порівнянні із випадком без кодування.

Для підтримки швидкості передавання біт інформації від джерела V_s слід відповідним чином пришвидшити швидкість передавання символів в каналі V_c [6]. В Таблиці 3 наведено відносно необхідне зростання швидкості передавання сигналів із надлишковим кодуванням в залежності від довжини блоку коду n .

Таблиця 3.

Відносне зростання швидкості передавання сигналів із надлишковим кодуванням в залежності від довжини блока коду n .

Довжина блоку n	(7)	(15)	(31)
Швидкість кодування	4/7	11/15	26/31
Відносне прискорення передачі сигналів (V_c / V_s)	1,7 5	1,36	1,19

Важливо уявити, за рахунок чого слід підвищувати швидкість передавання сигналів в каналі зв'язку і як це має позначитися на показниках, наведених у Таблиці 2.

Для відповіді слід звернутися до співвідношення (6), із якого можна побачити, що зміна значення відношення сигнал / шум (h^2) залежить виключно від зміни пропорції між потужністю сигналу P_s і швидкості передавання V_c :

$$h^2 = \frac{P_s}{N_0 \cdot V_c} . \quad (13)$$

Тому **прискорення** швидкості передавання сигналів по каналу V_c по відношенню до швидкості виходу біт інформації від джерела V_s у відповідності до даних Таблиці 3 має супроводжуватися пропорційним **зростанням потужності сигналу P_s** в точці реєстрації закодованого сигналу. При цьому зберігаються відомості про потрібні відношення сигнал шум при відповідному виді коду Хеммінга (Таблиця 2) із збереженням потрібної достовірності.

Отже, в підсумку слід констатувати, що за інтегральним показником – енергетичних витрат для досягнення потрібної достовірності без втрат швидкості передавання інформації від джерела – перевагу слід віддати коду Хеммінга із параметрами (7, 4) із виграшем у відношенні сигнал/шум $\Delta h^2 = 2,72$ дБ у порівнянні із випадком без кодування.

Висновки

1. Специфічний інтерес до кодів Хеммінга викликаний тим, що вони мають унікальну властивість: вони – за класичним визначенням – є досконалими кодами, тобто із мінімальною кількістю надлишкових розрядів.

Мета статті полягає у визначенні їх найбільш ефективної області використання в різноманітних сферах. Складність її досягнення полягає в тому, що поняття «найбільш ефективна область використання» в кожній із теорій кодування, завадостійкості та інформації є особливим, тому комплексний аналіз властивостей коду Хеммінга в суперпозиції цих теорій є актуальною і цікавою задачею дослідження.

2. З позицій теорії кодування коди Хеммінга мають однакову виправну здатність – незалежно від довжини блоку n виправляють лише одну помилку у блоці. В той же час, із зростанням довжини блоку n спадає надлишковість коду, наближаючи швидкість кодування до 1. Одночасно зростає міра наближеності кодів Хеммінга до границі Плоткіна, яка визначає теоретично досяжну межу співвідношень між швидкістю кодування, довжиною блоку і виправною здатністю коду. Тому перевага за цими критеріями має надаватися кодам Хеммінга з найбільшою довжиною блока коду n .

3. З позицій теорії завадостійкості найкращим є код, який надає найкращі показники коригування при найменшому значенні параметру каналу передавання даних h^2 (відношення сигнал/ шум) за умов забезпечення потрібної достовірності $p_{\text{потр}}$.

На прикладі значення $p_{\text{потр}} = 10^{-6}$, використання когерентної обробки модульованих сигналів ФМ-2 показано, що при однаковій виправній здатності всіх кодів Хеммінга перевага має бути віддана коду, який досягає достовірності $p_{\text{потр}} = 10^{-6}$ при найменшому значенні параметра h^2 . У порівнянні із випадком відсутності кодування найбільший вигреш дає найкоротший код – код Хеммінга (7,4).

4. З позицій теорії інформації кожен із кодів Хеммінга є надлишковим, тобто при передаванні сигналів по каналу зв'язку частина інформації має втрачатися, якщо швидкість передавання по каналу зв'язку дорівнює швидкості джерела інформації. Втрати зменшуються, якщо зростає

швидкість кодування по мірі зростання довжини блока коду n . Для підтримки швидкості передавання біт інформації від джерела слід відповідним чином пришвидшити швидкість передавання символів в каналі. При цьому енергетичні співвідношення для кодів Хеммінга з різними параметрами є незмінними.

5. В підсумку слід констатувати, що за інтегральним показником – енергетичних витрат для досягнення потрібної достовірності без втрат швидкості передавання інформації від джерела – перевагу слід віддати коду Хеммінга із параметрами (7, 4) із виграшом у відношенні сигнал/шум $\Delta h^2 = 2,72$ дБ у порівнянні із випадком без кодування.

Література

1. https://en.wikipedia.org/wiki/Hamming_code.
2. https://en.wikipedia.org/wiki/Hamming_bound.
3. Uryvsky L. Peshkin A. Assessment of information efficiency of error-correcting codes in Plotkin bound. 2017 International Conference Radio Electronics & Info Communications (UkrMiCo). IEEE Xplore Digital Library. P. 1 - 4.
4. Основи теорії телекомунікацій. Підручник. Під заг. ред. Ільченка М.Ю.К.: ICC3I, КПІ, 2010.786 с.

Refereses

1. https://en.wikipedia.org/wiki/Hamming_code.
2. https://en.wikipedia.org/wiki/Hamming_bound.
3. Uryvsky L. Peshkin A. Assessment of information efficiency of error-correcting codes in Plotkin bound. 2017 International Conference Radio Electronics & Info Communications (UkrMiCo). IEEE Xplore Digital Library. P. 1 - 4.
4. Osnovy teorii telekomunikatsii. Pidruchnyk. Pid zah. red. Ilchenka M.Iu.K.: ISSZI, KPI, 2010.786 s.
5. Chiani M. Win M.; Zanella A. Error probability for optimum.

УДК 621.391:537.86:519

ДЕКОМПОЗИЦІЯ МОДЕЛІ НЕСТАЦІОНАРНОГО ВИПАДКОВОГО ПРОЦЕСУ

DOI 10.36994 / 2788-5518-2023-01-05-11

Тихонов В.А., д.ф.-м.н., проф. Харківський національний університет радіоелектроніки, Україна.
vyacheslav.tykhonov@nure.ua

Безрук В.М., д.т.н., проф. Харківський національний університет радіоелектроніки, Україна.
valeriy_bezruk@ukr.net

Анотація: Розглядаються особливості виконання декомпозиції моделі нестационарного випадкового процесу, яка полягає в роздільному визначенні довільного тренду, сезонної складової та стаціонарної складової. Це дозволяє вирішувати завдання прогнозування, що означає на кілька кроків уперед передбачення значень випадкового процесу. Показано деякі проблеми, що виникають при використанні моделі авторегресії-проінтегрованого ковзного середнього для декомпозиції моделі випадкового нестационарного процесу. Наведено результати досліджень на реальних даних, що демонструють процес декомпозиції моделі нестационарного випадкового процесу.

Ключові слова: випадковий процес, модель, декомпозиція, тренд, сезонна складова, стаціонарна складова, прогнозування.

DECOMPOSITION OF THE NON-STATIONARY RANDOM PROCESS MODEL

Vyacheslav Tikhonov, Dr.habil., Prof. Kharkiv National University of Radio Electronics, Ukraine.
vyacheslav.tykhonov@nure.ua

Valeriy Bezruk, Dr.habil., Prof. Kharkiv National University of Radio Electronics, Ukraine.
valeriy_bezruk@ukr.net

Abstract: In the work, the peculiarity of the task of constructing the decomposition of a non-stationary model of a random process is studied using the examples of the analysis of two types of real data. The peculiarities of the decomposition of the model of a non-stationary random process are considered. It consists in the separate definition of an arbitrary trend, a seasonal component and a stationary component. This allows solving, for example, forecasting tasks. This task means several steps ahead in predicting the values of a random process. Some problems that arise when using the autoregression-integrated moving average (ARIMA) model for the decomposition of the model of a random non-stationary process are shown. The results of research on real data are given. They demonstrate the process of decomposition of a model of a non-stationary random process. The analyzed classes of random non-stationary processes with a trend allow us to solve two main problems with acceptable accuracy. The first task involves the decomposition of a model of a random non-stationary process. It includes the calculation of an arbitrary trend, a quasi-periodic seasonal component and a stationary component of the process. The second task involves the actual forecasting of a non-stationary time series using the ARIMA model. In the ARIMA model, it is assumed that the trends are deterministic and are linear, quadratic or higher order polynomials. The seasonal component is periodic and has equal counts in each period, which can be removed through the subtraction period. If the seasonal component does not satisfies these requirements, then when it is removed by difference operators, problems arise and the ARIMA model may lose accuracy.

Key words: random process, non-stationary model, decomposition, trend, seasonal component, stationary component, forecasting

Вступ

При вирішенні практичних завдань обробки реальних даних, які мають випадковий характер, виникає необхідність у роздільному визначенні тренду, сезонної складової та стаціонарної складової аналізованого випадкового процесу. Питання побудови математичних моделей випадкових процесів розглядаються у ряді робіт [1-3]. Однак цим питанням не приділено належної уваги. У даній статті детально розглядаються практичні особливості виконання декомпозиції моделі нестационарного випадкового процесу, яка полягає визначенні тренду, сезонної складової та стаціонарної складової процесу. При цьому може вирішуватися завдання прогнозування, тобто на кілька кроків уперед передбачати значення випадкового процесу. Наведено результати досліджень на реальних даних, які демонструють процес декомпозиції

моделі нестационарного випадкового процесу. Досліджено питання побудови моделі авторегресії-проінтегрованого ковзного середнього (АРПКС) та її зв'язок з декомпозицією нестационарного процесу. Розглянуто вплив оператора взяття різниці складових нестационарного процесу на точність декомпозиції та прогнозування. Показано, що точна компенсація тренду та сезонної складової можлива у разі їх детермінованості та строгої періодичності. У разі невиконання цих вимог модель АРПКС та декомпозиція неточні, а похибка компенсуються моделлю стаціонарної складової, що представляється у загальному випадку моделлю авторегресії-ковзного середнього.

Виконання досліджень

В роботі проведено дослідження особливостей завдання декомпозиції нестационарної моделі випадкового процесу на прикладах аналізу двох типів реальних даних. Проаналізовані класи випадкових нестационарних процесів з трендом дозволяють вирішувати з прийнятною точністю дві основні задачі. До першої задачі відноситься декомпозиція моделі випадкового нестационарного процесу, тобто обчислення довільного тренду, квазіперіодичної сезонної складової та стаціонарної складової процесу. До другої задачі належить власне прогнозування нестационарного часового ряду з використанням моделі АРПКС. У моделі АРПКС вважається, що тренди детерміновані і є лінійними, квадратичними або є поліномами вищого порядку. Сезонна складова періодична і має рівні відліки в кожному періоді, які можна видалити через період віднімання. Якщо сезонна складова не задовольняє цим вимогам, то при її видаленні операторами взяття різниці виникають проблеми і модель АРПКС може втратити точність. Приклади реальних нестационарних даних з трендом, що досить сильно відрізняються, представлені на рис. 1 та рис. 2. Перший випадковий процес (рис. 1) є трафік передачі даних у системі мобільного зв'язку [4]. Довжина часового ряду складає 365 відліків, що відповідає 365 дням. Він має значний зростаючий тренд. Дані містять сезону складову з періодом, рівним приблизно 7 днів, і стаціонарний процес, який описується в загальному випадку моделлю АРКС. Другий випадковий процес (рис. 2) є даними про середньомісячні температури, зафіксовані у метеорологічній обсерваторії міста Пулково. Повна довжина ряду становить 3108 значень температури, отриманих обсерваторією за 259 років з 1752 по 2010 роки. Дані містять слабкий тренд, яскраво виражені сезонні коливання температур з періодом 12 місяців та стаціонарну складову процесу. У даній статті шляхом усереднення знайдено тренди процесів, що представлені на рис. 1 та рис. 2. Для цього щоденний трафік передачі даних (рис. 1) представлено у вигляді ряду середньотижневих даних:

$$x^n[t] = \overline{x_0[t]}, \overline{x_1[t]} \dots \overline{x_{m-1}[t]}, \quad (1)$$

де

$$\overline{x_k[t]} = \frac{1}{n} \sum_{i=1}^n x_i[t], \quad n = 7. \quad (2)$$

У процесах (1) та (2) час визначено як номер даних середнього значення трафіку за тиждень. Для оцінки довготривалого тренду середньорічні температури, що представлені на рис. 2, також перетворені за допомогою виразу (1), де середньорічні значення температур обчислюються як:

$$\overline{x_k[t]} = \frac{1}{n} \sum_{i=1}^n x_i[t], \quad n = 12. \quad (3)$$

Для даних у виразі (3) час визначено як послідовність років.

Для отримання трендів середньотижневі значення даних (1) згладжені низькочастотним фільтром, що ковзає. Щоб накласти тренд на вихідні дані, тобто щоб довжина вихідних даних та даних, отриманих усередненням збігалася, продовжені значення трендів на 7 або 12 відліків відповідно. Тоді отримаємо часові ряди трендів, що дорівнюють початковій довжині даних.

Отримані тренди показані на рис. 1 та рис. 2. Зауважимо, що процедура ковзного усереднення не дає можливості отримати на початку часового ряду і в його кінці таке ж точне усереднення як і у середині часового ряду. Це зв'язано з очевидними особливостями процедури усереднення,

коли на початку часового ряду відсутні необхідні для усереднення попередні значення часового ряду, а кінці ряду – наступні значення. Як видно із графіків на рис. 1 та рис. 2, знайдені тренди мають довільну форму, а не детерміновану (лінійну, параболічну), як це визначено в моделі АРПКС. У цілому нині тренди неоднорідні, тк як у них є ділянки з різною швидкістю зміни.

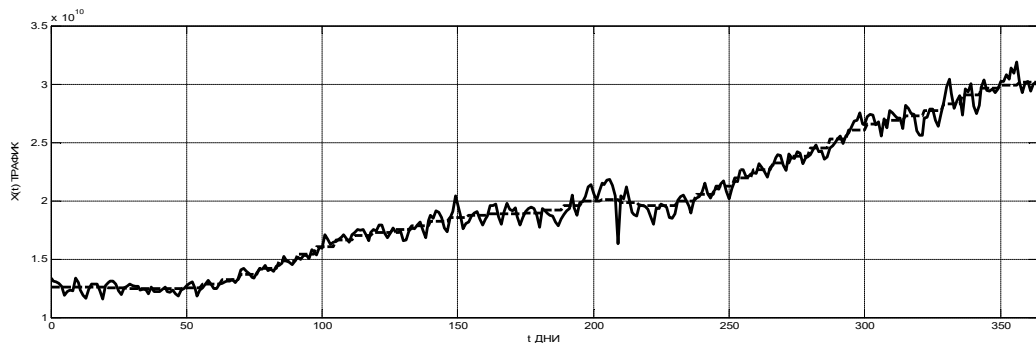
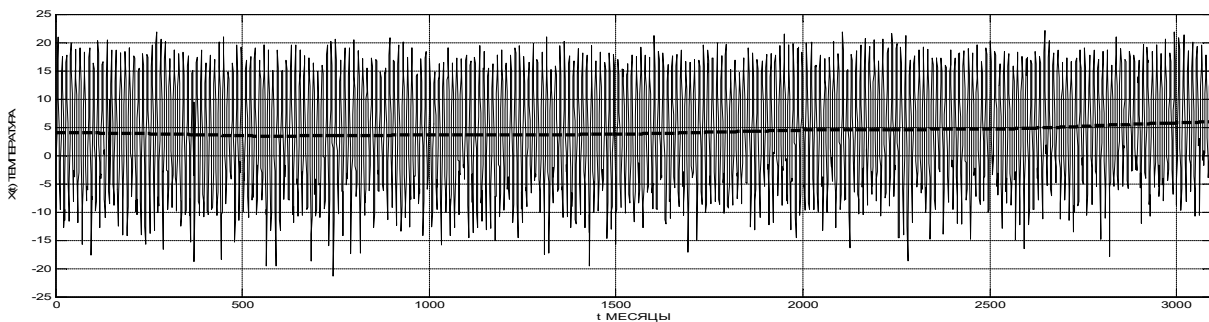


Рис. 1. Трафік передачі даних у системі стільникового мобільного зв'язку та оцінка його тренду

Аналіз тренду на рис. 2 показує, що з 1752 до 1803 температура падала з $4,05^{\circ}\text{C}$ до $3,45^{\circ}\text{C}$. Спад температури склав $0,6^{\circ}\text{C}$. З 1804 року спостерігалось зростання температури з $3,45^{\circ}\text{C}$ до $5,96^{\circ}\text{C}$



у 2010 році. Приріст температури за цей термін становив $2,51^{\circ}\text{C}$

Рис.2. Дані середньомісячних температур, зафіксованих у метеорологічній обсерваторії міста Пулково, та оцінка тренду

Як показує аналіз рис. 1 та рис. 2, отримані тренди мають довільний характер.

Модель АРПКС нестационарних даних та їх декомпозиція

Розглянемо особливості побудови моделей АРПКС для нестационарних процесів. Відповідно до [1] моделюємий нестационарний процес може мати детермінований тренд. У рамках моделі нестационарних процесів сезонна складова має бути періодичною функцією з певним періодом. Завданням побудови моделі АРПКС є видалення з вихідних даних тренду та сезонної складової за допомогою відповідних операторів взяття різниці та побудова моделі лінійного передбачення АРКС стаціонарного залишку. Після цього можна вирішувати завдання прогнозування [1] шляхом відновлення тренду та сезонної складової нестационарного процесу, а також побудови моделі АРКС стаціонарної складової.

Модель АРПКС нестационарного процесу представляється у вигляді мультиплікативного процесу $\omega[t]$ [2]:

$$\omega[t] = \nabla^d \nabla_s x[t] = (1-z)^d (1-z^s) x[t] = \nabla^d \omega_1[t]. \quad (4)$$

Оператори взяття різниці, що входять до (4), враховують сезону складову та тренд шляхом їх віднімання. Для видалення сезонної складової застосовується операція $\nabla_s = 1 - z^s$, где z^s - оператор зсуву, дія якого визначається виразом $z^s x[t] = x[t-s]$. Процес видалення сезонної складової за допомогою оператора зсуву визначається виразом:

$$\omega_1[t] = \nabla_s x[t] = (1 - z^s)x[t]. \quad (5)$$

Детермінований тренд у $\omega_1[t]$ визначається оператором $\nabla^d = (1 - z)^d$. Тоді процес без тренда можна представити у вигляді:

$$\omega[t] = (1 - z)^d \omega_1[t]. \quad (6)$$

Як правило, обмежуються найпростішими випадками лінійного та квадратичного трендів. Для цих випадків вважають $d = 1$, $d = 2$. Тоді:

$$\begin{aligned} \omega[t] &= \nabla \omega_1[t] = \omega_1[t] - \omega_1[t-1], \\ \omega[t] &= \nabla^2 \omega_1[t] = \nabla(\omega_1[t] - \omega_1[t-1]) = (\omega_1[t] - 2\omega_1[t-1] + \omega_1[t-2]). \end{aligned}$$

Таким чином, тренд (5) та сезонна складова (6), що визначаються операторами зсуву, видаляються. Під час прогнозування ці детерміновані складові відновлюються.

При ширшому аналізі нестационарних процесів інтерес представляє також процедура знаходження довільного тренду та сезонної складової, що виходить за межі побудови АРПКС моделі. При цьому форма тренду та властивості сезонної складової можуть змінюватися у ширших межах, ніж у моделі АРПКС. Це завдання називають декомпозицією моделі нестационарного часового ряду, що полягає у його розкладанні на складові [3].

Декомпозиція у моделі нестационарного процесу розуміється як представлення у вигляді:

$$x[t] = m[t] + c[t] + y[t], \quad (7)$$

где $x[t]$ - нестационарний процес, $m[t]$, $c[t]$, $y[t]$ - відповідно тренд, сезонна складова та корельована стаціонарна складова.

Похибки декомпозиції моделі нестационарного процесу

Зауважимо, що оператор $\nabla_s = 1 - z^s$, що усуває сезонні коливання в (4), діє не тільки на цю складову процесу, а й на трендову складову. При цьому стаціонарна складова нестационарного процесу також змінюється. Після дії оператором $\nabla^d = (1 - z)^d$ на тренд інші складові моделі декомпозиції нестационарного процесу можуть змінитися. Найбільший вплив оператора видалення сезонної складової. Після видалення сезонної складової оператором $\nabla_s = 1 - z^s$ декомпозиція (7) моделі нестационарного процесу набуває вигляду:

$$\nabla_s x[t] = m[t] - m[t-s] + c[t] - c[t-s] + y[t] - y[t-s], \quad (8)$$

де $\nabla_s x[t]$ - нестационарний процес без сезонної складової.

Якщо сезонна складова строго детермінована, то $c[t] = c[t-s]$ та декомпозиція (8) спрощується [2] :

$$\nabla_s x[t] = m[t] - m[t-s] + y[t] - y[t-s]. \quad (9)$$

Однак, як видно із рис. 1 та рис. 2, сезонна складова не детермінована і тому вираз (9) є несправедливим для цих даних.

На рис. 3 та рис. 4 показані дані трафіку та середньомісячних температур після видалення сезонних складових. На даних, що представлені на рис. 3, не видно тренд, який добре проглядається на рис. 1. За графіком на рис. 4 також важко визначити довготривалий тренд температур. Таким чином, видалення сезонної складової оператором $\nabla_s = 1 - z^s$ з вихідних даних також суттєво впливає і на тренд.

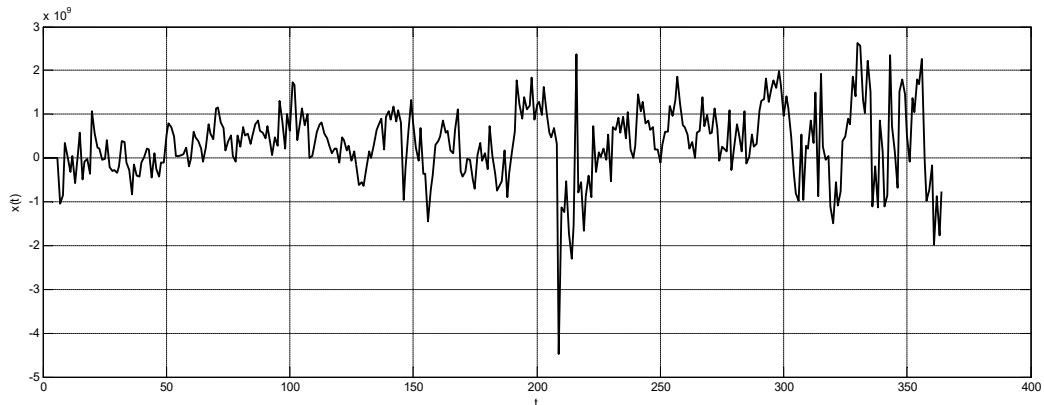


Рис. 3 Трафік передачі даних у системі стільникового мобільного зв'язку після видалення сезонної складової

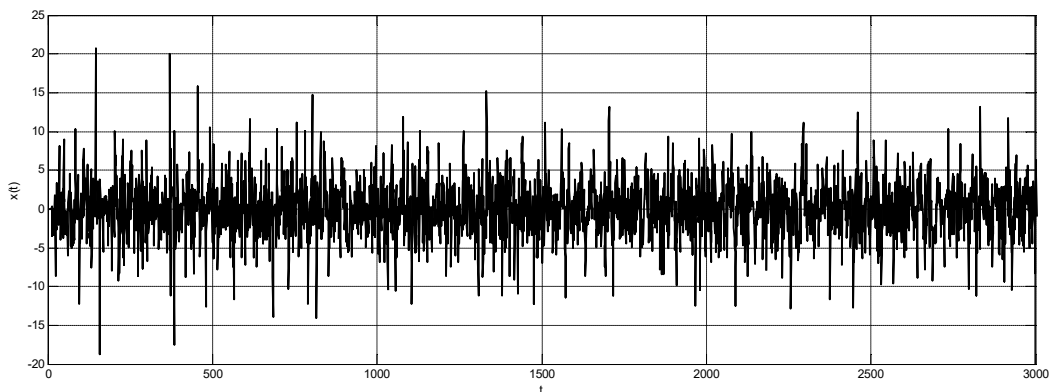


Рис. 4. Вибірка даних середньомісячних температур після видалення сезонної складової

Операція видалення тренду оператором $\nabla^d = (1 - z)^d$ при $d = 1$ менше спотворює дані. Однак слід мати на увазі, що дія оператора $\nabla = (1 - z)$ по суті зводиться до диференціювання випадкового процесу. Таким чином, доводиться досліджувати властивості не безтрендового процесу, а його дискретної похідної. На рис. 5 та рис. 6 показані дані трафіку та середньомісячних температур після видалення тренду. Аналіз даних на рис. 5 та рис. 6 показує, що тренд дійсно видаляється.

Наявність сезонної складової даних, можна визначити за графіками спектральної щільності потужності (СЩП) аналізованих даних [3]. Вони показані на рис. 7 та рис. 8. Аналіз графіка спектра, поданого на рис. 7 показує, що сезонна складова трафіку даних менше виражена. Форма спектру показує, що в отриманих після диференціювання даних трафіку, є постійна складова, сезонна складова з періодом 6,8 днів, а також слабкі коливання з періодами 3,4 і 2,3 дня. Як видно із рис. 8 більш явно сезонна складова виражена в даних температур. З аналізу спектра, показаного на рис. 8, слідує, що період коливань сезонної складової дорівнює 11,9 днів. Зазначимо, що для деяких форм сезонної складової диференціювання не змінює період коливань.

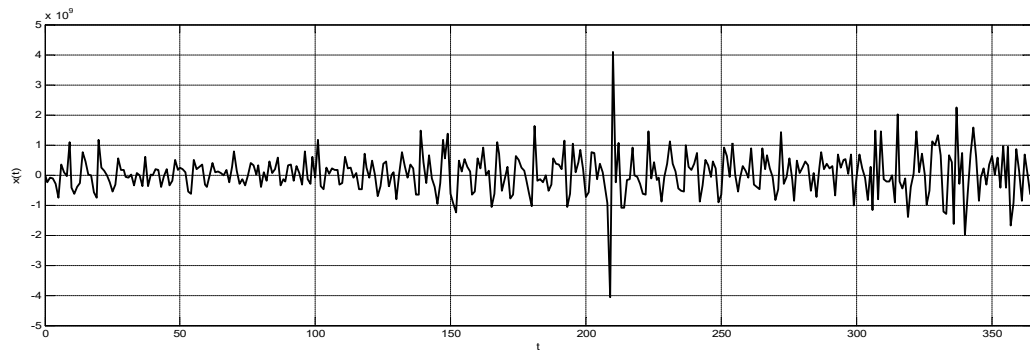


Рис. 5. Трафік передачі даних у системі мобільного зв'язку після видалення тренда

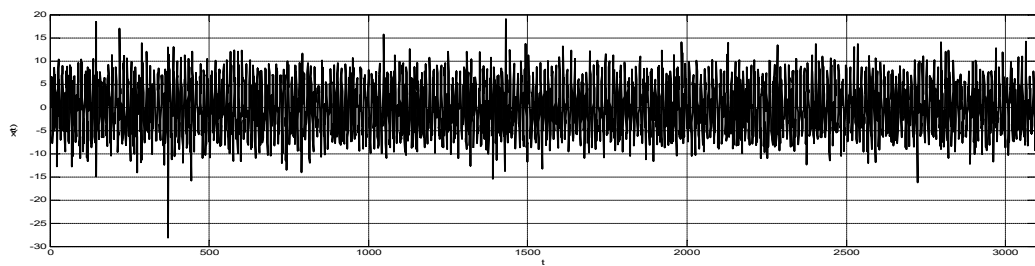


Рис. 6. Дані про середньомісячні значення температури після видалення тренда

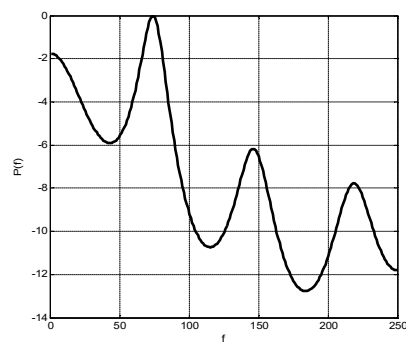


Рис. 7. СЩП після видалення тренду в даних трафіку у системі мобільного зв'язку

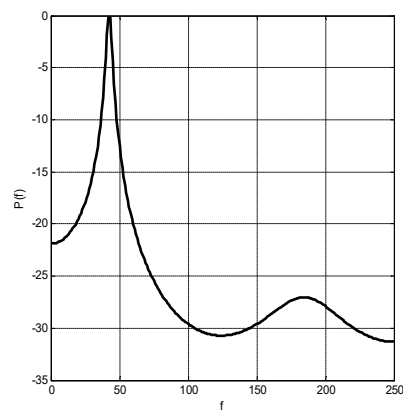


Рис. 8. СЩП після видалення тренду з даних температур

Висновки

У роботі розглянуто теоретичні і практичні особливості декомпозиції моделі нестационарного процесу, яка полягає у роздільному визначенні тренду, сезонної складової та стаціонарної складової. Це дозволяє вирішити задачу прогнозування випадкового процесу. Проведено дослідження особливостей задачі декомпозиції нестационарної моделі випадкового процесу на прикладах аналізу двох типів реальних даних. Розглянуто випадковий процес, що являє собою трафік передачі даних у системі мобільного зв'язку. Розглянуто також випадковий процес, що являє собою дані про середньомісячні температури, зафіксовані у метеорологічній обсерваторії міста Пулково. Показано деякі проблеми, що виникають при використанні моделі у вигляді АРПКС та декомпозицію моделі для аналізу нестационарних даних, що містять трендову, сезонну та стаціонарну складові. Використання в моделі АРПКС операторів, що визначають ці складові, обмежує допустимі властивості сезонної складової та тренду. Вплив на модель аналізованого процесу операторів зсуву продемонстровано на задачі декомпозиції нестационарного процесу. Застосування цих операторів зводиться до диференціювання вихідних даних, що спотворює характеристики вихідного процесу та призводить до неточностей. Для виділення тренду використані усереднені за деякий період дані з подальшим усередненням низькочастотним фільтром, що ковзає. Наведені результати декомпозиції моделей для вибраних реальних даних, що мають нестационарний характер.

Література

1. Бокс Дж., Дженкинс Г. Анализ временных рядов: Пер. с. англ. М. : Мир, 1974. 406 с.
2. Brockwell P. J., Davis R. A. Introduction to Time Series and Forecasting. Springer, 2002. 434 p.
3. Марпл С. Л. Цифровой спектральный анализ и его приложения. М. : Мир, 1990. 584 с.
4. Тихонов В. А., Чеботарёва Д. В. Прогнозирование потока данных в сетях мобильной связи // *Наукоємні технології в інфокомунікаціях*. Матеріали III Міжнародн. наук.-практ. конф. Харків : Друкарня «Мадрид», 2019, С. 126-127.

References

1. Boks Dzh., Dzhenskyns H. Analyz vremennykh riadov: Per. s. anhl. M. : Myr, 1974. 406 s.
2. Brockwell P. J., Davis R. A. Introduction to Time Series and Forecasting. Springer, 2002. 434 p.
3. Marpl S. L. Tsyfrovoi spektralnyi analiz y eho prylozheniya. M. : Myr, 1990. 584 s.
4. Tykhonov V. A., Chebotarëva D. V. Prohnozyrovanye potoka dannykh v setiakh mobylnoi svyazy // *Naukoiemni tekhnolohii v infokomunikatsiiakh*. Materialy III Mizhnarodn. nauk.-prakt. konf. Kharkiv : Drukarnia «Madrid», 2019, S. 126-127.

УДК 520.6.05/621.3.084.6

ОПТИМІЗАЦІЯ АНАЛОГОВОГО МОДУЛЮ ЗА ШУМОВИМИ ХАРАКТЕРИСТИКАМИ У БОРТОВОМУ ДЕТЕКТОРІ РЕЄСТРАТОРА-АНАЛІЗАТОРА ДЛЯ ІДЕНТИФІКАЦІЇ ЗАРЯДЖЕНИХ ЧАСТИНОК

DOI 10.36994 / 2788-5518-2023-01-05-12

Єзерський Н. В. КПІ імені Ігоря Сікорського, Київ, Україна. billytalent7777@gmail.com
Мовчанюк А. В., к.т.н., доц. КПІ імені Ігоря Сікорського, Київ, Україна. Movchanuk@rtf.kpi.ua

Анотація: В аналогових підсилювачах слабких сигналів постає питання пошуку оптимального співвідношення сигнал/шум за шумовими характеристиками для подальшої дискретизації сигналу цифровими методами. Використання комп'ютерних симуляцій не є рішенням цього питання, адже розробники симуляторів електроніки зазвичай не розкривають хід обчислень, або надають можливість самостійно вводити формули для розрахунку з ймовірною непередбачуваністю результатів прогнозування. В даній роботі розроблено метод розрахунку та оптимізації рівня шумів, який пропонується застосувати при налагодженні роботи схем на базі операційних підсилювачів, що заощадить час та кошти при подальшому тестуванні та оптимізації схеми ще до її виготовлення, оскільки розробники у супровідній документації вказують шумові параметри без врахування схеми включення (з розімкнутим колом зворотного зв'язку та інше). Тому варто розробити підхід по проектуванню підсилювача враховуючи вплив шумових параметрів схеми включення. Ключовими параметрами каскадів, що впливають на достовірність дискретизації серед інших типів обрано власні шуми операційних підсилювачів та теплові шуми елементів схеми включення, що оптимізовано зведенням до одного базису та перевірено відповідності поставленим задачам дослідження. Додатково описана послідовність дій, яка встановлює точність дискретизації системи з впливом шумової напруги. В результаті пришивдшено рішення задачі мінімізації рівня шуму спрощеним комп'ютерним розрахунком. Співвідношення сигнал/шум оптимізовано із застосуванням методу нелінійного програмування. Запропонований метод можна використовувати як для аналізу дискретних, так і до інтегрованих кіл. Результати роботи використано для оптимізації спроектованого аналогового модулю реєстратора-аналізатора MiPA_ер, що планується до запуску на низьку навколоземну орбіту для вивчення природи мікросплесків високоенергетичних електронів та протонів.

Ключові слова: оптимізація; CubeSAT; нелінійне програмування; шумова модель, операційний підсилювач.

OPTIMIZATION OF THE ANALOG MODULE ACCORDING TO NOISE CHARACTERISTICS IN THE ON-BOARD DETECTOR OF THE RECORDER-ANALYZER FOR THE IDENTIFICATION OF CHARGED PARTICLES

Nikita. Yezerskyi . Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine. billytalent7777@gmail.com
Andriy.Movchanuk., Ph.D., Ass..Prof, Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine.
Movchanuk@rtf.kpi.ua

Abstract: In analog amplifiers of weak signals, the question arises of finding the optimal signal/noise ratio based on noise characteristics for further discretization of the signal by digital methods. Using computer simulations is not a solution to this problem. Because the developers of electronics simulators do not reveal the progress of calculations as usual, or provide the opportunity to independently enter formulas for calculation with the unpredictable probability of forecasting results. In current work, a method for calculating and optimizing the noise level has been developed. It is proposed to be used for debugging circuits based on operational amplifiers, which will save time and money during further testing and optimization of the circuit before its manufacture. And the developers in the datasheet specify the noise parameters without taking into account all the cascade (with open feedback loop, etc.). That is why, a method is needed that takes into account the noise parameters of the switching circuit. The key parameters of the cascades, which have an impact on the reliability of sampling, are among other types: inherent noises of operational amplifiers and thermal noises of the elements of the switching circuit, which are optimized by reduction to a single basis and checked for compliance with the research objectives. In addition, the sequence of actions that establishes the accuracy of discretization of the system with the influence of noise voltage is described. As a result, the solution of the noise level minimization problem by simplified computer calculation is accelerated. Signal-to-noise ratio optimization achieved by non-linear

programming method. The proposed method can be used both for analysis discrete and integrated circuits. The results of the work have been used for optimization of the existing analog scheme of the small-caliber spectrometer MIRA_ep that is planning to be launched into low near-Earth orbit to study the nature of the high-energy electron and proton microbursts.

Keywords: *optimization; CubeSAT; non-linear programming; noise model, operational amplifier.*

Вступ

При проектуванні малошумної апаратури не складено однозначного алгоритму зменшення прогнозованого рівня шумів. Така задача зводиться до пошуку максимального співвідношення сигнал/шум SNR (signal-to-noise ratio) при відомому динамічному діапазоні (доступна розрядність АЦП).

Не дивлячись на те, що основні джерела шумів радіоелектронної апаратури є відомими (флікер-шуми, дробові та теплові) та факт, що перші каскади підсилювачів роблять найбільший внесок у загальний рівень шумів, питання побудови схеми за шумовими параметрами залишається дискусійним.

При математичному підході до вирішення даної задачі необхідно скласти багатовимірну функцію оптимізації з коефіцієнтами впливу, де відображено внесок кожного елементу в загальний рівень шумів. Надалі отримана функція в часткових похідних, що прирівнюється до нуля зі знаходженням її екстремуму. Недоліками такого підходу є витрачений час на складання моделі та неможливість застосування у інших випадках.

Такий підхід складно імплементувати при спробі зменшення шуму багатокаскадного підсилювача слабкого сигналу. Прикладом такого підсилювача є модуль аналогової обробки сигналу (АОС) наносупутника для вивчення сейсмічно-магнітосферної кореляції [6]. У такому модулі АОС присутні три спектрометричні канали по 9 каскадів, здебільшого виконаних на операційних підсилювачах (ОП). При цьому рівень шумів впливає на достовірність даних, що призводить отримання невірної кількості заряджених частинок у записі, захоплених коліматором з енергіями в околі границь енергетичних діапазонів.

В даній статті використано метод приведення шумів до входу підсилювача [4] з метою подальшого зменшення рівня шумів за сталого коефіцієнта підсилення, що можна буде застосувати для будь-якого каскаду, виконаного на базі ОП.

Виконання досліджень

Задачею, що постала перед авторами, є мінімізація шумів детектора заряджених частинок, що планується встановити на борту штучного супутника Землі. Умовами роботи завдано допустимий рівень шумів для кожного спектрометричного каналу: 18 мВ для сигналів, отриманих з тонких детекторів п-терфенілу, та 67 мВ для сигналів, отриманих з товстого детектора п-терфенілу [1]. Виділимо головні особливості джерел шумів радіоелектронної апаратури:

1) Дробовий шум, що виникає при перетині р-п переходу. Спектральна щільність цього шуму не залежить від частоти та температури, тобто його середньоквадратичне значення однакове на всіх відрізках. Він проявляється у напівпровідниках набагато сильніше, ніж у провідниках, та є відсутнім без протікання струму [3].

Середньоквадратичне значення визначається згідно [3]:

$$E_{SH} = k * T * \sqrt{\frac{2B}{q * I_{DC}}} \quad (1)$$

де E_{SH} — середньоквадратична напруга дробового шуму, k — стала Больцмана, T — кімнатна температура [К]; B — смуга частот [Гц], q — заряд електрона [Кл]; I_{DC} — струм через р-п перехід [А].

2) Теплові шуми, або шуми Джонсона. Генеруються тепловим рухом електронів у провідниках. Чим сильніше нагрітий провідник, тим більше ЕРС шуму. Тепловий шум зникає тільки при температурі, яка дорівнює абсолютному нулю за шкалою Кельвіна. Зменшення температури знижує цей шум не в значній мірі (поки не йдеться про криогенні температури), але зменшення опору резистора дає більш відчутні результати [3]. З іншого боку, не можна

зменшувати номінали резисторів до надто малих значень через порушення критерію узгодження по шумам, що буде наслідком такого рішення.

Скористаємось співвідношеннями Найквіста, що справедливі для частот до 100 МГц [3]:

$$E_{TH} = \sqrt{4kTRB} \quad (2)$$

де E_{TH} — середньоквадратична напруга теплового шуму [В], R — опір провідника [Ом].

3) Флікер-шум, або так званий шум «1/f». З'являється через дефекти кристалічної решітки напівпровідників. Йому властиве збільшення рівня при зменшенні частоти та постійний рівень для кожної декади, або октави [3]:

$$E_n = K_e * \sqrt{\ln \frac{F_{max}}{F_{min}}}, \quad I_n = K_i * \sqrt{\ln \frac{F_{max}}{F_{min}}} \quad (3)$$

де K_i та K_e — коефіцієнти пропорційності за напругою [В] та струмом [А] на частоті 1 Гц, F_{max} та F_{min} — максимальна і мінімальна частота [Гц].

Оскільки активними елементами в схемі, що підлягає оптимізації, є ОП, то основними джерелами шуму будуть шуми вхідного каскаду ОП, шуми джерела живлення, та зворотного зв'язку (ЗЗ).

Зазвичай, у паспортній документації наводиться спектральна щільність напруги шуму на частоті 1 кГц (нВ/√Гц) і графік залежності еквівалентної напруги шуму від частоти без врахування схеми включення ОП, тобто з розімкненим колом ЗЗ. Значення цих параметрів є істотно заниженими відносно рівня шумів, отриманих при дослідженні реальної схеми. Тому авторами пропонується оцінити рівень шумів підсилювача на базі ОП та створити методику його зменшення.

Найважливішою є оптимізація першого каскаду формуючого підсилювача, оскільки його шуми отримають найбільше підсилення в подальших каскадах до дискретизації.

У [4] розглянуто шумову модель каскаду інвертуючого ОП з джерелом сигналу у вигляді шуму джерела струму з високим внутрішнім опором. В даному випадку маємо джерело напруги шуму з малим внутрішнім опором, тому необхідно розробити алгоритм оптимізації шуму каскаду з іншими параметрами.

Перелічимо задачі, які вирішуються в даній роботі:

1. Оптимізація теплових шумів, як основної компоненти, присутньої в суміші сигналу та шуму.
2. Покращення SNR за рахунок оптимізації рівня шумів каскадів на ОП.
3. Досягнення мінімального рівня шумів при сталому коефіцієнті підсилення та динамічному діапазоні підсилювача для стабільної роботи підсилювачів.
4. Порівняння розрахованих значень з моделюванням у програмних пакетах проектування схем для підвищення довіри отриманим результатам.

Авторами розроблено метод прогнозування допустимого коефіцієнту шуму при відсутності даних про загальний коефіцієнт шуму, коефіцієнт підсилення та вхідний рівень шумів. Суть методу полягає у встановленні основних характеристик аналогового тракту за допомогою другорядних характеристик: відомої форми вхідних/вихідних сигналів функційних елементів схеми та наявного значення динамічного діапазону спектрометричного каналу.

Щоб доповнити відсутніми числовими характеристиками спектрометричних каналів скористаємось наявними епюрами тестових сигналів (рис. 1, 2). Це дозволяє визначити допустимий коефіцієнт шуму тракту за відсутності явно вказаного параметру в документації на виріб (рис. 1), [1].

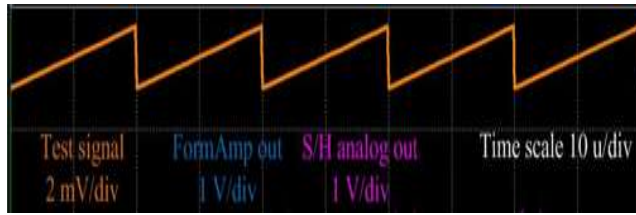


Рис. 1. Осцилограма тестового сигналу для оцінки працездатності каналу

Виходячи зі значення амплітуди тестового сигналу 1 мВ (рис. 1) визначимо допустимий вхідний рівень шумів за відомого значення динамічного діапазону, який дорівнює 46 дБ за напругою [1], тож:

$$U_{\text{ШВх}} = \frac{U_{\text{Тест}}}{D} = \frac{0.001}{80} = 12,5 \text{ [мкВ]} \quad (4)$$

Оскільки допустимий шум каналу 18 мВ, отримаємо коефіцієнт підсилення, який мають забезпечувати аналогові каскади схеми:

$$G = \frac{U_{\text{ШДоп}}}{U_{\text{ШВх}}} = \frac{18 \text{ мВ}}{12,5 \text{ мкВ}} = 63 \text{ [дБ]} \quad (5)$$

що збігається зі значеннями сигналу, продемонстрованого на рис. 2 (масштаб 1 В/кл), [1].

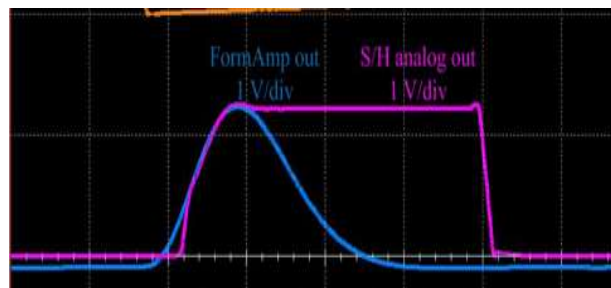


Рис. 2. Форма сигналу після проходження формуючого підсилювача та пікового детектора

Отримаємо загальний коефіцієнт шуму тракту до АЦП за визначенням:

$$N_{\text{ШЗар}} = \frac{\frac{U_{\text{Тест}}}{U_{\text{ШВх}}}}{G * \frac{U_{\text{Тест}}}{U_{\text{ШДоп}}}} = -20 \text{ [дБ]}, \quad (6)$$

Тобто SNR погіршиться на -20 дБ при проходженні аналогової частини схеми.

Для оптимізації каскаду з точки зору підвищення SNR побудуємо еквівалентну схему першого каскаду формуючого підсилювача на ОП. Вихідний шум будь-якого підсилювача є функцією шуму, створеного джерелом сигналу і шуму, створеного всередині підсилювача. Модель шуму підсилювача включає всі внутрішні джерела шуму, приведені до входу. Щоб ці джерела були незалежними від імпедансу джерела сигналу потрібно два джерела шуму —

послідовне джерело напруги v_s і джерело шунтованого струму i_n (рис. 3) [4].

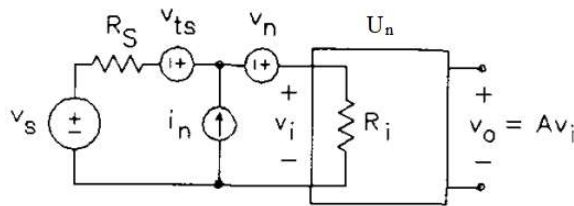


Рис. 3. V_n - I_n , модель шуму підсилювача

На рис. 3 використані наступні позначення: v_s – миттєва напруга джерела сигналу, R_s – опір джерела сигналу, v_{ts} – миттєва напруга теплового шуму створена R_s , v_n – напруга шуму каскаду перерахована до входу ОП.

Для R_s та v_{ts} справедливі співвідношення Найквіста, також це класична ілюстрація шумової моделі опору і є дійсною для резисторів, виготовлених за будь-якими типами технологій [4]. Значення вихідної напруги такого підсилювача (рис. 3) визначається як [4]:

$$v_o = \frac{A * R_i}{R_s + R_i} [v_s + (v_{ts} + v_n + i_n R_s)], \quad (7)$$

де A – коефіцієнт підсилення за напругою, R_i – вхідний опір ОП.

Приведена до входу напруга шуму характеризується еквівалентною вхідною напругою шуму v_{ni} . Тобто це напруга, що підключена послідовно з входом підсилювача, яка створює такий самий шум, як напруга на виході як і всі джерела шуму в колі [4]. Ця напруга є сумою всіх напруг шумів, приведених до входу (рис. 3).

$$v_n = \sqrt{U'_{ш_{RF}} + U_{ш_{ОП}} + (v_{ts})^2 + (v_n)^2 + (i_n R_s)^2} \quad (8)$$

де $U'_{ш_{Rf}}$ — напруга шуму кола ЗЗ приведена до входу, $U_{ш_{ОП}}$ — еквівалентна вхідна напруга шуму ОП без врахування схеми включення.

Розглянемо формуючий підсилювач (рис. 4, 5), який підсилює слабкий сигнал та формує смугу пропускання аналогового тракту наносупутника. На вході стоїть RC-коло з частотою зрізу 3 кГц. Далі перший каскад формуючого підсилювача (ОП U1.2), який формує підйом АЧХ у 20 дБ/декаду, за допомогою пасивного Т-подібного RC мосту у колі зворотного зв'язку. Через застосування такого рішення можна знехтувати флікер-шумами через їхній несуттєвий вплив на частотах, що перевищують сотні Гц. Мостова схема в колі ЗЗ дозволяє уникнути використання великих номіналів резисторів та зменшує чутливість схеми до наводок та завад.

Напруга регулювання включеного пікового детектору формується шляхом подання зміщення на неінвертуючий вхід другого каскаду підсилювача U1.3 (рис. 5). Ця напруга порівнюється з напругою ЗЗ, що регулюється ввімкненням потенціометру в колі ЗЗ ОП U1.4. Таким чином досягається обмеження напруги коливання, яке буде дискретизувати АЦП.

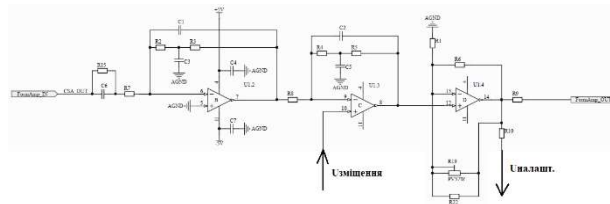


Рис. 4. Схема формуючого підсилювача

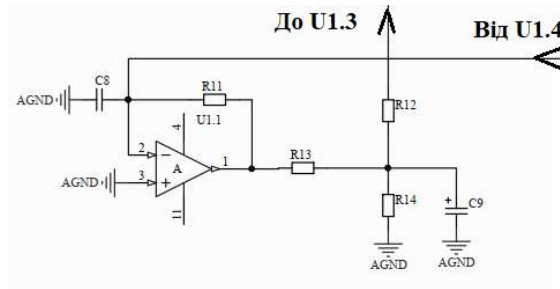


Рис. 5. Схема зміщення ОП U1.3

Вихід формуючого підсилювача підключено до пікового детектора та диференційного підсилювача з одиничним коефіцієнтом підсилення і диференційним виходом для покращення завадостійкості при передачі сигналу з аналогової до цифрової плати.

При використанні послідовного включення багатокаскадних підсилювачів (рис. 4) теоретично можна досягти яких завгодно величин коефіцієнту підсилення, але кожна ланка буде вносити власні шуми, що зменшує можливість підсилювати слабкі сигнали.

Тому опишемо алгоритм дій, щодо зменшення власних шумів каскаду на базі ОП:

1. Визначити основні джерела шуму каскаду
2. Привести джерела шуму до входу підсилювача
3. Звести задачу оптимізації до однієї функціональної залежності
4. Скористатися комп'ютерним розрахунком методом узагальненого приведенного градієнта для оптимізації схеми за шумами.

Розглянемо перший каскад формуючого підсилювача (рис. 6) як приклад реалізації алгоритму:

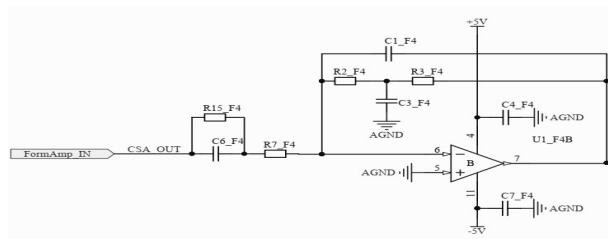


Рис. 6. Перший каскад формуючого підсилювача

Внаслідок перетворень складено еквівалентну схему для визначення коефіцієнту підсилення шуму

(рис. 7).

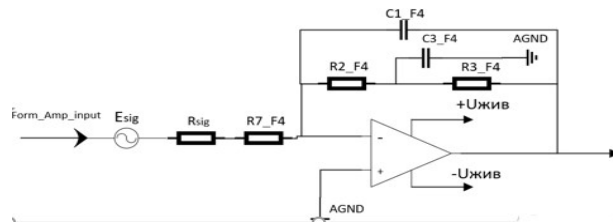


Рис. 7. Еквівалентна схема першого каскаду з врахуванням джерела сигналу

Першим кроком визначимо напруги, що відповідають шумовим напругам елементів схеми, рис. 7 ($U_{ш_R7_F4}$, $U_{ш_R2_F4}$ та $U_{ш_R3_F4}$), які приймаємо рівними значенню теплових шумів резисторів, бо флікер шуми вже придушені RC-колом на вході каскаду. Вважаємо, що резистор на вході каскаду R_{15_F4} не вносить суттєвих шумів через його шунтування конденсатором C_{6_F4} (рис. 6). Дробовими шумами нехтуємо через те, що їхній рівень у провідниках на порядок менший за рівень у напівпровідниках через суттєво менший вміст будь-яких домішок та неоднорідностей, які характерні для напівпровідників [3].

Другим кроком визначимо еквівалентну вхідну напругу шуму за допомогою наступних перетворень та відобразимо хід перетворень на схемах (рис. 8 (а), (б), (в)):

$$I_{ш_Rf} = \frac{\sqrt{4kTR_F B}}{R_F}; U'_{ш_Rf} = \frac{I_{ш_Rf}}{2} * R_F = \frac{U_{ш_Rf}}{2} \quad (9)$$

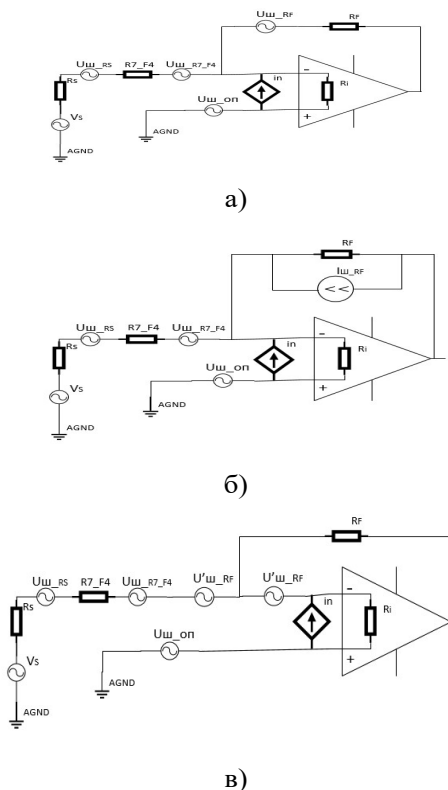


Рис. 8. Приведення до входу всіх джерел шумів зображене без елементів живлення: а) початкове відображення джерел ЕРС шуму елементів; б) перетворення ЕРС джерела шуму на джерело струму в колі ЗЗ; в) перенесення джерела шуму через вузол на вхід підсилювача

Шумові параметри вхідного кола ОП (I_n , R_i , $U_{ш_оп}$) без врахування схеми включення візьмемо з документації на ОП МС33079.

Визначимо шумову смугу підсилювача: оскільки смуга пропускання каналу за технічними вимогами дорівнює 200 кГц, тож шумова смуга є вдвічі меншою. Конденсатори на таких низьких частотах не будуть вносити суттєвих шумів, тому не відображені в схемах та не включені в розрахунок. Вважаємо, що джерела шуму між

собою некорельовані, хоча одне або кілька джерел шуму в підсилювачі можуть взаємодіяти як з v_n , так і з i_n .

Надалі скористаємось властивістю адитивності та визначимо середньоквадратичне значення напруги шуму. За такого підходу у вираз v_{ni} додається доданок $2 \cdot \rho \cdot i_n \cdot v_n \cdot R_s$, де ρ — коефіцієнт кореляції між v_n та i_n , що знаходиться як:

$$\rho = \frac{1}{I_n \cdot V_n} \quad (10)$$

де $\langle v_n i_n \rangle$ — середнє значення v_n та i_n за часом. Тому SNR обчислюється за виразом [8]:

$$SNR = 20 \log \left[\frac{V_s}{V_{ni}} \right] \quad (11)$$

Розробниками задано коефіцієнт підсилення каскаду при $R_{7_F4} = 3.3$ кОм та $R_{3_F4} = R_{2_F4} = 16$ кОм, який дорівнює 10. Отже, намагатимемось досягти його при меншій напрузі шумів, що приведено до входу підсилювача завдяки зміні номіналів елементів кола без врахування шумів вхідних каскадів, що трансформують опір між фотодіодом та формуючим підсилювачем. Таким чином сформована перша умова оптимізації. Обмежимося врахуванням впливу вихідного опору зарядо-чутливого підсилювача, який є джерелом сигналу у обмеженні номіналу вхідного резистора каскаду формуючого підсилювача. Виробник мікросхеми AD829 гарантує досягнення паспортних параметрів при $R_n = 50$ Ом та вище. Це є другою умовою.

Третім кроком складемо функцію залежності напруги шуму каскаду, яку необхідно мінімізувати з джерелами шуму, приведеними до входу підсилювача:

$$SNR = f(A, R_F, R_{7_F4}) \quad (12)$$

Представимо вираз (12) у розгорнутому вигляді (рис. 8(в)):

$$SNR = 20 \lg \left(\frac{V_s}{\sqrt{(i_n \cdot R_s)^2 + (U_{\text{ш}R_{7_F4}})^2 + (U'_{\text{ш}RF})^2 + (U_{\text{ш}RS})^2 + (U_{\text{ш}OP})^2}} \right) \quad (13)$$

Без зміни номіналів елементів каскаду розраховане значення шуму каскаду, приведене до входу, становить 178 нВ/√Гц (55,7 мкВ) та $SNR = 45,1$ дБ.

Приведенням шумів до входу ОП (рис. 8) задача мінімізації (12) зведена до одного базису, тож з'явилася можливість використати традиційний метод оптимізації (13) для нового застосування. У (13) чітко виділена цільова функція коефіцієнта підсилення та обмеження на номінали резисторів.

Четвертим кроком алгоритму проведенням оптимізації на мінімізацію рівня шумів за обмеження в коефіцієнті підсилення та номіналах резисторів здійснимо методами нелінійного програмування комп'ютерним розрахунком [5].

Складемо математичну модель введених обмежень:

$$\begin{cases} \frac{R_F}{R_{7_F4}} \leq 10; \\ R_{7_F4} \geq 50 \Omega \\ R_{7_F4}, R_F \leq 100 \text{ k}\Omega \end{cases} \quad (14)$$

У загальному випадку в методі узагальненого приведенного градієнта розглядається градієнт, або нахил цільової функції. Розраховані часткові похідні прирівнюються до нуля та при зміні вхідних значень отримуємо відповідь: чи перебуваємо ми в точці екстремуму на даній ітерації при заданій збіжності?

Скориставшись цим методом отримаємо наступні результати: $R_F = 485 \text{ Ом}$; $R_{7_F4} = 50 \text{ Ом}$; $\text{SNR} = 63,3 \text{ dB}$. Рівень шуму приведений до входу підсилювача склав $50 \text{ нВ}/\sqrt{\text{Гц}}$, або $15,5 \text{ мкВ}$, що задовольняє умові достатності для достовірної дискретизації сигналу, що наведена нижче.

Для перевірки результатів оптимізації використаємо одне з відомих середовищ моделювання електронних схем, що дозволяє виключити з процесу оптимізації вплив топології, методи виготовлення та монтажу друкованої плати, які складно врахувати при проведенні вимірювань шумів макетуванням пристрою.

Змодельована схема та спектральна щільність шумів у програмі Multsim приведено на рис. 9 та 10, відповідно.

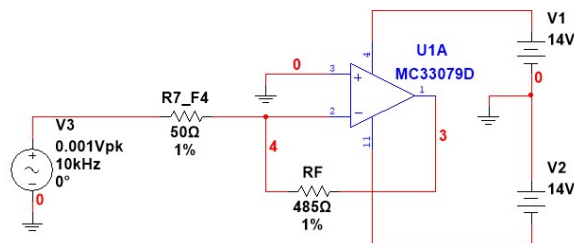


Рис. 9. Схема змодельованого підсилювача

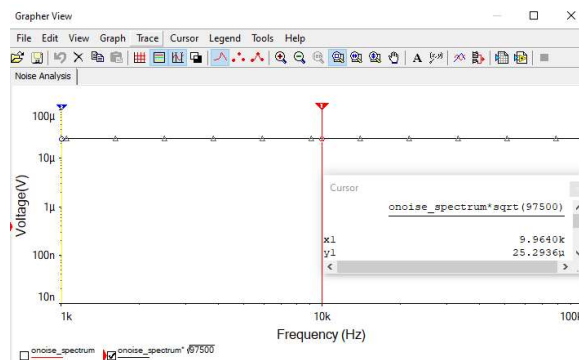


Рис. 10. Модельована спектральна щільність шумів

де чорна лінія є результатом розрахунку в діапазоні частот від 1 до 100 кГц.

Порівняємо отримані значення та встановимо похибку розрахованого значення відносно змодельованого (рис. 10):

$$\Delta = \frac{U_{\text{шумів_розраховане}}}{U_{\text{шумів_змодельоване}}} = 100 - \left(\frac{15,5 \text{ мкВ}}{25,3 \text{ мкВ}} * 100 \right) = 39,5\% \quad (15)$$

Похибка у 39,5% є суттєвою, але збіжність результатів в одному порядку мікрвольт є непоганим результатом. В будь-якому випадку, не існує цілковитої певності, що результат симулятора можна використовувати, як еталонне значення.

З метою усереднення рівня шумів по діапазону робочих частот можливо додати ще один каскад до підсилювача, наприклад ФНЧ. Але таке рішення в рамках дослідження можна віднести до зайвих через витрати часу та коштів на інтеграцію каскаду у модуль АОС та через невеликий вплив на прогнозований рівень шумів в області малих опорів резисторів, які надають можливість зменшити потенційний рівень шумів.

Вимогою забезпечення рівня шумів спектрометричного каналу є 18 мВ, що означає

рівень помилок визначення частинок в окремому діапазоні складає максимум 0,42%. Поглянемо, як проведена оптимізація впливає на достовірність результатів. На рис. 11, 12 приведено обраховані похибки у випадку додання білого Гаусового шуму до згенерованої послідовності (рис. 13) з SNR каскаду до та після оптимізації. Така послідовність складає 25000 відліків умовно розбитих на 15 під діапазонів [2]. Число відліків у кожному під діапазоні не є рівномірним через особливості роботи детекторної системи, продовженням якої є модуль АОС, що оптимізується.

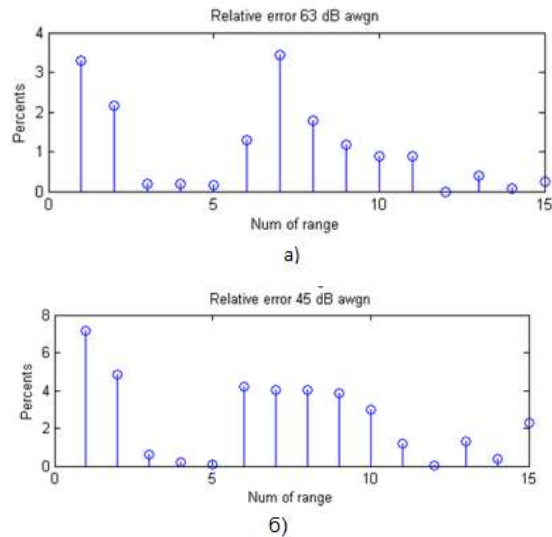


Рис. 11. Відносна прогнозована похибка вимірювань в кожному діапазоні після (а) та до (б) оптимізації.

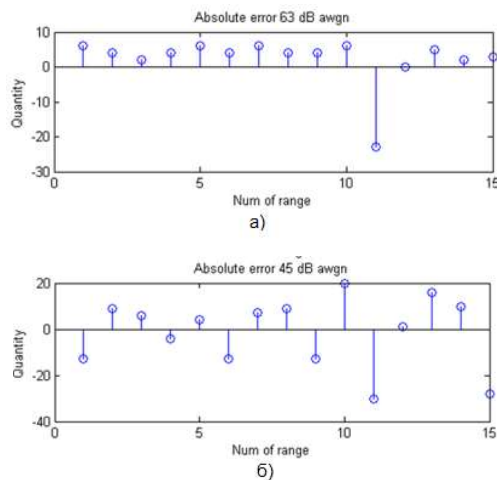


Рис. 12. Абсолютна прогнозована похибка вимірювань в кожному діапазоні після (а) та до (б) оптимізації.

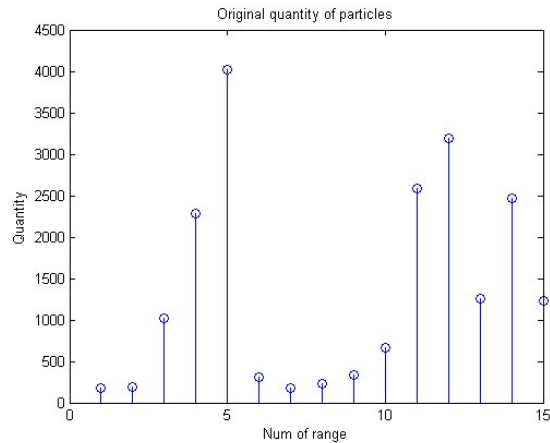


Рис. 13. Початкова кількість заряджених частинок у енергетичних діапазонах без додання шумів

Порівняємо кількість заряджених частинок на границях енергетичних діапазонів, що будуть підраховані, як частинки інших діапазонів до оптимізації (рис. 11 (б), 12 (б)):

$$\delta_{45dB} = \left(\frac{N_{ч.п.45dB}}{N_{\Pi}} \right) * 100\% = \frac{187}{25000} * 100\% = 0,75\% \quad (16)$$

Отриманий результат (16) не задовольняє вимозі за рівнем шумів, тому обчислимо кількість помилково обрахованих частинок після оптимізації (рис. 11а, 12а):

$$\delta_{63dB} = \left(\frac{N_{ч.п.63dB}}{N_{\Pi}} \right) * 100\% = \frac{91}{25000} * 100\% = 0,36\% \quad (17)$$

де $N_{ч.п.45dB}$ — кількість частинок, що підрахована як частинки інших діапазонів енергій при $SNR = 45$ дБ, $N_{ч.п.63dB}$ — кількість частинок, що підрахована як частинки інших діапазонів енергій при $SNR = 63$ дБ, N_{Π} — загальна кількість частинок, що міститься у 15 енергетичних діапазонах.

Отримане значення (17) задовольняє вимозі за рівнем шумів.

Введемо критерій достатності придушення рівня шумів у суміші сигналу та шуму у вигляді типової точності АЦП, що складає $\pm$ величину молодшого значущого розряду. Тобто, якщо рівень шумів менший за половину опорної напруги, то мета оптимізації досягнута.

Вирахуємо максимальну вхідну шумову напругу каскаду за якої можлива достовірна дискретизація сигналу:

$$E_{ш_{вх_{підс}}} \leq \frac{U_{опорне_{АЦП}}}{2 * 2^N * G} = \frac{2,54}{2 * 4096 * 1440} = 0,2 \text{ [мкВ]} \quad (18)$$

Також вирахуємо максимальну вихідну шумову напругу каскаду за якої можлива достовірна дискретизація сигналу:

$$E_{ш_{вих_{підс}}} \leq \frac{U_{опорне_{АЦП}}}{2 * 2^N} = \frac{2,54}{4096 * 2} = 310 \text{ [мкВ]} \quad (19)$$

де $N = 12$ — розрядність АЦП.

Подальші каскади (рис. 4) мають коефіцієнт підсилення 21 та 7, тож перерахуємо рівень

вхідних шумів до виходу підсилювача. Спершу визначимо рівень шумів каскаду до оптимізації:

$$E_{\text{ш}_{\text{вих}_{\text{підс}}}} = E_{\text{ш}_{\text{вх}_{\text{підс}}}} * K_1 * K_2 * K_3 = 55,88 * 10^{-6} * 9,7 * 21 * 7 = 0,04 \text{ [В]} \quad (20)$$

Значення (20) перевищує 18 мВ, що не відповідає вимогам за рівнем шумів. Перерахуємо рівень вхідних шумів, приведених до виходу підсилювача після оптимізації:

$$E_{\text{ш}_{\text{вих}_{\text{підс}}}} = E_{\text{ш}_{\text{вх}_{\text{підс}}}} * K_1 * K_2 * K_3 = 15,88 * 10^{-6} * 9,7 * 21 * 7 = 0,011 \text{ [В]} \quad (21)$$

Отримане значення (21) не перевищує 18 мВ шумів за вимогою. Визначимо точність дискретизації вхідного сигналу такої системи:

$$\frac{U_{\text{опорне}}}{2^N} = \frac{E_{\text{ш}_{\text{вих}_{\text{підс}}}}}{\text{Похибка дискретизації}} \Rightarrow 178 \text{ рівнів квантування} \quad (22)$$

Це означає, що АЦП буде дискретизувати дані з похибкою ± 89 рівнів квантування з 4096.

Висновки

Оптимізацією обраного каскаду підсилювача методом нелінійного програмування вдалося покращити SNR на 18 дБ, при зниженні розрахованого рівня шумів на 11 дБ. Це дозволило вдвічі підвищити достовірність даних у спектрометричному каналі.

Продемонстровано позитивний вплив методики оптимізації шумів, оскільки потенційні шуми, що склали ± 89 рівнів квантування з 4096, призводять до похибки дискретизації $\pm 2,5\%$.

Відносна похибка розрахованого значення шумів оптимізованої схеми, проти змодельованого значення склала 39%, або 10 мкВ.

Метод прогнозування та зменшення рівня шумів каскадів на базі ОП може бути використано при проектуванні кіл, чутливих до рівня шумів, які впливають на результат дискретизації підсиленого слабкого сигналу. Це особливо актуально у вимірювальній апаратурі.

Подальшим розвитком може бути створення подібного методу зменшення рівня шумів для кіл на базі транзисторів.

Література

1. Дудник О., Курбатов Є. Використання наносупутників для вивчення природи мікросплексів високоенергетичних частинок у магнітосфері землі: ідея космічного експерименту. *Космічна наука і технології*. 2018. Т. 24, № 2. С. 36–42. URL: <https://doi.org/10.15407/knit2018.02.036> (дата звернення: 19.06.2023).
2. Yezerskyi N., Movchanuk A., Dudnik O. Electronic simulator of signals derived from scintillation detectors in spectrometric identification of high-energy charged particles. *Radio-technical fields, signals, apparatus and systems* : Abstracts International Conference, Kyiv, 9 November 2021. Kyiv, 2021. P. 75–77.
3. Operational amplifiers for everyone : Design Reference / ed. by R. Mancini. Dallas : Texas Instrument, 2002. 464 p.
4. Leach W. Fundamentals of low-noise analog circuit design. *Proceedings of the IEEE*. 1994. Vol. 82, no. 10. P. 1515–1538. URL: <https://doi.org/10.1109/5.326411> (date of access: 19.06.2023).
5. Семенченко А. Методичний підхід щодо максимізації доходів населення з нестандартними формами зайнятості. *Бізнес інформ*. 2016. С. 137–141. URL: http://repository.hneu.edu.ua/bitstream/123456789/15977/1/Семенченко_стаття%20Ал.%20В..pdf (дата звернення: 19.06.2023).
6. Fidani C., Battiston R., Burger W. A Study of the Correlation between Earthquakes and NOAA Satellite Energetic Particle Bursts. *Remote Sensing in Seismology*. 2010. Vol. 2, no. 9. P. 2170–2184. URL: <https://doi.org/10.3390/rs2092170> (date of access: 19.06.2023).

References

1. Dudnyk O., Kurbatov Ye. Vykorystannia nanosuputnykiv dlia vyvchennia pryrody mikrosplektiv vysokoenerhetychnykh chastynok u mahbitosferi zemli: ideia kosmichnoho eksperymentu. *Kosmichna nauka i tekhnolohii*. 2018. T. 24, № 2. S. 36–42. URL: <https://doi.org/10.15407/knit2018.02.036> (data zvernennia: 19.06.2023).
2. Yezerskyi N., Movchanuk A., Dudnik O. Electronic simulator of signals derived from scintillation detectors in spectrometric identification of high-energy charged particles. *Radio-technical fields, signals, apparatus and systems* : Abstracts International Conference, Kyiv, 9 November 2021. Kyiv, 2021. P. 75–77.
3. Operational amplifiers for everyone : Design Reference / ed. by R. Manchini. Dallas : Texas Instrument, 2002. 464 p.
4. Leach W. Fundamentals of low-noise analog circuit design. *Proceedings of the IEEE*. 1994. Vol. 82, no. 10. P. 1515–1538. URL: <https://doi.org/10.1109/5.326411> (date of access: 19.06.2023).
5. Cemenchenko A. Metodychnyi pidkhid shchodo maksymizatsii dokhodiv naselennia z nestandardnymy formamy zainiatosti. *Biznes inform.* 2016. S. 137–141. URL: http://repository.hneu.edu.ua/bitstream/123456789/15977/1/Семенченко_стаття%20Ал.%20В..pdf (data zvernennia: 19.06.2023).
6. Fidiani C., Battiston R., Burger W. A Study of the Correlation between Earthquakes and NOAA Satellite Energetic Particle Bursts. *Remote Sensing in Seismology*. 2010. Vol. 2, no. 9. P. 2170–2184. URL: <https://doi.org/10.3390/rs2092170> (date of access: 19.06.2023).

КОМП'ЮТЕРНІ ТЕХНОЛОГІЇ

УДК 004.056

МЕТОДИ ПІДВИЩЕННЯ ЯКОСТІ ПЕРЕТВОРЕННЯ МОВИ НА ТЕКСТ В СИСТЕМАХ БІОМЕТРИЧНОЇ АУТЕНТИФІКАЦІЇ

DOI 10.36994 / 2788-5518-2023-01-05-13

Корчинський В.В., д.т.н.,проф. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна. vkadkorchin@ukr.net

Стайкуца С. В., к.ф.н.,доц. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна. s_t_a@ukr.net

Виноградов І.В Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна, ipvinner@gmail.com

Швець О. В. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна. ovshvets@ukr.net

Бєлова Ю. В. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна. bilovaulia@gmail.com

Анотація: У статті розглянуто методи та алгоритми перетворення мови на текст, сучасні відкриті та комерційні системи для створення систем, а також використання цих технологій у сфері кібербезпеки. Пропонується створити систему перетворення мови на текст високої якості. Проведено аналіз математичних алгоритмів, які використовуються для скорочення коефіцієнта помилок, що дає змогу створювати унікальні голосові відбитки та підвищити захист від підроблення. Описана структура сучасних систем перетворення мови на текст. Внаслідок зміни даних сетів, параметрів прихованих Марківських моделей, якісного словника фонем, використання мовних моделей існує можливість зменшити процент помилок при розпізнаванні мови, також використання системи для багатомовності типу "Суржик".

Розглянуто математичні методи оцінки якості системи перетворення мови на текст (WER), а також різні методи розрахунку, що важливо для їхнього подальшого вдосконалення і оптимізації. Розглянуто структуру сучасних систем, а саме попередня обробка сигналу, витягування ознак, акустичне моделювання, моделювання мови, декодування, пост-обробка. Для кожного з етапів запропоновані вектори дослідження, які можуть зменшити коефіцієнт помилок системи в цілому.

Зменшення помилок розпізнавання мови та можливості підробити голос досягається за допомогою різних методів: глибокі нейронні мережі, приховані Марківські моделі, алгоритм Баума-Велша, N-gram моделі, Моделі з увагою, створення якісного словника фонем, датасету, філерів. Технологія перетворення мови в текст може бути використана в системах біометричної аутентифікації для виявлення та аналізу унікальних особливостей голосу користувача. Однак сучасні системи перетворення мови на текст для української, російської та "суржику" потребує вдосконалення, як акустичного, так і мовного блоку. Наукові роботи, які присвячені дослідженню та оптимізації цих систем для біометричної аутентифікації, не повною мірою висвітлюють ці питання. Це стало приводом для подальшого дослідження в цьому напрямку, тому метою цієї роботи є розробка системи розпізнавання мови з мінімальним коефіцієнтом помилок.

Ключові слова: перетворення мови на текст, мовні моделі, моделі прихованого марківського процесу, захист інформації, біометрична аутентифікація.

METHODS OF IMPROVING THE QUALITY OF SPEECH-TO-TEXT CONVERSION IN BIOMETRIC AUTHENTICATION SYSTEMS

Korchynskiy V.V., Dr. habil., Prof. State University of Intellectual Technologies and Communication, Odessa, Ukraine, vkadkorchin@ukr.net

Staikuca S.V., Ph.D., Ass. Prof. State University of Intellectual Technologies and Communication, Odessa, Ukraine, s_t_a@ukr.net

Vynogradov I.V. State University of Intellectual Technologies and Communications, Odessa, Ukraine, khaled@alfaiomi.com

Shvets O.V. State University of Intellectual Technologies and Communication, Odessa, Ukraine, ovshvets@ukr.net

Bielova I.V. State University of Intellectual Technologies and Communications, Odessa, Ukraine, bilovaulia@gmail.com

Abstract: The article discusses the methods and algorithms of speech-to-text conversion, modern open and commercial systems for creating systems, as well as the use of these technologies in the field of cyber security. It is proposed to create a high-quality speech-to-text conversion system. An analysis of the mathematical algorithms used to reduce the error rate, which makes it possible to create unique voice prints and increase protection against forgery, has been carried out. The structure of modern speech-to-text conversion systems is described. By changing datasets, parameters of hidden Markov models, a high-quality dictionary of phonemes, and the use of language models, there is an opportunity to reduce the percentage of errors in language recognition, as well as the use of a system for multilingualism such as "surzhyk".

The mathematical methods of assessing the quality of the system of speech to text (WER), as well as various methods of calculation, which is important for their further improvement and optimization, are considered. The structure of modern systems is considered, namely, signal pre-processing, feature extraction, acoustic modeling, speech modeling, decoding, post-processing. For each of the stages, study vectors have been proposed that can reduce the error rate of the system as a whole.

Reducing speech recognition errors and the ability to fake a voice is achieved using various methods: deep neural networks, hidden Markov models, Baum-Welch algorithm, N-gram models, models with attention, creation of a high-quality phonemes dictionary, dataset, and fillers. Speech-to-text conversion technology can be used in biometric authentication systems to detect and analyze the unique features of the user's voice. However, modern speech-to-text conversion systems for Ukrainian, Russian, and "surzhyk" need improvement in acoustic and language units. Scientific works, which are devoted to research and optimization of these systems for biometric authentication, do not fully cover these issues. This became the reason for further research in this direction, so this work aims to create a speech recognition system with a minimum error rate.

Keywords: speech-to-text conversion, language models, hidden Markov process models, information protection, biometric authentication.

Вступ

Технології перетворення мови на текст є важливим напрямком досліджень у сфері штучного інтелекту і мовознавства. Ці технології відіграють важливу роль у багатьох сферах[2], включаючи автоматичний переклад, синтез мовлення, підтримку осіб з обмеженими можливостями, аналіз тексту, автоматизації центрів обробки дзвінків та багато інших.

Одним з аргументів, які обґрунтовують необхідність дослідження цих технологій, є їх вплив на покращення комунікації між людьми з різними мовними навичками та культурними контекстами. Завдяки перетворенню мови на текст, можливо збільшити доступність інформації для широкого кола людей, незалежно від їхньої рідної мови чи здатності сприймати усну мову.

Дослідження[1,10] в цій області також допомагають вдосконалювати системи автоматичного перекладу, що має велике значення в глобалізованому світі, де спілкування між людьми різних національностей та мов є необхідністю. Застосування технологій перетворення мови на текст в автоматичному перекладі сприяє полегшенню комунікації, розвитку міжнародних бізнес-відносин та сприяє культурному обміну.

Технології перетворення мовлення на текст (Speech to Text) мають великий потенціал і можуть бути використані в Call центрах. Ось декілька способів, наприклад:

1) автоматичне розпізнавання мовлення: це може допомогти операторам Call центру ефективно працювати з великою кількістю дзвінків та забезпечувати точність та швидкість обробки клієнтських запитів;

2) розпізнавання інформації клієнта: ця технологія дозволяє автоматично виділяти ключову інформацію та використовувати її для подальшої обробки або маршрутизації дзвінків до відповідних відділів або операторів;

3) аналіз настрою та емоцій: ця технологія дозволяє оцінити задоволеність клієнта обслуговуванням та вчасно реагувати на негативні ситуації або проблеми, що виникають під час розмови;

4) автоматичне створення запису дзвінка: застосування технологій Speech to Text дозволяє автоматично транскрибувати дзвінки та створювати текстовий запис розмови.

Створення голосових асистентів передбачає застосування технологій перетворення мови на текст, що особливо актуально для певних завдань в галузі кібербезпеки декількома способами:

1) моніторинг спілкування: використовуючи технологію STT, компанії можуть відстежувати аудіо- та відеоконференції для виявлення потенційно шкідливої активності або обговорення

конфіденційної інформації. Це може бути особливо корисним для запобігання витоку інформації;

2) розпізнавання голосових команд: у сучасних системах, які використовують голосове управління, технології STT можуть аналізувати команди, щоб переконатися, що вони не містять потенційно шкідливих інструкцій;

3) біометрична аутентифікація: технології STT також можуть бути використані разом з технологією розпізнавання голосу для аутентифікації користувачів за їх голосом. Це може бути додатковим шаром захисту, на додаток до традиційних методів аутентифікації.

4) аналіз соціальних медіа: технології STT можуть бути використані для перетворення аудіо- та відео контенту в текст, який потім може бути проаналізований з метою виявлення загроз кібербезпеки. Наприклад, обговорення шкідливої активності або витоку конфіденційної інформації.

5) автоматизація реагування на інциденти: STT може бути використаний для автоматичного ведення записів в процесі реагування на інциденти, звільняючи спеціалістів кібербезпеки для важливіших завдань.

Таким чином, метою роботи є розробка системи розпізнавання мови з мінімальним коефіцієнтом помилок.

Виконання досліджень

Слід відзначити, що сучасні системи перетворення мови на текст (Speech to Text) мають певні недоліки, які можуть створювати також загрози інформаційної безпеки користувачів. Ось декілька з них:

1) недостатня точність: навіть найкращі системи перетворення мови на текст можуть містити неточності, особливо якщо розмова відбувається в шумному середовищі або якщо розмовники вимовляють слова невірно;

2) відсутність контексту: наприклад, якщо розмовники використовують слова з множинним значенням або вживають сленгові вирази, система може не зрозуміти, що саме має на увазі розмова;

3) питання безпеки: системи перетворення мови на текст можуть збирати та зберігати конфіденційну інформацію, яка може бути використана для шахрайства або ідентифікації користувачів;

4) обмеженість мов: багатомовні системи перетворення мови на текст можуть бути обмежені у підтримці різних мов та діалектів. Це може ускладнювати комунікацію з користувачами, які говорять іншими мовами або використовують місцеві діалекти. Особливо це актуально для України, де українська мова є державною мовою та є однією з основних мов, якою користуються українці. Російська мова також широко використовується. Крім того, в Україні проживає ряд етнічних груп, які зберігають свої рідні мови та мови-комуніканти, такі як румунська, польська, угорська, білоруська та інші;

5) помилки в інтерпретації мови.

Існуючі пропріетарні та відкриті системи перетворення мови на текст мають певні недоліки та переваги. Пропріетарні системи мають наступні особливості:

1) Google Speech-to-Text API – це сервіс від Google, який використовує передові алгоритми для перетворення мови на текст. Перевагами такого сервісу є висока точність розпізнавання, підтримка багатьох мов, Інтеграція з іншими сервісами Google. До недоліків можна віднести: вартість використання для великого обсягу даних; залежність від інтернет-з'єднання, питання приватності;

2) Nuance Communications Dragon – це один з найпопулярніших пропріетарних продуктів для перетворення мови на текст. До переваг цього сервісу можна віднести високу точність розпізнавання, широкий вибір версій для різних професій, наявність програмного забезпечення для комп'ютерів та мобільних пристроїв. Недоліки сервісу – це висока вартість, обмежена підтримка мов, необхідність установки на локальний пристрій;

Можна відзначити наступні проекти Open source, які характеризують системи перетворення мови на текст:

1) Mozilla DeepSpeech – це проект від Mozilla, який розробляється на базі технологій машинного навчання для перетворення мови на текст. Перевагами такого проекту є відкритий код, можливість працювати на локальному пристрої, наявність пре-тренованих моделей для англійської мови. До недоліків можна віднести: нижча точність розпізнавання, порівняно з комерційними рішеннями, обмежена підтримка мов, можливі труднощі при встановленні та налаштуванні;

2) Kaldi – це відкритий проект для розробки систем розпізнавання мови на основі глибокого навчання. Перевагами проекту є відкритий код, гнучкість та налаштованість, підтримка різних алгоритмів та методів машинного навчання;

3) CMU Sphinx – це один з найбільш відомих open source проектів для перетворення мови на текст, розроблений в Університеті Карнегі-Меллон. Перевагами цього проекту є: відкритий код, що дозволяє налаштовувати систему під конкретні потреби, наявність версій для різних мов та доменів, можливість працювати на локальному пристрої без потреби в постійному інтернет-з'єднанні. До недоліків слід віднести: точність розпізнавання може бути нижчою в порівнянні з комерційними рішеннями, встановлення та налаштування може вимагати певних технічних знань, потреба в додаткових ресурсах для тренування моделей для нових мов або доменів[9].

Таким чином, одним із пунктів дослідження в роботі буде вибір відкритої системи для побудови власного алгоритму перетворення мови на текст. Як критерії ефективності виберемо математичні алгоритми, параметри яких планується удосконалити для отримання кращих результатів перетворення мови на текст, та знання мов програмування в автора.

Для завдання дослідження розглянемо наступні математичні алгоритми розпізнавання мови:

1) Hidden Markov Models (HMMs) може використовуватися для моделювання послідовності звуків у словах або фраз[5];

2) Gaussian Mixture Models (GMMs) використовується для представлення розподілу акустичних ознак в мові. Кожний компонент суміші може моделювати різні види звуків, такі як голосні або приголосні;

3) Viterbi Algorithm використовується для пошуку найбільш ймовірного шляху через HMM, що відповідає даному набору акустичних ознак[6];

4) Mel Frequency Cepstral Coefficients (MFCCs) забезпечує витягування ознак, який використовується для перетворення сигналу мови в набір ознак, які можна використовувати для тренування і тестування моделей розпізнавання мови;

5) N-gram Language Models використовуються для представлення ймовірностей послідовностей слів у мові. Наприклад, у моделі біграм, ймовірність кожного слова залежить від попереднього слова[5];

6) Baum-Welch Algorithm використовується для тренування параметрів HMM.

Таким чином, обґрунтованим є дослідити та за можливості оптимізувати один, або декілька алгоритмів для отримання кращих результатів перетворення мови на текст, а саме зниження WER, про який буде сказано далі.

Структура розпізнавання мови може бути розбита на декілька основних етапів:

1) попередня обробка: на цьому етапі, аудіосигнал, який містить мову, перетворюється в цифрову форму. Це включає в себе відлік сигналу, його фільтрацію та, можливо, зменшення шуму;

2) витягування ознак: одразу після попередньої обробки, аудіо сигнал аналізується з метою витягування корисних ознак, що характеризують мову. Тут використовуються техніки, такі як MFCC (Mel Frequency Cepstral Coefficients), що дозволяють перетворити сигнал в набір числових даних[2];

3) акустичне моделювання: на цьому етапі, використовуються акустичні моделі, зокрема Hidden Markov Models (HMMs) або Deep Neural Networks (DNNs), для визначення відповідності між ознаками аудіосигналу та фонемами, що є найменшими значущими звуковими одиницями мови[5];

4) моделювання мови: цей етап використовує моделі мови (зазвичай на основі n-грамів), які оцінюють ймовірність послідовності слів. Це допомагає вирішувати неоднозначності в розпізнаванні[5];

5) декодування: на цьому етапі, алгоритм вибирає найбільш ймовірну послідовність слів, враховуючи інформацію від акустичної моделі та моделі мови. Для цього часто використовується алгоритм Вітербі[8];

5) пост-обробка: після декодування, можливі деякі корекції або доповнення до тексту, такі як виправлення помилок, пунктуація тощо.

Система розпізнавання мови може бути розбита на два основних блоки: акустико-фонетичний та лінгвістичний.

Акустико-фонетичний блок – це частина системи, яка відповідає за обробку вхідного аудіо сигналу і його перетворення в послідовність фонем або інших звукових одиниць. Він складається з наступних етапів:

1) попередня обробка: на цьому етапі вхідний звуковий сигнал перетворюється в цифрову

форму, яка може бути оброблена комп'ютером. Це може включати відлік, фільтрацію та зменшення шуму;

2) витягування ознак: аудіо сигнал аналізується для витягування корисних ознак, що характеризують мову. Зазвичай використовується метод Mel Frequency Cepstral Coefficients (MFCC) для перетворення аудіо сигналу в набір числових даних;

3) акустичне моделювання: використовуються акустичні моделі, такі як Hidden Markov Models (HMMs) або Deep Neural Networks (DNNs), щоб визначити відповідність між ознаками аудіосигналу та фонемами[5].

Лінгвістичний блок – це частина системи, яка відповідає за моделювання структури мови та генерацію тексту. Він включає наступні етапи:

1) моделювання мови: на цьому етапі використовуються моделі мови (часто на основі n-грамів), щоб визначити ймовірність послідовності слів;

2) декодування: декодер використовує інформацію від акустичної моделі та моделі мови, щоб визначити найбільш ймовірну послідовність слів.

Розглянемо умову забезпечення якісного спектра сигналу для обробки та перетворення його в текст. Математичні моделі перетворення акустичного сигналу не є частиною цього дослідження, але оскільки вірогідно датасети будуть записані в умовах без шумів, треба дослідити можливості зменшення шуму.

Глибокі нейронні мережі (Deep Neural Networks, DNNs) можуть бути ефективними інструментами для зменшення шуму в акустичних сигналах, включаючи ті, які використовуються в системах автоматичного розпізнавання мови (ASR). Ці системи зазвичай можуть виправити або відфільтрувати шум в сигналі перед тим, як сигнал передається на більш високий рівень ASR для визначення фонем або слів[7].

Існує певні способи, за якими DNN можуть використовуватися для зменшення шуму:

1) Speech Enhancement: моделі DNN можуть використовуватися для визначення характеристик шуму в сигналі і зменшення його впливу. Наприклад, є така техніка як "Deep Noise Suppression", яка використовує глибокі нейронні мережі для визначення і видалення шуму з аудіосигналу;

2) Robust Acoustic Modeling: DNN також можуть бути натреновані для розпізнавання фонем в сигналах, які містять шум. Такі моделі, називаються стійкими акустичними моделями, намагаються визначити корисну інформацію в акустичному сигналі, незважаючи на наявність шуму;

3) Feature Compensation: корекція або компенсація ознак, які були спотворені шумом. Це включає використання DNN для виправлення або видалення шумових складових в ознаках акустичного сигналу.

Пошук, або створення якісного датасету для створення моделі розпізнавання мови є одним зі складних завдань цієї роботи. Нам знадобиться великий набір даних, який містить широкий спектр дикторів, акцентів, фраз та умов запису. Особливо це стосується суржиків та інших змішаних варіацій мови. Щоб створити модель розпізнавання мови для суржиків необхідно мати спеціалізований датасет, що містить аудіозаписи людей, які говорять цим діалектом. Такий датасет повинен включати велику кількість записів з різноманітними дикторами, контекстами, акцентами та швидкостями мовлення. Крім того, потрібні відповідні транскрипції цих записів, щоб модель могла вивчити відповідність між аудіосигналами та текстом. Певна робота вже була зроблена проектом UA Silero, але на відміну від англійської чи російської мов, для української та суржиків ще не існує якісних датасетів, які можна було б використовувати без доопрацювання. Також потрібно дослідити датасети, які вже є в open source, наприклад Common Voice від Mozilla або Voxforge і чи достатньо цього буде для завдання дослідження. Датасети відіграють важливу роль у створенні ефективних моделей розпізнавання мови з наступних причин та завдань:

1) навчання моделі: датасети використовуються для навчання моделей розпізнавання мови, які вивчають взаємозв'язки між аудіосигналами та відповідними транскрипціями тексту. Моделі використовують ці відповідності, щоб "вивчити" як перетворити аудіосигнали на текст;

2) різноманітність даних: ефективні моделі розпізнавання мови мають бути здатні розпізнавати мову в різних контекстах. Це означає, що вони повинні бути навчені на датасетах, які включають аудіо від різних дикторів з різними акцентами, голосами, швидкістю мовлення та іншими факторами;

3) тестування та валідація: датасети також використовуються для тестування та валідації моделей під час та після навчання. Це допомагає забезпечити, що модель працює правильно та здатна ефективно розпізнавати мову в реальних умовах.

Після отримання спектра, який простіше використовувати для системи розпізнавання мови потрібно буде знайти фонему, філери та транскрипції. Для цього потрібно застосування словника фонем, словника філерів та словника транскрипцій:

1) словник фонем – це набір усіх фонем, що використовуються в певній мові. Фонема - це найменша одиниця звуку, яка може змінювати значення слова в мові. Одним із напрямків роботи буде спроба знайти або створити словник фонем для двох або більше мов, та спробувати використовувати його для побудови багатомовної системи перетворення мови на текст;

2) філери в системах Speech to Text (розпізнавання голосу) – це слова або звуки, які люди використовують у своєму повсякденному мовленні, але які не мають особливого значення для змісту висловлювання. Це слова, як "емм", "ну", "типу", "такий" і так далі. Використання філерів може підвищити якість розпізнавання на 10-15%;

3) словник транскрипцій – це ресурс, який використовується в процесі розпізнавання мови, щоб вказати, як слова мають бути вимовляються. Він зазвичай містить список слів разом з їхніми фонетичними транскрипціями.

Приховані марковські моделі (ПММ або НММ) вже тривалий час використовуються в системах розпізнавання мови[5] у розв'язанні ключових проблем статистичного розпізнавання, включаючи розпізнавання вимовленого слова, розпізнавання мови та розпізнавання диктора, а саме:

1) статистичний аналіз: ПММ дозволяють проводити статистичний аналіз розподілу ймовірностей для послідовностей слів, що допомагає у розпізнаванні мови.

2) моделювання часу: ПММ відображають часові залежності в даних. Це дозволяє їм моделювати послідовність звуків в словах, що є ключовим фактором в розпізнаванні мови.

3) розпізнавання слів: ПММ можуть бути навчені представляти конкретні слова або набори слів, що дозволяє системі розпізнавання мови ідентифікувати та класифікувати вимовлені слова.

4) розпізнавання диктора: ПММ можуть бути використані для ідентифікації особливостей голосу диктора, що поліпшує здатність системи розпізнавати та інтерпретувати мову різних людей.

Оптимізація НММ включає в себе ряд проблем, які вивчаються сучасними науковцями. Можна відзначити наступні проблеми, вирішення яких може удосконалити існуючі моделі перетворення мови на текст:

1) покращення алгоритму вивчення: в НММ використовується алгоритм Баума-Велша для вивчення параметрів моделі. Цей алгоритм є емпіричним, і сучасні науковці постійно шукають способи його покращення, щоб він був швидшим і точнішим;

2) оптимізація структури моделі: НММ може мати різні структури, залежно від контексту проблеми. Вибір найкращої структури моделі є однією з ключових проблем оптимізації НММ;

3) спрощення моделі: у багатьох ситуаціях, використання більш простих моделей може поліпшити швидкість та ефективність обчислень. Тому науковці досліджують способи спрощення моделей без значного зниження їхньої точності;

2) масштабування моделей: при використанні великих датасетів, може виникнути необхідність в масштабуванні моделей, що є іншою важливою проблемою оптимізації;

3) інтеграція з іншими моделями: інтеграція НММ з іншими статистичними або машинними моделями є ще однією областю дослідження;

3) універсальні моделі: розробка універсальних прихованих марковських моделей, які можуть бути ефективно використаними для перетворення;

4) мовні моделі є основними компонентами багатьох сучасних систем обробки природної мови.

За допомогою цих моделей вивчається структура тексту та генеруються ймовірні прогнози наступних слів або речень. Існує декілька типів мовних моделей:

1) N-gram моделі – це найпростіші мовні моделі, які використовують статистику з текстових даних для прогнозування ймовірності наступного слова на основі n-1 попередніх слів. Вони доволі прості у використанні, але мають обмеження, що стосуються віддалених залежностей в тексті та розріджених даних[5].

2) моделі прихованого марковського процесу (НММ) – ці моделі можуть бути використані для моделювання послідовностей слів. Такі моделі здатні обробляти віддалені залежності, але їхній алгоритм навчання може бути складним;

3) рекурентні нейронні мережі (RNN): RNN можуть обробляти довільну довжину послідовностей, що робить їх дуже корисними для мовних моделей. Однак, вони мають проблеми з віддаленими залежностями через проблему затухання градієнта;

3) моделі з увагою, як Transformer – засновані на механізмі "уваги", тобто ці моделі дозволяють моделювати залежності між словами, незалежно від їх відстані одне від одного в тексті[7];

4) моделі на основі BERT: BERT (Bidirectional Encoder Representations from Transformers) використовує Transformer архітектуру для створення потужних мовних моделей, які можуть моделювати контекст обох напрямків;

5) GPT (Generative Pre-training Transformer) – це серія моделей, розроблених OpenAI, що використовуються для розуміння і генерації природної мови.

Розглянемо можливості дослідження та оптимізації мовних моделей на основі N -gram. Ця модель є одна з найпростіших та найпоширеніших, в якій використовується машинне навчання для розпізнавання мови. При цьому відбувається поділ тексту на послідовність N слів та використання їх з метою вивчення ймовірностей мови.

Основні дослідницькі зусилля для оптимізації N -грам моделей включають:

1) згладжування: проблема з N -грамами полягає в тому, що деякі послідовності слів можуть не з'явитися в тренувальному наборі даних, що призводить до нульової ймовірності. Згладжування допомагає вирішити цю проблему, шляхом призначення маленького значення ймовірності для таких послідовностей[5];

2) Backoff та interpolation: ці методи є способами управління ситуаціями, коли в тренувальних даних не вистачає інформації для оцінки ймовірності певного N -граму;

3) використання контексту: покращення способу, яким N -грами використовують контекст, може підвищити їхню точність. Замість того, щоб тільки використовувати безпосередній контекст $N-1$ попередніх слів, деякі дослідники дивляться на більш широкий контекст;

4) покращення обчислювальної ефективності: N -грам моделі можуть бути великими та складними, особливо для великих значень N . Дослідники шукають способи зменшення розміру моделі або покращення швидкості її обчислень[5];

5) використання глибокого навчання: нейронні мережі також можуть бути використані для побудови мовних моделей на основі N -грамів, здатних моделювати складніші залежності в даних;

6) комбінування з іншими методами: N -грам моделі часто використовуються в комбінації з іншими методами машинного навчання для створення більш потужних систем розпізнавання мови.

WER, або Word Error Rate, є загальноприйнятою мірою якості в задачах перетворення мови на текст, таких як автоматичне розпізнавання мови (ASR). WER вимірює кількість помилок, зроблених системою під час перетворення мовлення в текст, в порівнянні з референтним текстом.

WER розраховується на основі кількості вставок (I), видалень (D) і заміन (S) необхідних для того, щоб перетворити гіпотезу (текст, згенерований системою ASR) в референсний текст. Він вимірюється за формулою:

$$WER = (S + D + I) / N,$$

де N – це кількість слів в референтному тексті.

Чим нижче WER, тим точніше система перетворення мови на текст. Проте WER також має свої обмеження, наприклад, він не враховує семантичний контекст або порядок слів, що можуть бути важливими в деяких застосуваннях.

Треба визначити кількість вставок, видалень та модифікацій, які необхідні для того, щоб розпізнана послідовність слів співпадала з еталонною послідовністю слів у текстовому корпусі. Цей процес обраховується на основі відстані Левенштейна.

Передові системи перетворення мови на тексти можуть досягати WER близько 5-10% на чистих, добре структурованих даних, наприклад, новинні відгуки або транскрипції подкастів. Однак в умовах реального світу, зі шумом у фоновому середовищі, акцентами, неформальною мовою, ймовірність помилки може збільшитись. Мета дослідження - створення системи перетворення української мови(опціонально російської та "суржик") яка буде мати якомога менший WER.

Для оцінки треба використовувати перехресну валідацію. Кросвалідація, також відома як перехресна перевірка, це метод для оцінки якості моделі машинного навчання, що включає розділення даних на декілька підмножин та тестування моделі на одній підмножині після тренування на інших. В контексті автоматичного розпізнавання мови (ASR), кросвалідація може бути використана для оцінки того, наскільки добре модель може адаптувати своє навчання до нових даних. Наприклад, можна поділити набір даних для тренування на 10 підмножин (в методі,

що відомий як 10-кратна кросвалідація), потім навчати модель на 9 з них і тестувати її на 10-й. Це процес може бути повторений 10 разів, так що кожна підмножина використовується для тестування один раз.

Висновки

- 1) Хоча системи перетворення мови на текст вже досить ефективні, вони все ще мають ряд обмежень. Потрібно продовжувати дослідження і розробку, щоб поліпшити точність, ефективність і адаптивність цих систем.
- 2) Запропонований покроковий алгоритм створення системи speech to text, а також можливі шляхи покращення математичних, лінгвістичних та інженерних моделей для зменшення помилок перетворення мови на текст.
- 3) На першому етапі пропонується створити систему перетворення української мови з тих датасетів, які є у відкритому доступі. Збір даних - це не одноразовий процес. Для оптимальної роботи системи перетворення мови на текст датасет повинен бути постійно оновлюваний та покращуваний з урахуванням змін та покращення роботи системи.
- 4) На другому етапі пропонується оптимізувати математичні алгоритми, які були оглядово розглянуті в цій статті для підвищення ймовірності визначення фону.
- 5) На третьому етапі пропонується зі спеціалістами-філологами вдосконалити або створити якісні словники транскрипцій та словники фону.
- 6) Отримані результати будуть проаналізовані і використані для створення багатомовної системи українська, суржик.

Література

1. A.Gaydhani, V.Doma, S.Kendre, L.Bhagwat Виявлення зневажливих та образливих слів у Twitter за допомогою машинного навчання: підхід на основі n-грам та статистичного показника (tf-idf). URL: <https://arxiv.org/pdf/1809.08651.pdf>
2. M.Suzuki, N.Itoh, T.Nagano, G.Kurata Удосконалення моделі мови n-грам з використанням тексту, згенерованого з моделі нейронної мови. URL: <https://ieeexplore.ieee.org/document/8683481>
3. R.Shadiev, WY.Hwang, NS.Chen, YM.Huang Огляд технології розпізнавання мовлення в текст для покращення навчання. URL: https://www.researchgate.net/publication/267811277_Review_of_Speech-to-Text_Recognition_Technology_for_Enhancing_Learning
4. DA.Jones, F.Wolf, E.Gibson, E.Williams, E.Fedorenko Вимірювання читабельності автоматичних розшифровок мовлення в текст. URL: https://www.researchgate.net/publication/221489176_Measuring_the_readability_of_automatic_speech-to-text_transcripts
5. BH.Juang, LR.Rabiner Приховані марковські моделі для розпізнавання мовлення. URL: <https://www.jstor.org/stable/1268779>
6. J.Picone - журнал IEEE Assp Безперервне розпізнавання мови з використанням прихованих марковських моделей. URL: <https://ieeexplore.ieee.org/document/54527>
7. Підходи до глибокого навчання в системі синтезу мовлення: систематичний огляд і перспектива останніх досліджень. URL: https://www.researchgate.net/publication/363982582_A_deep_learning_approaches_in_text-to-speech_system_a_systematic_review_and_recent_research_perspective
8. Y.Kumar, A.Koul, C.Singh - Мультимедійні засоби та програми, 2023 – Springer. URL: <https://link.springer.com/article/10.1007/s11042-022-13943-4>
9. Джерела та документи Sphinx. URL: <https://www.sphinx-doc.org/en/master/>
10. Lawrence Rabiner, Biing-Hwang Juang Основи розпізнавання мови. URL: https://www.academia.edu/4924307/Fundamental_of_Speech_Recognition_Lawrence_Rabiner_Biing_Hwang_Juang

References

1. A Gaydhani, V Doma, S Kendre, L Bhagwat Detecting hate speech and offensive language on twitter using machine learning: An n-gram and tfidf based approach. URL: <https://arxiv.org/pdf/1809.08651.pdf>
2. M Suzuki, N Itoh, T Nagano, G Kurata Improvements to n-gram language model using text generated from neural language model. URL: <https://ieeexplore.ieee.org/document/8683481>
3. R Shadiev, WY Hwang, NS Chen, YM Huang Review of speech-to-text recognition technology for enhancing learning. URL: https://www.researchgate.net/publication/267811277_Review_of_Speech-to-Text_Recognition_Technology_for_Enhancing_Learning
4. DA Jones, F Wolf, E Gibson, E Williams, E Fedorenko Measuring the readability of automatic speech-to-text transcripts. URL: https://www.researchgate.net/publication/221489176_Measuring_the_readability_of_automatic_speech-to-text_transcripts
5. BH Juang, LR Rabiner Hidden Markov models for speech recognition. URL: <https://www.jstor.org/stable/1268779>
6. J Picone - IEEE Assp magazine Continuous speech recognition using hidden Markov models. URL: <https://ieeexplore.ieee.org/document/54527>

7. A deep learning approaches in text-to-speech system: A systematic review and recent research perspective. URL: https://www.researchgate.net/publication/363982582_A_deep_learning_approaches_in_text-to-speech_system_a_systematic_review_and_recent_research_perspective
8. Y Kumar, A Koul, C Singh - Multimedia Tools and Applications, 2023 – Springer. URL: <https://link.springer.com/article/10.1007/s11042-022-13943-4>
9. Sphinx sources and docs. URL: <https://www.sphinx-doc.org/en/master/>
10. Lawrence Rabiner, Biing-Hwang Juang Fundamentals of Speech Recognition. URL: https://www.academia.edu/4924307/Fundamental_of_Speech_Recognition_Lawrence_Rabiner_Biing_Hwang_Juang_

УДК 004.43:004.02

ОПИС ЗНАНЬ ПРО ВИХІДНИЙ КОД З ВИКОРИСТАННЯМ ОНТОЛОГІЇ

DOI 10.36994 / 2788-5518-2023-01-05-14

Ткачук А.В., аспірант, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна, <https://orcid.org/0000-0002-9127-6381>, andrewtkachuk@yahoo.com

Булах Б.В., к.т.н., доц., Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна. <https://orcid.org/0000-0001-5880-6101>, bogdan.bulakh@gmail.com

Анотація: Дане дослідження зосереджене на використанні онтології як формальної баз знань для представлення вихідного коду в автоматизованих системах. Шляхом створення онтології, яка відображає концепції, зв'язки та властивості елементів вихідного коду, стає можливим зберігання та організація знань про кодову базу. Це надає різноманітні функціональні можливості, такі як автоматизований аналіз коду та створення знань про структуру та поведінку коду. Метою цього дослідження є перевірка можливості використання онтології як формальної бази знань про вихідний код в автоматизованих системах. Огляд останніх статей показав, що існуючі інструменти для роботи з онтологіями не надають безпосередньо функціональності для перетворення коду в знання, тому ідея, запропонована в статті, має новизну. Стаття досліджує використання мови OWL як основної мови представлення онтології через її виразність, підтримку механізмів логічного виведення та сумісність зі стандартами семантичної мережі. Пропонується створення представлень онтології для різних елементів коду, з класами, що представляють основні типи сутностей, та окремими екземплярами, що представляють конкретні екземпляри всередині кодової бази. Для перевірки концепції та демонстрації очікуваних результатів будується система-прототип, яка складається з клієнта для розбору вихідного коду, RESTful API для зв'язку між компонентами та менеджера онтологій. Система використовує такі бібліотеки, як SwiftSyntax для розбору Swift-коду, Alamofire для HTTP-зв'язку та бібліотеки маніпулювання OWL (наприклад, OWL-API та ONT-API) для роботи з онтологіями. Система має на меті вилучення знань з вихідного коду, перетворення його на сутності та аксіоми онтології та обробку онтології за допомогою системи отримання знань (наприклад, HermiT). Зроблено висновок, що онтології ефективно зберігають знання про вихідний код та дозволяють різноманітні маніпуляції над ними. Однак, для реалізації збору більшого обсягу знань про вихідний код необхідно провести подальше дослідження для визначення оптимального набору властивостей та сутностей, які має підтримувати система перетворення коду в онтологію.

Ключові слова: онтологія; OWL; база знань; система перетворення коду в онтологію; REST; семантичний веб; reasoner.

DESCRIBING THE KNOWLEDGE ABOUT THE SOURCE CODE USING AN ONTOLOGY

Andrii Tkachuk, Ph.D. student, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine. <https://orcid.org/0000-0002-9127-6381>, andrewtkachuk@yahoo.com

Bogdan Bulakh, Ph.D., Ass. Prof., National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine. <https://orcid.org/0000-0001-5880-6101>, bogdan.bulakh@gmail.com

Abstract: This research focuses on the use of ontologies as formal knowledge bases for representing source code in automated systems. By creating an ontology that represents the concepts, relationships, and properties of source code elements, it becomes possible to capture and organize knowledge about the codebase. This enables various functional abilities such as automated code analysis and reasoning about code structure and behavior. The aim of this research is to verify the ability to use ontology as a formal knowledge base about source code in automated systems. A review of the recent articles showed that existing tools to work with ontologies does not provide direct functionality to convert the code into a knowledge, so the idea proposed in the article has novelty. The research explores the use of the Web Ontology Language (OWL) as the primary ontology representation language due to its expressiveness, reasoning capabilities, and compatibility with semantic web standards. The research proposes the creation of ontology representations for various code entities, with classes representing main entity types and individuals representing specific instances within the

codebase. To validate the concept and demonstrate the expected results, a proof-of-concept system is built, consisting of a parsing client, a RESTful API for communication, and an ontology manager. The system utilizes libraries such as SwiftSyntax for parsing Swift code, Alamofire for HTTP communication, and OWL manipulation libraries (such as OWL-API and ONT-API) for working with ontologies. The system aims to extract knowledge from the source code, convert it into ontology entities and axioms, and reason over the ontology using a reasoner (such as HermiT). The research concludes that ontologies can effectively store knowledge about source code and enable various further manipulations on that knowledge. However, further investigation is needed to determine the optimal set of properties and entities that should be supported by the code ontology extraction system to infer more knowledge about the source code.

Keywords: ontology; OWL; knowledge base; code ontology extraction; REST; semantic web; reasoner.

Introduction

An ontology is a formal representation of knowledge that defines concepts, their properties, and the relationships between them within a particular domain. It provides a structured framework for organizing and categorizing information, allowing for better understanding and communication between humans and machines. Ontologies are commonly used in information science and knowledge management to enable efficient data integration, reasoning, and semantic interoperability.

By creating an ontology that represents the concepts, relationships, and properties of source code elements such as classes, methods, variables, and their interactions, it becomes possible to capture and organize knowledge about the codebase. This can enable various useful features such as an automated code analysis, code documentation generation, semantic searching, and reasoning about code structure and behavior. Ontologies can help developers and software engineers better understand and manipulate source code by providing a structured representation of its semantics [1].

Advanced automated refactoring and source code analysis systems may benefit from properties of a knowledge base which are backed by the ontology. They are able to store references between code entities, strictly define their types, query the base, and perform any kinds of data reasoning. They may help to add or rewrite source code with insurance that it will compile (in contrast with AI systems with unspecified or fuzzy production rules) [2].

Problem definition. Using ontologies to store the knowledge about the source code can be complex to create and maintain, requiring domain expertise and ongoing effort. Additionally, capturing the full scope of a codebase and managing scalability can be challenging, making ontology-based approaches less suitable for large and diverse codebases.

The aim of this research is to investigate the possibility to use ontology as a formal knowledge base about source code in automated systems.

Recent research articles review. Ontologies can be represented using different formal languages, with the most commonly used one being the Web Ontology Language (OWL). OWL provides a standardized and expressive syntax to define concepts, properties, relationships, and axioms within an ontology. Other representations include RDF (Resource Description Framework), which allows for the creation of semantic triples, and graphical notations like UML (Unified Modeling Language) for visualizing ontologies [3].

OWL is considered better for representing ontologies due to its higher expressiveness, semantic web standards compliance, reasoning capabilities, and a well-developed tooling ecosystem, enabling precise modeling and interoperability. However, the choice of ontology representation language ultimately depends on the specific requirements and complexity of the domain being modeled [4; 5].

Taking into consideration all aforementioned facts, OWL was selected as a main tool to create a knowledge base about the source code.

OWL (Web Ontology Language) consists of classes for representing concepts, properties for defining relationships, and individuals to represent specific instances within a domain. It also includes axioms to express logical statements, such as subclass relationships and property characteristics, as well as logical constructs and reasoning capabilities for inferring new knowledge and checking

ontology consistency. Together, these components form a comprehensive framework for creating and utilizing ontologies to represent knowledge in a structured and semantically rich manner.

To verify the concept and prove that a knowledge base powered by an ontology will give some expected result, the following representations may be created: large group types (like classes, structures, enumerations, functions, etc.) may be presented by OWL classes, each entity in the source code may be represented as named individual belonging to its respective class, properties of entities may be expressed by data properties, and relations between entities may be expressed object properties. Also, to fulfill all the gaps, to build truly semantic representation of the codebase, and to enable reasoning, a large set of axioms may be used (including but not limited to subclass, equivalent class, disjoint class, data property, object property axioms).

To be able to get the missing axioms in ontology, to verify consistency with code smells (if any defined in ontology), and to make amends in codebase respectively, a reasoner should be used. A reasoner is a software tool or component that performs automated logical reasoning over an ontology. It utilizes the defined axioms, rules, and logical constructs within the ontology to derive new knowledge, check consistency, and answer queries. Reasoners can infer implicit relationships, identify inconsistencies, validate data against the ontology, and provide logical conclusions based on the ontological knowledge. For the purposes of this research, HermiT reasoner has been selected [6; 7].

Preliminary research results show that there are solutions available that aim to convert a codebase into an ontology or utilize ontologies to represent code-related knowledge. These solutions typically employ techniques from software analysis, natural language processing, and ontology engineering. Some examples include Code Ontology Extraction, Reverse Engineering Tools, Ontology Mapping and Integration, etc.

It's important to note that while these solutions exist, the process of converting a codebase into an ontology can be challenging, and the resulting ontology may require manual refinement and maintenance. The complexity and nuances of code often make it difficult to capture all the relevant information accurately. Additionally, the choice of specific tools or techniques may vary depending on the programming language, codebase size, and desired ontology representation.

Further research on possible applications of those tools (JOIE, SemTK (over Jena), SourcererCC) revealed that even though they provide some possibilities to easily manage RDF triplets or OWL ontologies, they do not provide API or any other ability to convert codebase into ontology. Such conclusion testifies that the aim of current research (to use the ontology as knowledge base about the source code) has some novelty.

Code ontology extraction is the process of automatically extracting knowledge from source code and representing it in the form of an ontology. It involves analyzing the codebase to identify various code elements such as classes, methods, variables, and their relationships. The goal is to capture the structure, semantics, and interdependencies within the codebase and transform them into a structured representation that adheres to ontology concepts, properties, and relationships.

Code ontology extraction typically involves techniques from software analysis, parsing, and semantic understanding. These techniques may include static analysis, abstract syntax tree (AST) traversal, type inference, and pattern recognition. By analyzing the codebase, extracting relevant information, and mapping it to ontology constructs, code ontology extraction aims to provide a formal representation of the code's structure and semantics.

The extracted code ontology can enable various applications such as code search, code reuse, documentation generation, and automated reasoning about code behavior. It enhances the understanding and utilization of the codebase by providing a structured and semantic representation of its elements and relationships.

In this research we are concentrating at code ontology extraction process and aiming at implementing the model which may be used by automated tools for refactoring and analysis. At first, the idea was to use the same programming language (which is used to write source code) to

represent the knowledge base [8; 9]. It will give the possibility to quickly interpret source code and convert it to OWL entities and axioms.

For such reason, several frameworks and tools were considered. As the main research language is Swift, only libraries with such language supported were taken into consideration. One of them is OWL-API for iOS [10] with respective reasoner – Mini-ME Swift [11]. Their authors state that there is a lack of semantic reasoners and OWL manipulation libraries for Swift which makes it reasonable to develop one. Porting existing solutions or use of multiplatform libraries may be hard due to architectural differences.

Unfortunately, provided frameworks and libraries are on their initial stage of development. They do not support ontology creation (only parsing) and specific OWL entities (such data properties) are still not available.

Also, using the same language for parsing client and knowledge base makes these two instances tightly coupled which may lead to unnecessary rework when one of them changes (Fig. 1).

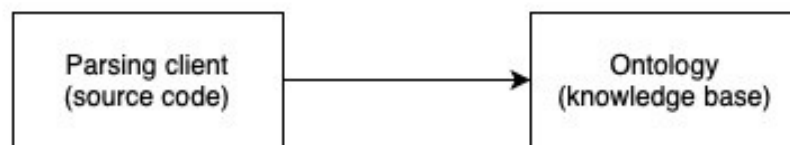


Fig. 1. Dependency between client and ontology

In order to break the coupling, an abstraction should be used (according to Dependency Inversion principle). It empowers creation of flexible and maintainable code by inverting the dependency direction, where higher-level modules depend on abstractions rather than concrete implementations, enabling easier substitution and extensibility of components. Data Transfer Object (DTO) may be used as intermediary in this case. DTO agreement may be established and DTO models to be built on both sides – client and ontology handler (Fig. 2). This will help to uncouple the instances and provide the ability to integrate abstract ontology with different parsing clients (several programming languages) or ontology management systems.

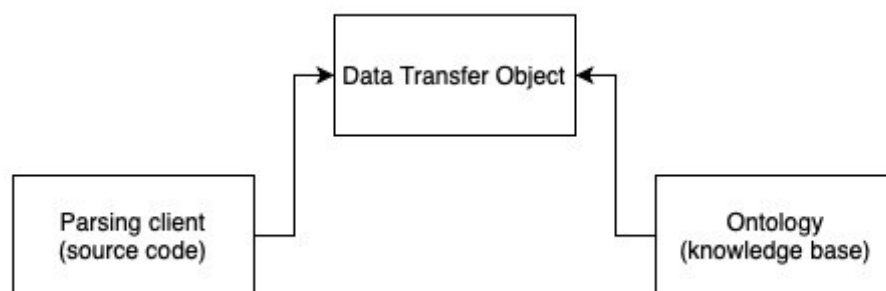


Fig. 2. Breaking dependency with DTO

It is worth saying that implementing our own library for OWL ontology manipulation is not reasonable. Firstly, it will require significant efforts and resources to build and maintain, though writing such library is not the primary aim of the research. Secondly, developers and designers should possess distinctive knowledge about those technologies. Lately, existing OWL libraries and frameworks provide well-tested and optimized solutions with extensive tooling support, which can save development effort and ensure compatibility with established standards and practices.

In order to support all the required traits of knowledge base generated by code ontology extraction system (and to build such system), specific libraries should be taken into consideration.

OWL-API [12] and ONT-API [13] are the two worth mentioning. They are Java libraries to work with ontologies. OWL-API provides support of OWL2 ontologies (ability to create, update, serialize/deserialize them). In contrast, ONT-API (which is built on top of OWL-API) provides support for RDF and is built over Apache Jena. These libraries have all the necessary ontology properties supported to be able to be used in code ontology extraction system.

Research Details

Taking into consideration all the finding from latest research and articles, it is worth building code ontology extraction system, which is language independent, uses abstractions, utilizes OWL manipulation libraries with sufficient functionality, easily extensible and able to be integrated with other systems (ontology or codebase manipulating ones).

In order to build a proof of concept and show that proposed idea makes sense, it is decided to build:

- the parsing client in Swift language which will parse the Swift codebase;
- RESTful API which will accept JSON DTOs of code entities [14];
- OWL ontology manager, which will handle all the interactions with ontology.

The principal component diagram is displayed on fig. 3. Due to the architecture, there may be several clients connecting to the Ontology service and Ontology service itself may work with different ontology managers or tools for ontology reasoning or visualization.

Before work with ontology can be started, it is worth designing its structure.

On Fig. 4 a structure of ontology is depicted. There were three classes created: *Class*, *Function*, *Property*. All the entities from a codebase which fit the type of class will be transformed into named individuals which belong to respective ontology class. Those instances will be linked with each other by the following object properties: "*hasMethod*", "*hasProperty*", "*hasArgument*". In order to support bidirectional links, two inverse object properties counterparts should be created: "*isInClass*", "*isInMethod*".

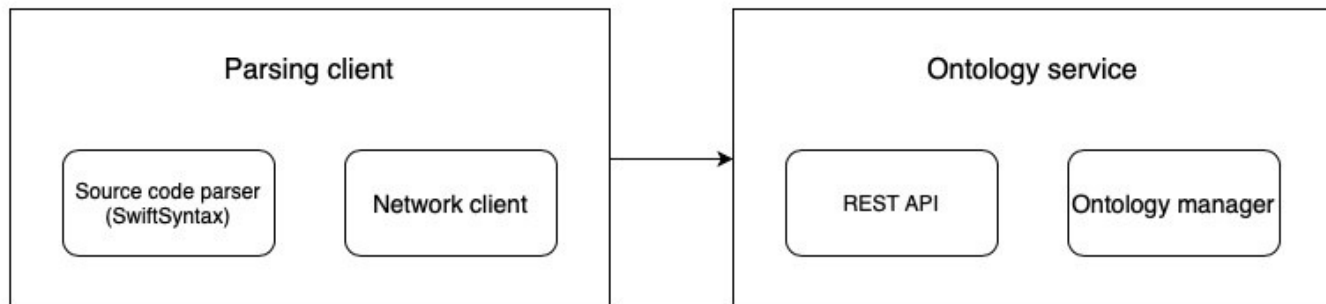


Fig. 3. Principal diagram of code ontology extraction system

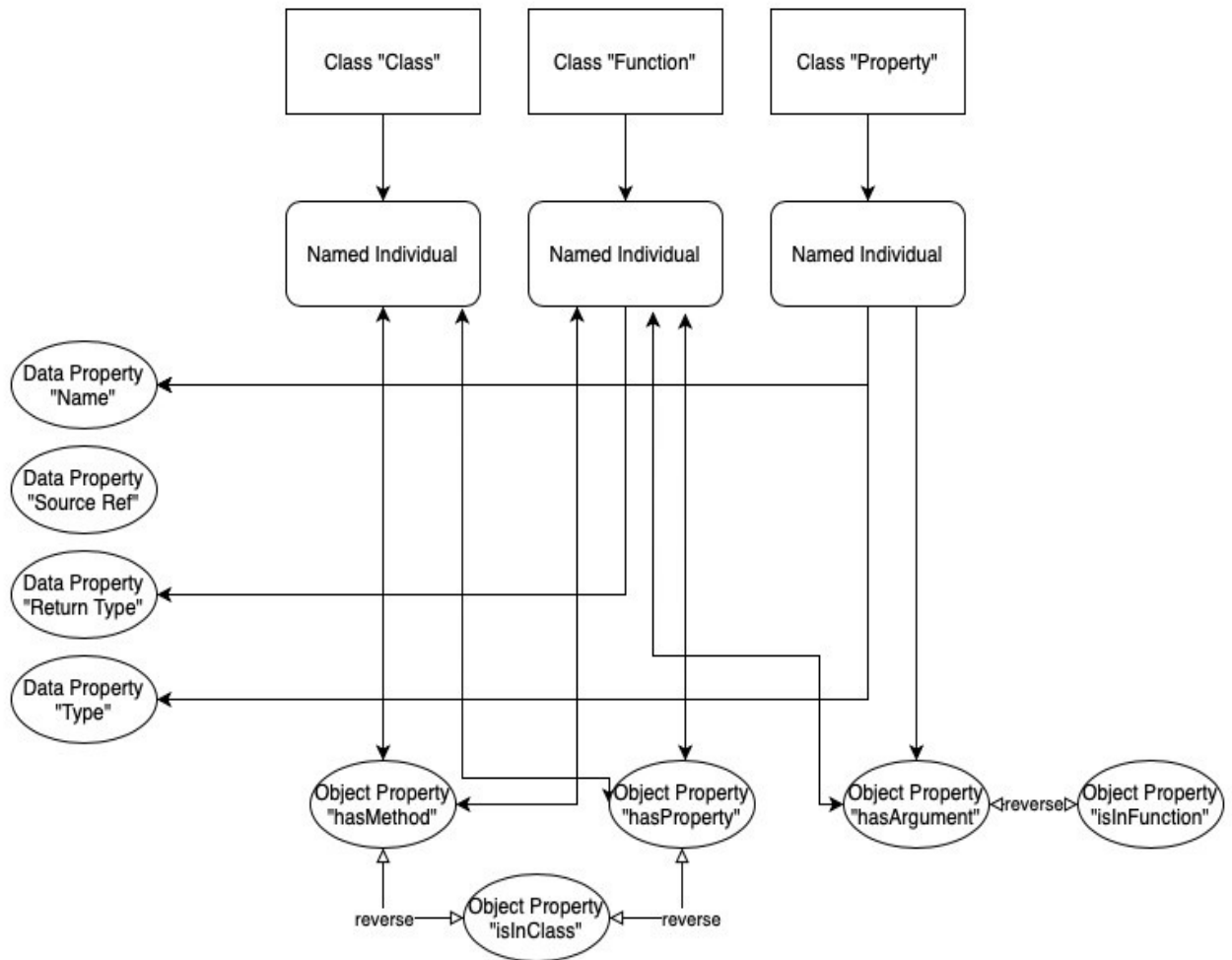


Fig. 4. Ontology structure

To be able to describe named individuals, a set of data properties is created. They are “*name*”, “*type*” (data type, like *integer*), “*returnType*” (data type). Additionally, there is a “*sourceRef*” (meaning “*SourceCodeReference*”) property which states where in the codebase specific entity is located. That property is beneficial as it will allow to uniquely identify the source code entity and provide backward compatibility – to be able to generate or alter code from the ontology.

Additionally, “*isInClass*” and “*isInMethod*” object properties are defined as disjoint, meaning that named individual which can have both these properties (Property in our case) cannot have these two relations associated at one time.

On the Parsing client side the SwiftSyntax library [9] is used to parse source code written in Swift language. Then parsed nodes of AST are converted into DTO – defined entities, which are known by Ontology service. The following code fragment depicts a piece of code which parses the VariableDeclSyntax node (represent properties of class, structure, enumeration, or regular variables) and transforms it to a Property object:

```

override func visit(_ node: VariableDeclSyntax) -> SyntaxVisitorContinueKind {
    let name = node.bindings.first?.pattern.description ?? ""
    let type = node.bindings.first?.typeAnnotation?.type.description ?? ""
    let property = Property(name: name,
        type: type,

```

```

        sourceCodeReference: node.sourceRange(
                                converter: converter).debugDescription)

    properties.append(property)
    return .skipChildren
}

```

As the last step, Alamofire library is used to connect to the Ontology service via HTTP and transfer all visited entities in provided source code. On the Ontology service side, a simple HTTP server from the Sun package provides REST API for communication. Here is a piece of Java code which handles the input and created new named individual that belongs to class "Class":

```

@Override
public String handleInput(String input) {
    Class classEntity = converter.fromJson(input, classOfT:Class.class);
    return InternalOntology.sharedInstance().addClass(classEntity);
}

```

Data which is transferred between Parsing client and Ontology service is in JSON format. The following code sample creates a new named individual:

```

private OWLNamedIndividual createFunction(Function func) {
    IRI identifier = IRI.create(iri, UUID.randomUUID().toString());
    OWLNamedIndividual funcDecl = dataFactory.getOWLNamedIndividual(identifier);
    ontology.addAxiom(dataFactory.getOWLDeclarationAxiom(funcDecl));
    ontology.addAxiom(dataFactory.getOWLObjectPropertyAssertionAxiom(
        functionEntity, funcDecl));
    ontology.addAxiom(dataFactory.getOWLObjectPropertyAssertionAxiom(
        nameProperty, funcDecl, func.name));
    ontology.addAxiom(dataFactory.getOWLObjectPropertyAssertionAxiom(
        returnTypeProperty, funcDecl, func.returnType));
    ontology.addAxiom(dataFactory.getOWLObjectPropertyAssertionAxiom(
        sourceCodeReferenceProperty, funcDecl, func.sourceCodeReference));

    for (Property prop: func.arguments) {
        OWLNamedIndividual propDecl = createProperty(prop);
        ontology.addAxiom(dataFactory.getOWLObjectPropertyAssertionAxiom(
            hasArgumentProperty, funcDecl, propDecl));
    }
    return funcDecl;
}

```

Firstly, a unique IRI is created for a new instance. Secondly, new named individual object is created, and all possible axioms are added. Lately, all dependent entities (Property named individuals in this case) are created in the same manner, which is described in first two steps.

To verify that expected ontology contains all the necessary axioms, a test should be conducted. For testing purposes, a simple code snippet was created. Lets assume that this snippet is a part of some bigger source file (note that we have added some line numbers for better understanding the idea of how *Source Code Reference* is used). This demo code snippet has some very simple class hierarchy yet demonstrating the inheritance and method overriding:

7 ...	27 override init() {
8 protocol Say {	28 lives = 9
9 func say() -> String	29 super.init()
10 }	30 }
11	31
12 class Animal: Say {	32 override func say() -> String {
13 let age: Int	33 return "meow"
14	34 }
15 init() {	35 }
16 age = 0	36
17 }	37 final class Dog: Animal {
18	38 private var senses: Int
19 func say() -> String {	39
20 return ""	40 override init() {
21 }	41 senses = 6
22 }	42 super.init()
23	43 }
24 final class Cat: Animal {	44
25 private var lives: Int	45 override func say() -> String {
26	46 return "woof"
	47 }
	48 }
	49 ...

After parsing the source code, ontology was created. Then, Hermit reasoner was run to infer all possible axioms. As a result, inversions, and basic ontology class hierarchy were added. Updated ontology is on fig. 5, where bold elements depict the inference results added to the initial ontology by Hermit reasoner.

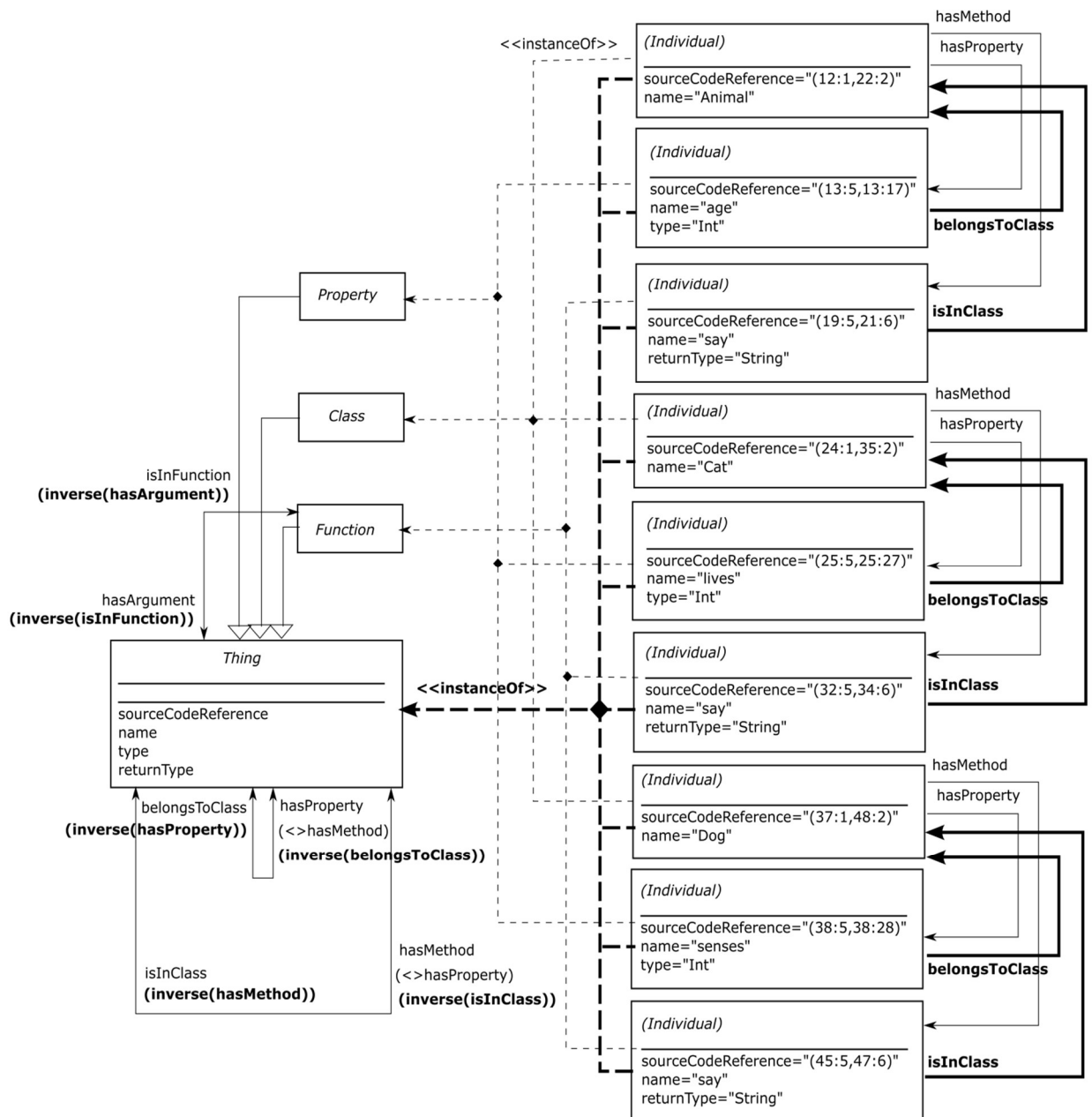


Fig. 5. Ontology after reasoning (inferred axioms are in bold, individual's identifiers are skipped for brevity)

Unfortunately, not all expected knowledge about the code appeared in the ontology. As it may be seen on the code snippet, there are hierarchical relations (inheritance and protocol conformance) in the code, but neither initial ontology nor updated one after reasoning do not contain those relations. This means that further investigation and design should be conducted to find and describe the optimal set of properties and entities which should be supported by code ontology extraction system to be able to infer more knowledge about source code.

Conclusions

Review of the recent research and articles has showed that ontology is a powerful tool to represent domain semantics and knowledge. It is safe to state that ontologies can be used to efficiently keep the knowledge about the source code in particular.

In order to be able to represent codebase efficiently, a wide range of the source code ontology's 'building blocks' must be defined. They include but not limited to classes, named individuals, data properties, object properties and wide variety of axioms.

When all the axioms in ontology are correctly representing the state of the codebase, a reasoner may help to infer implicit or unknown relations/axioms and add them to the initial ontology. Such ability of the proposed system allows to add rules to the ontology (such as code smell anti-patterns and approaches to fix them) to remove, refactor or generate code which will compile.

Several approaches to convert codebase to ontology were discussed. This article has shown that the code ontology extraction is quite suitable approach to transform the source code into a knowledge base represented by OWL2 ontology.

In order to verify the proposed concept, a design of simple basic knowledge base was developed, and a code ontology extraction system prototype was built. Its architecture and principal design points were discussed. As a result, a code snippet was transformed into a simple ontology. Implicit and unknown axioms were inferred by the reasoner and successfully added to the initial ontology.

That simple knowledge base structure used does not cover all the possible relations and code entities, so in order for the system be able to infer complex data (like protocol conformance) it is necessary to design a set of axioms which will allow the reasoner to do so and submit sufficient data about source code to the ontology manager. The further research will be devoted to the development of the more complex semantic foundation for the real-world source code knowledge extraction and manipulation.

Література

1. Happel, H.-J., Seedorf, S. Applications of ontologies in software engineering. *In Proc. of Workshop on Semantic Web Enabled Software Engineering (SWESE) on the ISWC*, 2006, pp. 5-9.
2. As we may code. [Електронний ресурс] - Режим доступу: <https://nshipster.com/as-we-may-code/#skip>
3. OWL 2 Web Ontology Language. [Електронний ресурс] - Режим доступу: <https://www.w3.org/TR/owl2-syntax/>
4. Semantic Web Tool. [Електронний ресурс] - Режим доступу: <https://www.w3.org/wiki/SemanticWebTools>
5. Resource Description Framework (RDF). [Електронний ресурс]. - Режим доступу: <https://www.w3.org/RDF/>
6. Glimm, B., Horrocks, I., Motik, B., Stoilos, G., Wang, Z. HermiT: an OWL 2 reasoner. *Journal of Automated Reasoning*, Springer Netherlands, 2014, No. 53, pp. 245 – 269.
7. Horrocks, I., Motik, B., Wang, Z. The HermiT OWL Reasoner. *CEUR Workshop Proceedings*, 2012, Vol. 3(858).
8. Tkachuk, A., Bulakh, B. Research of possibilities of default refactoring actions in Swift language. *Technology Audit and Production Reserves*, 2022, No. 5(2(67)), pp. 6–10. DOI: [10.15587/2706-5448.2022.266061](https://doi.org/10.15587/2706-5448.2022.266061)
9. Tkachuk, A., Bulakh, B. Usage of formalized knowledge about source code for refactoring actions in Swift. *Technology Audit and Production Reserves*, 2022, No. 6(2(68)), pp. 6–10. DOI: [10.15587/2706-5448.2022.267160](https://doi.org/10.15587/2706-5448.2022.267160)
10. Ruta, M., Scioscia, F., Di Sciascio, E., Bilenchi, I. OWL API for iOS: early implementation and results. *13th OWL: Experiences and Directions Workshop and 5th OWL reasoner evaluation workshop (OWLED - ORE 2016)*, Springer, 2016, vol. 10161, pp. 141–152, DOI: 10.1007/978-3-319-54627-8
11. Ruta, M., Scioscia, F., Gramegna, F., Bilenchi, I., Di Sciascio, E. Mini-ME Swift: The First Mobile OWL Reasoner for iOS. *In: The Semantic Web. ESWC 2019. Lecture Notes in Computer Science*, 2019, vol. 11503, Springer, Cham, pp. 298-313. DOI: 10.1007/978-3-030-21348-0_20
12. An Introduction to the OWL API. [Електронний ресурс] - Режим доступу: <http://syllabus.cs.manchester.ac.uk/pgt/2019/COMP62342/introduction-owl-api-msc.pdf>
13. ONT-API: an RDF-centric Java library to work with OWL2. [Електронний ресурс] - Режим доступу: <https://github.com/owlcs/ont-api>
14. Dirsumilli, R., Mossakowski, T. RESTful Encapsulation of OWL API. *5th International Conference on Data Management Technologies and Applications*, 2016, pp. 150 - 157. DOI: 10.5220/0005987201500157.

References

1. Happel, H.-J., Seedorf, S. (2006). Applications of ontologies in software engineering. *In Proc. of Workshop on Semantic Web Enabled Software Engineering (SWESE) on the ISWC*, pp. 5-9.
2. As we may code. [Електронний ресурс] - Режим доступу: <https://nshipster.com/as-we-may-code/#skip>
3. OWL 2 Web Ontology Language. [Електронний ресурс] - Режим доступу: <https://www.w3.org/TR/owl2-syntax/>
4. Semantic Web Tool. [Електронний ресурс] - Режим доступу: <https://www.w3.org/wiki/SemanticWebTools>
5. Resource Description Framework (RDF). [Електронний ресурс]. - Режим доступу: <https://www.w3.org/RDF/>
6. Glimm, B., Horrocks, I., Motik, B., Stoilos, G., Wang, Z. (2014). HermiT: an OWL 2 reasoner. *Journal of Automated Reasoning*, Springer Netherlands, No. 53, pp. 245 – 269.
7. Horrocks, I., Motik, B., Wang, Z. (2012). The HermiT OWL Reasoner. *CEUR Workshop Proceedings*, Vol. 3(858).
8. Tkachuk, A., Bulakh, B. (2022). Research of possibilities of default refactoring actions in Swift language. *Technology Audit and Production Reserves*, No. 5(2(67)), pp. 6–10. DOI: 10.15587/2706-5448.2022.266061

9. Tkachuk, A., Bulakh, B. (2022). Usage of formalized knowledge about source code for refactoring actions in Swift. *Technology Audit and Production Reserves*, No. 6(2(68)), pp. 6–10. DOI: 10.15587/2706-5448.2022.267160
10. Ruta, M., Scioscia, F., Di Sciascio, E., Bilenchi, I. (2016). OWL API for iOS: early implementation and results. *13th OWL: Experiences and Directions Workshop and 5th OWL reasoner evaluation workshop (OWLED - ORE 2016)*, Springer, vol. 10161, pp. 141–152, DOI: 10.1007/978-3-319-54627-8
11. Ruta, M., Scioscia, F., Gramegna, F., Bilenchi, I., Di Sciascio, E. (2019). Mini-ME Swift: The First Mobile OWL Reasoner for iOS, *In: The Semantic Web. ESWC 2019. Lecture Notes in Computer Science*, vol. 11503, Springer, Cham, pp. 298-313. DOI: 10.1007/978-3-030-21348-0_20
12. An Introduction to the OWL API. [Електронний ресурс] - Режим доступу: <http://syllabus.cs.manchester.ac.uk/pgt/2019/COMP62342/introduction-owl-api-msc.pdf>
13. ONT-API: an RDF-centric Java library to work with OWL2. [Електронний ресурс] - Режим доступу: <https://github.com/owlcs/ont-api>
- Dirsumilli, R., Mossakowski, T. (2016). RESTful Encapsulation of OWL API. *5th International Conference on Data Management Technologies and Applications*, pp. 150 - 157. DOI:

УДК 004.94

ПОРІВНЯННЯ АЛГОРИТМІВ КЛАСИФІКАЦІЇ BIG DATA МЕТОДАМИ ІМІТАЦІЙНОГО МОДЕЛЮВАННЯ

DOI 10.36994 / 2788-5518-2023-01-05-15

Одегов М.А., к.т.н., доц. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна. onick_64@ukr.net

Гаджиєв М.М., д.т.н., проф. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна. gadjievmm@ukr.net

Буката Л.М. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна. ygrikluda@gmail.com

Глазунова М.В., к.ф.-м.н. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна. lvglazun@gmail.com

Кочеткова М.В. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна., jubdvvg@gmail.com

Анотація: У статті вирішується задача порівняльного аналізу швидких алгоритмів класифікації, що можуть застосовувати для вирішення задач з надвеликими об'ємами даних (Big Data). Задача розв'язується методами імітаційного моделювання за допомогою програми Adaptive Metrics. Алгоритми найближчих сусідів, центрів класів та адаптивних правил порівнюються за критеріями надійності та продуктивності. Отримані результати дозволяють зробити висновок, що алгоритми, засновані на принципах M-means можуть ефективно використовуватись в задачах класифікації за певних умов, оскільки мають значну перевагу за критерієм продуктивності.

Ключові слова: алгоритми найближчих сусідів, алгоритм центрів класів, алгоритм адаптивних правил, критерій надійності, критерій продуктивності, методи імітаційного моделювання, програма імітаційного моделювання.

COMPARISON OF BIG DATA CLASSIFICATION ALGORITHMS BY SIMULATION METHODS

Nick Odegov, Ph.D., Ass. Prof. State University of Intellectual Technologies and Communication, Odesa, Ukraine. onick_64@ukr.net

Matin Hadzhyiev, Dr. habil., Prof. State University of Intellectual Technologies and Communication, Odesa, Ukraine. gadjievmm@ukr.net

Liudmyla Bukata. State University of Intellectual Technologies and Communication, Odesa, Ukraine. ygrikluda@gmail.com

Liudmyla Glazunova, Ph.D., Ass. Prof. State University of Intellectual Technologies and Communication, Odesa, Ukraine. lvglazun@gmail.com

Marina Kochetkova. State University of Intellectual Technologies and Communication, Odesa, Ukraine., jubdvvg@gmail.com

Abstract: With the development of information transmission and storage technologies, the volumes of data that require processing and analysis are growing rapidly. Therefore, the task of developing algorithms for solving various artificial intelligence problems for Big Data volumes is urgent. In our works, this informal term "Big Data" refers to situations when known processing algorithms do not allow solving a problem in a practically acceptable time. With regard to classification tasks, such conditions are possible when the first place is not even high reliability (that is, the minimum number of errors), but productivity (classification speed).

The well-known method of nearest neighbors is one of the most productive. However, the indicator of the order of growth (the number of typical operations) for it is $K \times M \times N$, where K is the number of nearest neighbors, M is the number of classes, N is the typical number of class elements.

Along with this, we propose to consider algorithms based on the principles of M-means, where classes are replaced by only a small number of their characteristics. Among such algorithms, the article considers: the algorithm of class centers and the algorithm of adaptive rules. The order of growth for these algorithms is only M according to the number of classes.

The comparative analysis of these algorithms is performed by the method of simulation modeling. Simulation models are implemented by the Adaptive Metrics program, developed at the Department of Software

Engineering at DUITZ. In this program, the classification problem is solved using the example of the dichotomy problem for classes A and B. The program has the possibility of very flexible setting of models. Problems can be solved in 1-dimensional, 2-dimensional,..., 6-dimensional spaces. The distribution of factor values for classes A and B can have quite different statistical characteristics - from uniform and triangular distribution functions to functions approaching a normal distribution. The graphical interface of the program allows you to dynamically observe the solution of the classification problem in one-dimensional, two-dimensional and 6-dimensional projections.

As a result of multiple runs of the program, it was established that the algorithms of the nearest neighbors slightly outperform the algorithms of class centers and adaptive rules according to the criterion of reliability, and also comply with the principle of compactness (concentration of the largest number of erroneous solutions in the hypercube of errors). Algorithms based on M-means principles significantly outperform this algorithm in terms of performance. Also, the algorithm of adaptive rules best corresponds to the principle of equality of classes and is the most productive of the considered ones

Key words: nearest neighbor algorithms, class centers algorithm, adaptive rule algorithm, reliability criterion, performance criterion, simulation modeling methods, simulation simulation program.

Вступ

З поняттям Big Data автори цієї статті зіткнулися вже у простій задачі дихотомії, коли базові сукупності даних у класі А і класі В складали лише 10000 об'єктів у 6-мірному просторі. Треба було вирішити задачу віднесення 1000 тестових об'єктів кожного класу за допомогою алгоритму К найближчих сусідів [1], який вважається одним з найбільш ефективних з точки зору продуктивності (швидкості прийняття рішень) [2]. Число К для обох класів визначалось рівним 100. У авторів було враження, що досить сучасний комп'ютер «зависає». Втім, він напружено працював над задачею. Дійсно, у даному випадку значення порядку росту алгоритму [3] складає $2 \times 6 \times 10000 \times 1000 \times 100 = 12 \times 10^9$ (дванадцять мільярдів!) типових операцій, кожна з яких – визначення метрики у евклідовому просторі.

Порядок росту алгоритму К найближчих сусідів складає:

$$P=M \cdot K \cdot N, \quad (1)$$

де M – кількість класів; K – кількість найближчих сусідів; N – кількість відомих (вже раніше класифікованих) об'єктів класів.

В статті [4] обґрунтовано ряд швидких алгоритмів класифікації, які засновані на принципах M-means, що зазвичай використовуються для вирішення задач кластеризації [1, 2]. Порядок росту для цих алгоритмів складає лише $P = M$.

Виконання досліджень

Метою даної статті є порівняльний аналіз алгоритмів К найближчих сусідів з алгоритмами центрів класів та адаптивних метрик [4] за критеріями надійності та продуктивності.

Оскільки реальні задачі класифікації можуть вирішуватись в умовах досить різних варіантів розподілення значень факторів, то аналітичної оцінки у вигляді замкнених формул отримати в загальному випадку неможливо. Тому у даній статті задача вирішується методами імітаційного моделювання [5].

З часу виходу статті [6] (Ніколас, Метрополіс, 1949 р.) сфера застосування методів імітаційного моделювання поширилась на майже безмежну кількість наукових та практичних задач. Наведемо лише декілька прикладів. Застосування у сільському господарстві [7], у сфері будівництва [8], в задачах управління водними ресурсами [9], в задачах оптимізації протоколів передачі даних [10], в задачах оптимізації проектування авіаційної техніки [11], та, навіть, в медицині [12]. Тобто, головна проблема застосування цих методів – не визначення сфери, а синтез адекватної моделі. У даній статті ця проблема вирішується за допомогою гнучкого налаштування моделі з можливістю обрання дуже великої кількості варіантів розподілень значень факторів.

У даній статті задача порівняння методів класифікації вирішується для задачі дихотомії як базової задачі [4]. Класи умовно позначаються як А («зелений») та В («синій»).

1. Загальний опис імітаційної моделі

Імітаційне моделювання здійснюється за допомогою програми Adaptive Metrics (мова програмування C# , IDE Visual Studio 2022), яка розроблена співробітниками кафедри ІПЗ ДУІТЗ. Загальний вигляд вікна програми у певному варіанті показаний на рис. 1.

Моделювання виконується у фактор-просторі розмірності $D=6$. Окремо для кожного фактору для класів А і В визначаються функції розподілення значень. Базовий метод імітації значень у даному випадку – отримання псевдовипадкових (надалі – просто випадкових) чисел. Ці числа формуються відомим конгруентним алгоритмом [5] спочатку у діапазоні $r = [0,1]$. Такі числа приблизно рівномірно розподілені у даному діапазоні. Для отримання значень факторів класу А для рівномірних розподілень використовуються формула:

$$1R_A = SA \cdot r \quad (2)$$

У формулі (2) і далі символ $1R$ розуміється як рівномірне розподілення. Конкретно у даній моделі параметр S_A дорівнює 300 (що не суттєво). Різні інші унімодальні розподілення для класу А позначаються як $2R, 3R, \dots, 10R$ і отримуються в результаті підсумовування відповідної кількості рівномірно розподілених чисел:

$$HR = \frac{1}{H} \sum_{h=1}^H 1R, \quad H=1,2,\dots,10. \quad (3)$$

Очевидно, що всі ці розподілення є симетричними відносно центру, який у графічному відображенні позначається як ім'я класу. У даному для «зеленого» класу завжди $A = 150$.

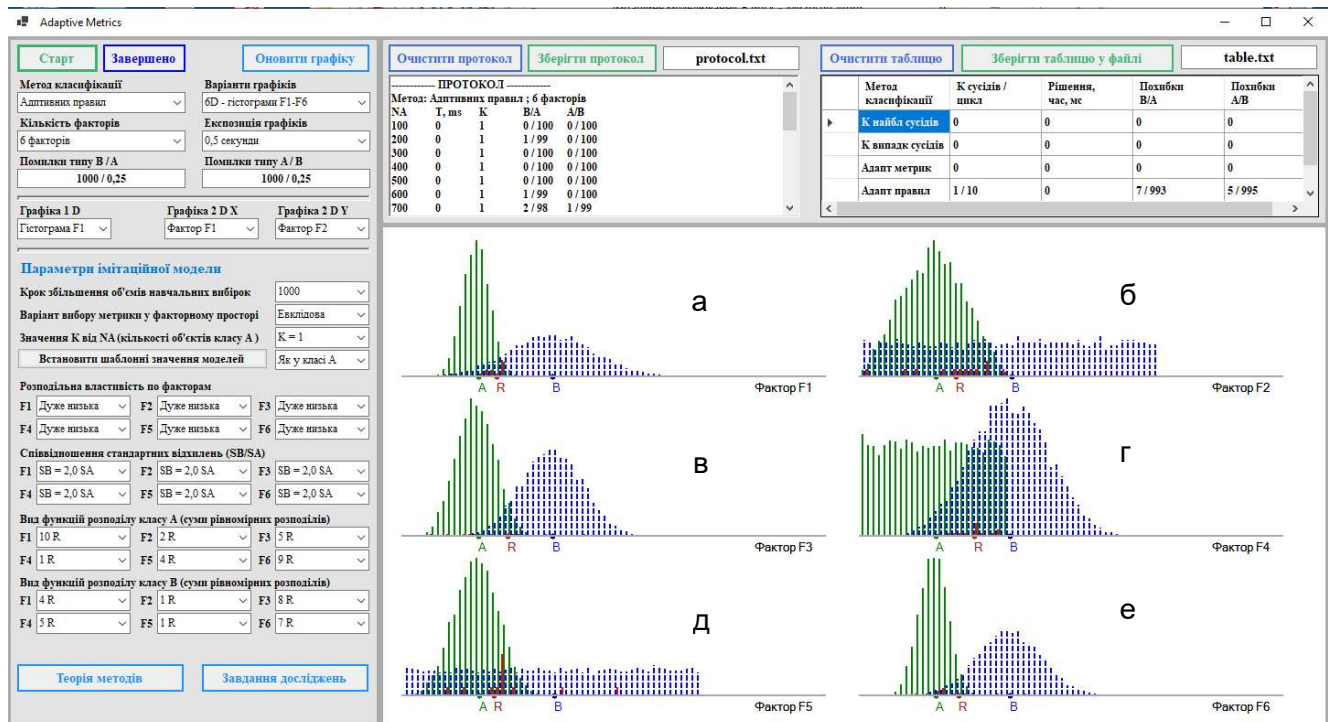


Рис. 1. Загальний вигляд вікна програми моделювання Adaptive Metrics з різними варіантами функцій розподілення окремих факторів для класів А і В:

а) А : 10R, В : 4R; б) А : 2R, В : 1R; в) А : 5R, В : 9R; г) А : 1R, В : 5R; д) А : 4R, В : 1R; е) А : 9R, В : 7R

Алгоритми генерування випадкових чисел намагаються зробити такими, щоб окремі послідовності були слабо корельованими, а краще – статистично незалежними. Відомо, що

щільність розподілення суми двох незалежних випадкових величин X та Y є згорткою їх щільностей розподілення [13]: $p(X+Y)=p(X)\otimes p(Y)$. Тому правило (3) формування послідовностей значень факторів для варіанту 2R можна представити у такій формі:

$$p(2R)=p(1R)\otimes p(1R), \quad (4)$$

а у загальному вигляді дана згортка виконується $H - 1$ разів:

$$p(HR)=p(1R)\otimes p(1R)\otimes \dots \otimes p(1R). \quad (5)$$

Із теорії спектрів [14] відомо, що операції згортки у часовому просторі відповідає операція множення у частотному просторі. Не важко провести аналогію між щільністю рівномірного розподілення випадкової величини та прямокутним відео імпульсом, спектр якого має форму функції $\text{sinc}(x)=\sin(x)/x$. Відповідно, спектр згортки двох прямокутних імпульсів буде мати форму функції $\text{sinc}^2(x)$. А такий спектр має трикутний імпульс [14]. Таким чином, розподілення типу 2R (4) повинно бути приблизно трикутним, якщо одномірні розподілення статистично незалежні. Приблизно такий варіант і спостерігається на рис. 1,б.

Функції виду $\text{sinc}^H(x)$ досить швидко сходяться до функцій виду $\exp(-x^2)$ (гаусової функції). Втім, зворотне перетворення Фур'є гаусової функції знову ж таки приводить до гаусової функції, а у термінах теорії ймовірностей – до нормальної щільності розподілення. Це добре видно на прикладі рис. 1, в.

Висновок: послідовності випадкових чисел, що використовуються у даній роботі, можна вважати приблизно статистично незалежними.

Розподілення (3) при різних значеннях H будуть відрізнятися стандартними відхиленнями, а також ексцесами. Відомо, що дисперсія суми незалежних випадкових величин дорівнює сумі дисперсій [15]. Дисперсія кожного доданку у сумі (4) складає $D = D(1R)/H^2$, тоді дисперсія суми дорівнює $D(HR)=D(1R)/H$ і для стандартних відхилень буде справедлива теоретична формула:

$$\sigma(HR)=\frac{\sigma(1R)}{\sqrt{H}}. \quad (6)$$

Розподілення факторів класу В («синього») відрізняються лише параметрами зсуву та масштабу, тобто залежність (3) модифікується таким чином:

$$1R_B = SB \cdot SA \cdot r - B(SB), \quad B(SB) > 0, \quad (7)$$

де параметр SB у даному варіанті програми може приймати значення 1, 1,5 або 2. Відповідно, у нашому випадку розмах значень для класу В може складати 300, 450 або 600.

2. Алгоритм навчання класів та тестування класифікаторів

Базові послідовності значень 6-ти факторів для кожного класу формуються за формулами (3) та (7) у такій послідовності: спочатку об'єми класів дорівнюють базовим значенням N_1 . В програмі передбачено два варіанти значень N_1 : 100 або 1000.

Навчання для алгоритму К найближчих сусідів полягає лише у тому, що значення факторів відносяться до відповідних класів. Тобто, моделюється ситуація, коли вже є передісторія досліджень, згідно якої визначено базові сукупності класів А і В.

Об'єми класів далі нарощуються з кроком N_1 9 разів. Тобто об'єм навчальних вибірок поступово збільшується: 100, 200, ..., 1000 (або 1000, 2000, ..., 10000).

Для алгоритмів центрів класів і адаптивних правил на кожному такому кроці за допомогою рекурентних формул визначаються параметри: середні значення A_d і B_d , стандартні відхилення $\sigma_{A,d}$ і $\sigma_{B,d}$, а також вирішальні границі R_d для кожного фактору $d=1,2\dots6$.

На кожному з десяти кроків отримання значень навчальних вибірок формуються за тими ж самими правилами тестові вибірки об'єму N_1 для кожного класу. Також на кожному з цих кроків обраним алгоритмом класифікації елементи тестових послідовностей відносяться до класу А або до класу В. При цьому визначається кількість вірних рішень ($A=A$, $B=B$) та кількість помилок 1-го та 2-го типів (A/B , B/A) [4].

На кожному кроці нарощування об'ємів вибірок проміжні результати заносяться у протокол, а останній варіант при максимальному об'ємі базових елементів класів – у таблицю (рис. 1).

3. Варіанти налаштувань імітаційної моделі

Програма Adaptive Metrics дозволяє обрати один з алгоритмів класифікації [4]:

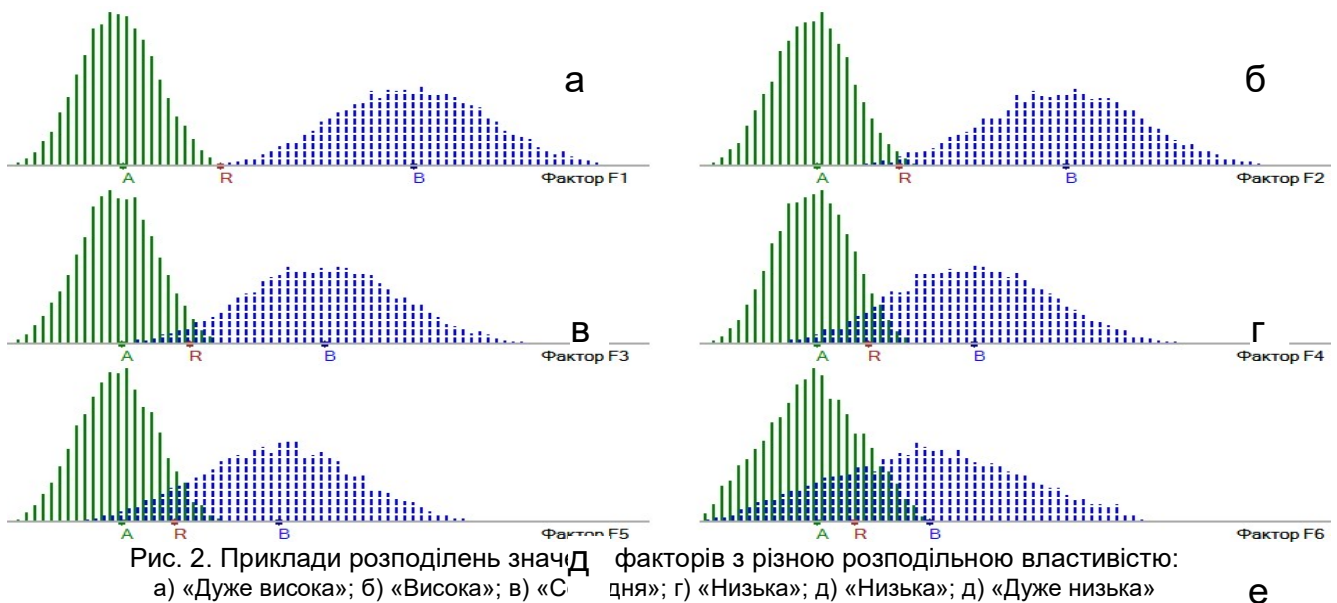
- К найближчих сусідів;
- К випадкових сусідів;
- центрів класів;
- адаптивних метрик;
- адаптивних правил.

Для перших двох алгоритмів число сусідів К можна обрати: або $K = 1$, або як відсоток від об'ємів базових елементів класів А і В. Програма дозволяє обирати відсотки від 1% до 10%. Тобто, максимальне значення К може бути від 10 до 1000.

Класифікація може здійснюватись лише з урахуванням першого фактору (F_1), лише з урахуванням факторів F_1 та F_2 , ..., з урахуванням всіх 6-ти факторів.

Для обчислення метрик передбачено два варіанти: або Евклідова метрика, або метрика Хемінга.

Для кожного фактора окремо може налаштовуватись параметр розподільної властивості, який визначає відстань між центрами значень факторів для класів А і В. Значення даного параметру: «Дуже висока», «Висока», «Середня», «Низька», «Дуже низька». Приклади розподілень відповідно значенням даного фактора наведено на рис. 2.



Про налаштування параметрів розподілень сказано у п.1. Загалом програма дозволяє обрати:

- значення параметрів розподільної властивості – один з 30 варіантів;

- значення параметру SB (окремо для кожного фактора) – один з 18 варіантів;
- варіантів вибору видів розподілень – по 60 окремо для кожного класу.

Тобто, всього можна обрати один з 64800 варіантів налаштування моделей. Тому імітаційна модель є достатньо гнучкою для вирішення задач порівняння характеристик алгоритмів класифікації.

Зауважимо, що серед моделей розподілень немає асиметричних функцій (для яких коефіцієнт асиметрії не дорівнює нулю). Це обмеження допущено навмисно: найбільше значення для алгоритмів класифікації має не стільки вид функцій розподілення, скільки доля елементів класів, які потрапляють в область перетину значень факторів.

4. Варіанти графічного відображення результатів

Графічному відображенню результатів навчання алгоритмів та класифікації закономірно приділяється значна увага [2]: звичайний (не штучний) інтелект людини дозволяє вирішувати різні задачі, якщо базова інформація представлена у логічному та наглядному вигляді. Тому програма Adaptive Metrics дозволяє представляти проміжні та кінцеві результати класифікації у дуже наочних графічних формах. А саме, передбачено три варіанти відображення:

- 1D – одновимірна проекція розподілень значень факторів у вигляді гістограм для класів A і B (рис. 3);
- 6D – відображення розподілень по всім 6-ти факторам у вигляді гістограм (рис. 1, 2);
- 2D – проекція значень базових елементів та помилок класифікації на площину будь-яких різних двох факторів (рис. 4).

У варіантах 1D та 6D гістограми помилок відображаються в умовному масштабі без розрізнення кольору варіантів A/B і B/A. У варіанті 2D базові елементи відображаються окремими пікселями зеленого (для класу A) та синіми (для класу B) кольору; помилки відображаються крупними значками відповідно червоного (B/A) та коричневого кольорів (A/B) (рис. 4).

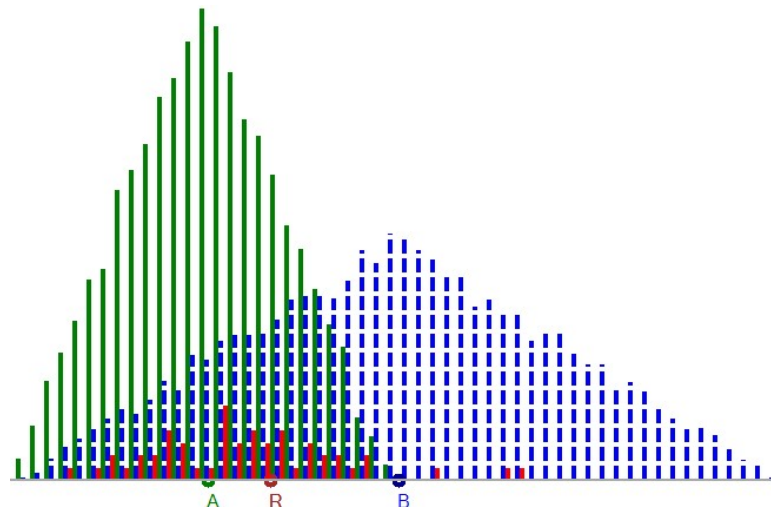


Рис. 3. Приклад проекції 1D на фактор F1

У кожному варіанті (рис. 1-4) на графіках букви A і B показують центри відповідних класів, а буквою R позначається вирішальна границя (у даному випадку від слова «Rule») для алгоритмів адаптивних метрик та адаптивних правил.

Всі результати розрахунків зберігаються в оперативній пам'яті, тому як під час виконання числових експериментів, так і після їх завершення можна вивести у вікно програми будь-який варіант графіки.

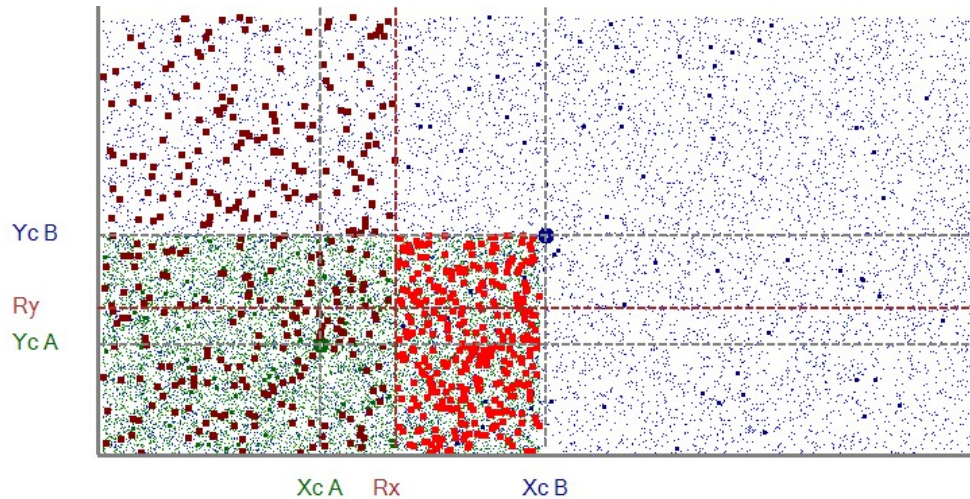


Рис. 4. Приклад 2D проекції: класифікація лише за фактором F1; метод адаптивних правил

5. Поняття гіперкубу помилок

У Евклідовому просторі значення факторів елементів класів A і B можна представити у вигляді D – розмірного прямокутного паралелепіпеду з обмеженими ребрами. Після приведення значень факторів до єдиної шкали даний паралелепіпед перетворюється у D – розмірний гіперкуб. Саме такий термін ми й будемо використовувати надалі.

На наше переконання, у практичних задачах значення факторів завжди обмежені. Обмеження можна встановити як на підставі вивчення властивостей елементів класів у задачах Big Data, так, навіть, за допомогою здорового глузду. Наприклад, у «класичній задачі» – видавати чи не видавати кредит [2] – дихотомія вирішується у двовірному просторі: F1 – вік людини; F2 – дохід. Тут, мабуть, мова йде про споживчий кредит: на покупку смартфона, ноутбука та т. ін. Більш коштовні покупки потребують застави і алгоритм прийняття рішення там дуже сильно відрізняється від цього.

Отже, ми не дуже помилились, якщо передбачимо можливі значення віку від 16 до 140 років, а місячний дохід – від мінімальної пенсії до, наприклад, 1 000 000 грн: навряд чи людина з великим доходом прийде за кредитом, який значно менше розміру його «кишенькових» грошей.

Таким чином, по кожному фактору для кожного класу можна визначити одинірну проекцію гіперкубу можливих значень як відрізок прямої між мінімальним та максимальним значенням, що ми позначимо як $HQ(A, d)$ та $HQ(B, d)$, де $d=1, 2, \dots, D$. Далі, користуючись звичайними позначеннями теорії множин та теорії ймовірнісних просторів [15], визначимо одинірні проекції гіперкубів всіх значень для обох класів $HQ_d(A \cup B)$ та одинірні проекції областей можливих помилок $HQE_d(A \cap B)$. Тоді гіперкуби можливих значень окремих класів, значень для обох класів та гіперкубу помилок визначаються як відповідні Декартові добутки:

$$\begin{cases} HQ(A) = HQ(A, 1) \times HQ(A, 2) \times \dots \times HQ(A, D) \\ HQ(B) = HQ(B, 1) \times HQ(B, 2) \times \dots \times HQ(B, D) \\ HQ(A \cup B) = HQ_1(A \cup B) \times HQ_2(A \cup B) \times \dots \times HQ_D(A \cup B) \\ HQE(A \cap B) = HQ_1(A \cap B) \times HQ_2(A \cap B) \times \dots \times HQ_D(A \cap B) \end{cases} \quad (8)$$

Також визначимо об'єми цих гіперкубів як звичайні добутки довжин відрізків одинірних проекцій гіперкубів (8):

$$V(A), V(B), V(A \cup B), V(A \cap B) \quad (9)$$

та відносні об'єми цих гіперкубів:

$$\delta V(A) = \frac{V(A)}{V(A \cup B)}, \quad \delta V(B) = \frac{V(B)}{V(A \cup B)}, \quad \delta V(A \cap B) = \frac{V(A \cap B)}{V(A \cup B)}. \quad (10)$$

Визначення (8-10) дозволяють виконувати експрес-аналіз наборів даних у практичних задачах класифікації, а також оцінювати надійність алгоритмів якісно, а в деяких випадках і кількісно, що ми покажемо на прикладах.

6. Приклади порівняльного аналізу алгоритмів класифікації методом імітаційного моделювання

По перше, розглянемо найважчий випадок, який тільки можливий у пропонованій імітаційній моделі: значення всіх факторів розподілені рівномірно, параметр SB=2, варіант розподільної властивості для всіх факторів «Дуже низька» (рис. 5). Алгоритми класифікації при цьому враховують всі 6 факторів.

З рис. 5 добре видно, що розмах значень для всіх факторів класу B вдвічі більша, ніж для класу A. Також видно, що мають місце досить складні для класифікації співвідношення між гіперкубами значень та помилок:

$$HQ(B) \equiv HQ(A \cup B), \quad HQ(A) \equiv HQE(A \cap B),$$

тобто гіперкуб класу B дорівнює гіперкубу всіх можливих значень обох класів, а гіперкуб класу A дорівнює гіперкубу помилок. Також видно, що по кожному фактору значення для класу A займають вдвічі менший відрізок, ніж для класу B. Тому для відносних об'ємів гіперкубів справедливі співвідношення:

$$\delta V(B) = 1, \quad \delta V(A) = \delta V(A \cap B) = 1/64. \quad (11)$$

Числові характеристики (11) вказують на те, що у даному випадку можна очікувати «перекіс» частоти помилок у сторону A/B (реальний клас елементу тесту B, але його віднесено до класу A). Саме така закономірність простежується у табл. 1.

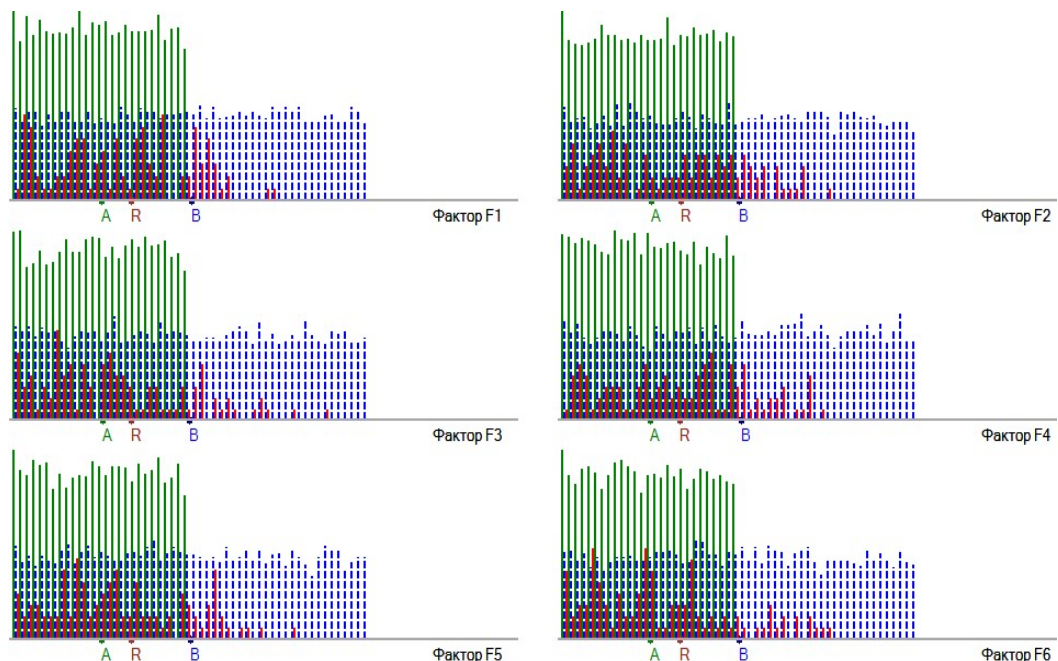


Рис. 5. Приклад класифікації алгоритмом К найближчих сусідів

Для досліджень якості алгоритмів за показником відповідності принципу «рівноправності класів» [4] доцільно ввести показник асиметрії помилок: співвідношення кількості помилок типу

В/А та А/В. При однаковій апіорній ймовірності появи у тестах представників класів А і В «хороший» алгоритм кластеризації повинен давати значення цього показника близьке до 1.

Таблиця 1.
Результати порівняльного аналізу алгоритмів класифікації

Метод	Найближчих сусідів, K=1			Центрів класів			Адаптивних метрик		
Об'єми класів	Час, мс	В/А	А/В	Час, мс	В/А	А/В	Час, мс	В/А	А/В
1000	18560	7	115	18	16	137	0	116	87
2000	36100	3	102	39	12	157	1	101	89
3000	53650	20	99	42	12	155	1	93	99
4000	73230	8	85	35	20	127	2	115	94
5000	96630	17	73	27	17	134	1	95	104
6000	120220	15	68	42	10	147	1	100	105
7000	129590	22	61	33	12	150	2	104	98
8000	154870	25	69	44	25	134	1	99	104
9000	165870	17	64	26	20	143	2	93	94
10000	190590	23	67	37	18	155	1	108	86
Загальний процент помилок, %			4,80	8,01			9,92		
Середнє значення показника асиметрії помилок			5,11	8,88			0,94		

Аналіз значень у табл. 1 дозволяє для даного прикладу налаштувань імітаційної моделі зробити наступні висновки:

- за критерієм надійності алгоритм найближчих сусідів (при K=1) приблизно вдвічі переважає алгоритми, що засновані на принципах M-means (алгоритми центрів класів та адаптивних правил);
- за критерієм продуктивності алгоритми, що засновані на принципах M-means суттєво переважають алгоритм K найближчих сусідів: алгоритм одного сусіда відпрацював більш як за 3 хвилини, тоді як алгоритм центрів класів приблизно за 20 мілісекунд, тоді як алгоритм адаптивних правил іноді вирішує задачу менше, ніж за 1 мілісекунду;
- за показником асиметрії помилок («рівноправності» класів) найкращі результати дає алгоритм адаптивних правил.

Як бачимо, у даному випадку алгоритми, засновані на принципах M-means поступаються алгоритму найближчих сусідів. Але не важко екстраполювати отримані результати, якщо у класах буде не 10000 об'єктів, а мільйон, або мільярд. Чи вистачить часу та терпіння дослідників для вирішення задачі класифікації, особливо у випадку, якщо кількість класів $M > 2$?

У даному випадку метод K найближчих сусідів має ще одну перевагу: не потребує навчання: базові елементи класів А і В апіорно задані як статичні множини (хтось колись у попередній історії вже відніс елементи класів до базових сукупностей, як у прикладі з наданням кредиту). Але визначення оптимального числа K сусідів все ж таки потребує навчання алгоритму [1]. Тому виконаємо аналіз лише для алгоритму K найближчих сусідів при зміні значення K, що також дозволяє програма Adaptive Metrics. Приклади протоколів розрахунків для того ж самого варіанту розподілень значень факторів (рис. 5) наведено у табл. 2.

Таблиця 2.

Результати порівняльного аналізу алгоритму К найближчих сусідів при $K=1\%$ та $K=2\%$ від кількості елементів класів

Кількість елементів класів	К сусідів	Помилки А/В	Помилки В/А	Процент помилок	К сусідів	Помилки А/В	Помилки В/А	Процент помилок
1000	10	57	128	9,25	20	9	133	7,1
2000	20	1	103	5,2	40	0	123	6,15
3000	30	0	107	5,35	60	0	106	5,3
4000	40	0	96	4,8	80	0	131	6,55
5000	50	0	116	5,8	100	0	108	5,4
6000	60	0	118	5,9	120	0	112	5,6
7000	70	0	109	5,45	140	0	92	4,6
8000	80	0	92	4,6	160	0	117	5,85
9000	90	0	82	4,1	180	0	120	6,0
10000	100	0	77	3,85	200	0	98	4,9

Аналіз значень у табл. 2 дозволяє зробити наступні висновки:

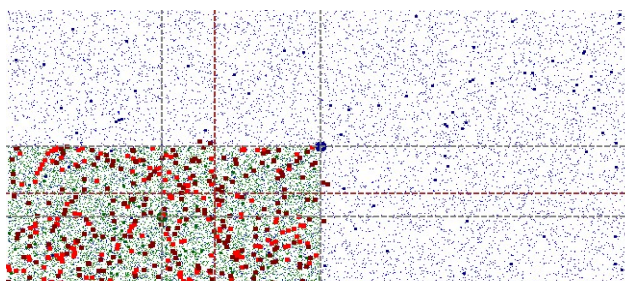
- процент помилок має незначну тенденцію зменшення при зростанні кількості К найближчих сусідів;

- загальний процент помилок, навіть при великій кількості К найближчих сусідів, незначно менший, ніж для випадку $K=1$, тобто при даному налаштуванні моделі підвищення кількості сусідів не призводить до суттєвого покращення за критерієм надійності;

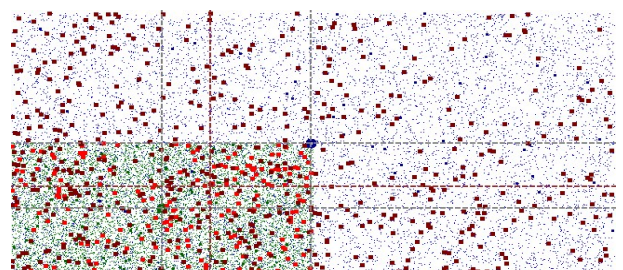
- варіація значень та відсутність чіткої функціональної залежності між кількістю К найближчих сусідів свідчить лише про те, що використані датчики випадкових чисел забезпечують досить впливову ймовірнісну складову, що можна побачити з рис. 5, де гістограми розподілень для дуже великих вибірок мають не зовсім рівномірний характер.

Наочно продемонструвати особливості алгоритмів, що розглядаються можна на прикладі того ж самого варіанту розподілень значень факторів, але у випадку класифікації лише з врахуванням двох факторів: F1 та F2. У даному випадку вичерпну інформацію у графічному вигляді можна отримати у варіанті відображення 2D (рис. 6, 7).

Можна прийняти без доказу, що «хороший» алгоритм класифікації забезпечує концентрацію основної кількості помилкових рішень всередині гіперкубу помилок (умова компактності помилок). Графічні відображення результатів класифікації показують, що алгоритм К найближчих сусідів (у даному випадку) відповідає цій умові. Алгоритм випадкових сусідів [4], хоча й дозволяє суттєво зменшити обсяг обчислень, але може давати значну кількість помилок класифікації (табл. 3). Також він зовсім не відповідає умові компактності помилок.



а) алгоритм найближчих сусідів ($K=1$)



б) алгоритм випадкових сусідів ($K=1$)

Рис. 6. Результати класифікації алгоритмами К сусідів

Результати класифікації, що представлені на рис. 7, демонструють особливості досліджених алгоритмів, заснованих на принципах M-means. Алгоритм центрів класів фактично задає регресійну поверхню між класами (штрихова лінія на рис. 7. а). У найпростішому варіанті даного

алгоритму регресійна поверхня перпендикулярна напрямленню від центру класу А до центру класу В. Недолік: дана регресійна поверхня не чутлива до масштабування значень факторів для різних класів і завжди проходить по середині між центрами цих класів. Відповідно, знаходиться певна кількість помилкових рішень, які у даному варіанті моделювання виходять за межі гіперкубу помилок.

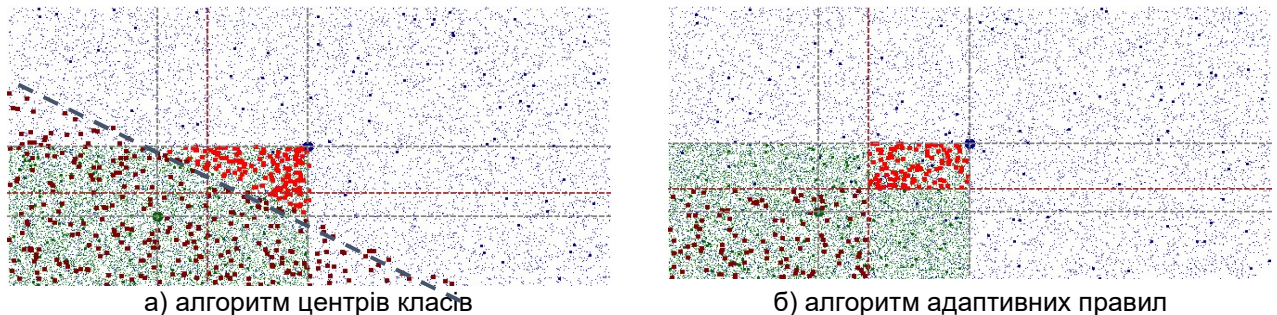


Рис. 7. Результати класифікації алгоритмами, заснованими на принципах M-means

Алгоритм адаптивних правил у даному випадку (рис. 7, б) дає коректні результати класифікації: помилки розподіляються згідно обраного адаптивного правила [4] та цілком зосереджуються у гіперкубі помилок.

Таблиця 3.

Результати порівняльного аналізу алгоритмів найближчих сусідів та алгоритмів, заснованих на принципах M-means у двомірному фактор-просторі

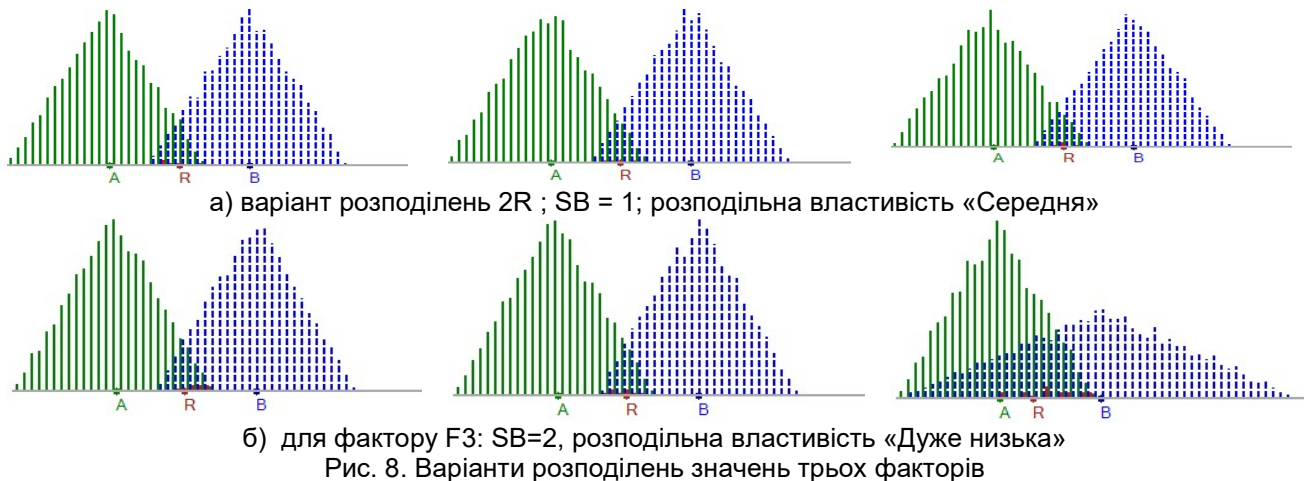
Алгоритм класифікації	Час прийняття рішень, мс	Помилки типу А/В	Помилки типу В/А	Загальний процент помилок, %	Значення показника асиметрії
К найближчих сусідів	53450	188 / 812	225 / 775	20,65	1,20
К випадкових сусідів	15	181 / 819	494 / 506	33,75	2,73
Центрів класів	14	147 / 853	277 / 723	21,2	1,88
Адаптивних метрик	0	109 / 891	104 / 896	10,65	0,95

Аналіз даних у табл. 3 дозволяє зробити певні висновки у даному варіанті моделювання:

- найкращі результати за усіма показниками дає алгоритм адаптивних правил;
- алгоритм К найближчих сусідів ($K=1$) та алгоритм центрів класів дають приблизно одні й ті ж результати за кількістю помилок, але алгоритм центрів класів показує суттєво більшу ефективність по критерію продуктивності;
- алгоритм К випадкових сусідів практично дорівнює алгоритму центрів класів, оскільки використовує ті ж самі правила обчислення метрик, але за іншими показниками поступається всім іншим алгоритмам: помилки типу В/А навіть скоріше визначити підкиданням монетки. Також зауважимо, що цей алгоритм чутливий до кількості К сусідів: при збільшенні К пропорційно зростатиме час прийняття рішень.

Програма Adaptive Metrics дозволяє вирішувати задачі аналізу впливу розподільних властивостей окремих факторів на кількісні показники якості класифікації. Розглянемо лише три варіанти:

- класифікація здійснюється лише з врахуванням двох факторів F1 та F2, розподілення для обох класів і для обох факторів однакові, функції розподілення обрано за варіантом 2R (трикутні функції розподілення ймовірностей), варіант розподільної властивості «Середня»;
- варіант тих самих співвідношень по факторам F1 та F2, але додається ще третій фактор F3 з такими ж характеристиками (рис. 8, а);
- третій варіант відрізняється тим, що розподілення для класу В має вдвічі більший розмах значень ($SB=2$), а також розподільна властивість цього фактору встановлена як «Дуже низька» (рис. 8, б).



Результати прогонів імітаційної моделі дано у табл. 4, де кількість помилок дана як типові значення з протоколів за формою табл. 2.

Таблиця 4.
Помилки внаслідок впливу фактору з низькою розподільною властивістю

Алгоритм класифікації	Варіант налаштування імітаційної моделі					
	2 перші фактори		3 фактори, варіант а)		3 фактори, варіант б)	
	A/B	B/A	A/B	B/A	A/B	B/A
К найближчих сусідів	9	11	2	1	7	11
Центрів класів	6	4	1	0	3	5
Адаптивних правил	5	3	1	0	16	19

Аналіз даних у табл. 4 дозволяє зробити певні висновки з урахуванням того, що кількість помилок у даному випадку дуже мала, але:

- додавання інформативних факторів з хорошою розподільною властивістю призводить до підвищення надійності класифікації;
- додавання факторів з низькою розподільною властивістю може призводити до зниження надійності класифікації;
- найкращі результати за критерієм надійності у даному випадку дає алгоритм центрів класів;
- алгоритм адаптивних правил найбільш чутливий до включення в процес класифікації факторів з низькою розподільною властивістю.

З цих висновків витікає не дуже тривіальне твердження: алгоритм адаптивних правил завдяки його найвищій продуктивності може ефективно використовуватись у алгоритмах штучного інтелекту для попереднього відбракування «поганих факторів» (з низькою розподільною властивістю).

У даній статті розглянуто лише обмежену кількість варіантів налаштування імітаційної моделі, втім наведені приклади дозволяють зробити певні висновки стосовно застосування різних алгоритмів класифікації саме для випадків Big Data.

Висновки

1. Отримані результати дозволяють підтвердити основний висновок у статті [4:]: різні методи (алгоритми) класифікації заздалегідь не є «поганими» чи «добрими», оскільки практика їх застосування потребує у кожному випадку вибору найбільш ефективного алгоритму для даної конкретної задачі.

2. При відносно невеликих об'ємах даних (кількість класів – одиниці або десятки, кількість елементів класів - тисячі) кращі результати за критерієм надійності дає алгоритм К найближчих

сусідів, але при зростанні об'ємів даних ефективність даного алгоритму нівелюється за критерієм продуктивності.

3. Ще одною перевагою алгоритмів K найближчих сусідів є відповідність принципу компактності помилок: помилкові рішення найбільше концентруються в межах гіперкубу помилок.

4. Алгоритми, засновані на принципах M -means дещо поступаються алгоритмам K найближчих сусідів за критерієм надійності, але суттєво переважають за критерієм продуктивності.

5. Алгоритм адаптивних правил не має конкурентів за критерієм продуктивності, оскільки виконує лише найбільш швидкі операції порівняння значень факторів тестових об'єктів з вирішальними границями.

6. Ще однією перевагою алгоритму адаптивних правил є відповідність принципу рівноправності класів: навіть у дуже складних для класифікації умовах частота помилок типу A/B приблизно дорівнює частоті помилок типу B/A.

7. Втім, алгоритм адаптивних правил найбільш чутливий до включення в обробку факторів з низькою розподільною властивістю. Даний факт можна використати для алгоритмів автоматизації відбракування малоінформативних факторів.

8. Окремо відмітимо можливість застосування розробленого програмного забезпечення і подібних розробок для вирішення дидактичних задач. Розроблене програмне забезпечення завдяки значній варіабельності налаштувань та наочності відображення результатів дозволяє вирішувати різні навчальні задачі:

- демонструвати особливості різних алгоритмів класифікації для вирішення задач класу Big Data на лекційних заняттях;

- сприяти розвитку абстрактного мислення: здобувачі освіти за 121 спеціальністю «Інженерія програмного забезпечення» та 122 спеціальністю «Комп'ютерні науки» повинні розвивати абстрактне мислення – навчитись мислити у багатомірних просторах;

- оскільки лише варіантів налаштування імітаційної моделі 64800, то якщо до цього додати варіанти вибору алгоритмів (5 різних), варіанти кроку зростання кількості елементів базових класів (100 або 1000), варіанти вибору метрик (Евклідова чи метрика Хемінга), то загальна кількість варіантів досліджень буде 1 296 000, тобто здобувачам освіти буде чим зайнятися на парах лабораторних занять.

Література

1. Марченко О.О., Россада Т. В. Актуальні проблеми Data Mining: навч. посібн. для студентів факультету комп'ютерних наук та кібернетики. Київ: КНУ, 2017. 150 с.
2. Чубукова І.О. Data Mining. Київ: КНЕУ ім. В. Гетьмана, 2016. 326 с. URL: http://kist.ntu.edu.ua/textPhD/Chubukova-Data_Mining.pdf
3. Кромен Т., Лейзерсон Ч., Ривест Р. Алгоритмы: построение и анализ / пер. с англ. под ред. А. Шеня. М.: МЦНМО, 2002. – 960 с.
4. Одогов М.А., Гаджиев М.М., Буката Л.М., Глазунова Л.В., Кочеткова М.В. Обґрунтування швидких алгоритмів класифікації на множинах Big Data за критеріями надійності і продуктивності. *Інфокомунікаційні та комп'ютерні технології*. Київ, 2023. No 1 (стаття у даному випуску журналу).
5. Буртняк І.В. Імітаційне моделювання. Івано-Франківськ: ПНУ ім. В. Стефаника, 2019. 97 с.
6. Nicholas; Metropolis. The Monte Carlo Method. *Journal of the American Statistical Association*, 1949. Vol. 44, no. 247. P. 335—341. ISSN 0162-1459. DOI: [doi:10.2307/228023](https://doi.org/10.2307/228023).
7. Andrea Parenti, Giovanni Cappelli, Walter Zegada-Lizarazu, Carlos Martín Sastre, Myrsini Christou, Andrea Monti, Fabrizio Ginaldi. SunnGro: A new crop model for the simulation of sunn hemp (*Crotalaria juncea* L.) grown under alternative management practices. *Biomass and Bioenergy*, Vol. 146, 2021, 105975, ISSN 0961-9534. DOI: <https://doi.org/10.1016/j.biombioe.2021.105975>.
8. Mayuri Rajput, Mostafa Reisi Gahrooei, Godfried Augenbroe. A statistical model of the spatial variability of weather for use in building simulation practice. *Building and Environment*, Vol. 206, 2021, 108331, ISSN 0360-1323. DOI: <https://doi.org/10.1016/j.buildenv.2021.108331>.
9. A. Araya, P.V.V. Prasad, I.A. Ciampitti, P.K. Jha. Using crop simulation model to evaluate influence of water management practices and multiple cropping systems on crop yields: A case study for Ethiopian highlands. *Field Crops Research*, Vol. 260, 2021, 108004, ISSN 0378-4290. DOI: <https://doi.org/10.1016/j.fcr.2020.108004>.
10. El Arbi Abdellaoui Alaoui, Stephane Cedric Koumetio Tekouabou, Yassine Maleh, Anand Nayyar. Corrigendum to "Towards to intelligent routing for DTN protocols using machine learning techniques. *Simulation*

Modelling Practice and Theory. Vol. 123, 2023, 102689, ISSN 1569-190X. DOI: <https://doi.org/10.1016/j.simpat.2022.102689>.

11. Jean-Charles Mare. Best practices for model-based and simulation-aided engineering of power transmission and motion control systems. *Chinese Journal of Aeronautics*. Vol. 32, Issue 1, 2019, P. 186-199, ISSN 1000-9361. DOI: <https://doi.org/10.1016/j.cja.2018.07.015>.

12. Tian Zhao, Xi Chen, Ke-Lun He, Qun Chen. A standardized modeling strategy for heat current method-based analysis and simulation of thermal systems. *Energy*, Vol. 217, 2021, 119403, ISSN 0360-5442. DOI: <https://doi.org/10.1016/j.energy.2020.119403>.

13. Закон распределения суммы двух случайных величин. URL: https://bstudy.net/717926/sotsiologiya/zakon_raspredeleniya_summy_dvuh_sluchaynyh_velichin_kompozitsiya_dvuh_zakonov_raspredeleniya.

14. Харкевич А.А. Спектры и анализ. М.: «ЛИБКОМ», 2009. 240 с.

15. Ширяев А.Н. Вероятность. М.: Наука, 1980. 576 с.

References

1. Marchenko O.O., Rossada T.V. Aktualni problemi Data Mining: navch. posibn. dlya studentiv fakultetu komp'yuternih nauk ta kibernetiki. Kiyiv: KNU, 2017. 150 s.

2. Chubukova I.O. Data Mining. Kiyiv: KNEU im. V. Getmana, 2016. 326 s. URL: http://kist.ntu.edu.ua/textPhD/Chubukova-Data_Mining.pdf

3. Kromen T., Lejzerson Ch., Rivest R. Algoritmy: postroenie i analiz / per. s angl. pod red. A. Shenya. M.: MCNMO, 2002. – 960 s.

4. Odegov M.A., Hadzhyiev M.M., Bukata L.M., Glazunova L.V., Kochetkova M.V. Obgruntuvannya shvidkikh algoritmiv klasifikatsiyi na mnozhinah Big Data za kriteriyami nadijnosti i produktivnosti. *Infokomunikacijni ta komp'yuterni tehnologiyi*. Kiyiv, 2023. No 1 (stattya u danomu vipusku zhurnalu).

5. Burtnyak I.V. Imitacijne modelyuvannya. Ivano-Frankivsk: PNU im. V. Stefanika, 2019. 97 s.

6. Nicholas; Metropolis. The Monte Carlo Method. *Journal of the American Statistical Association*, 1949. Vol. 44, no. 247. P. 335—341. ISSN 0162-1459. DOI: [doi:10.2307/228023](https://doi.org/10.2307/228023).

7. Andrea Parenti, Giovanni Cappelli, Walter Zegada-Lizarazu, Carlos Martin Sastre, Myrsini Christou, Andrea Monti, Fabrizio Ginaldi. SunnGro: A new crop model for the simulation of sunn hemp (*Crotalaria juncea* L.) grown under alternative management practices. *Biomass and Bioenergy*, Vol. 146, 2021, 105975, ISSN 0961-9534. DOI: <https://doi.org/10.1016/j.biombioe.2021.105975>.

8. Mayuri Rajput, Mostafa Reisi Gahrooei, Godfried Augenbroe. A statistical model of the spatial variability of weather for use in building simulation practice. *Building and Environment*, Vol. 206, 2021, 108331, ISSN 0360-1323. DOI: <https://doi.org/10.1016/j.buildenv.2021.108331>.

9. A. Araya, P.V.V. Prasad, I.A. Ciampitti, P.K. Jha. Using crop simulation model to evaluate influence of water management practices and multiple cropping systems on crop yields: A case study for Ethiopian highlands. *Field Crops Research*, Vol. 260, 2021, 108004, ISSN 0378-4290. DOI: <https://doi.org/10.1016/j.fcr.2020.108004>.

10. El Arbi Abdellaoui Alaoui, Stephane Cedric Koumetio Tekouabou, Yassine Maleh, Anand Nayyar. Corrigendum to "Towards to intelligent routing for DTN protocols using machine learning techniques. *Simulation Modelling Practice and Theory*. Vol. 123, 2023, 102689, ISSN 1569-190X. DOI: <https://doi.org/10.1016/j.simpat.2022.102689>.

11. Jean-Charles Mare. Best practices for model-based and simulation-aided engineering of power transmission and motion control systems. *Chinese Journal of Aeronautics*. Vol. 32, Issue 1, 2019, P. 186-199, ISSN 1000-9361. DOI: <https://doi.org/10.1016/j.cja.2018.07.015>.

12. Tian Zhao, Xi Chen, Ke-Lun He, Qun Chen. A standardized modeling strategy for heat current method-based analysis and simulation of thermal systems. *Energy*, Vol. 217, 2021, 119403, ISSN 0360-5442. DOI: <https://doi.org/10.1016/j.energy.2020.119403>.

13. Закон распределения суммы двух случайных величин. URL: https://bstudy.net/717926/sotsiologiya/zakon_raspredeleniya_summy_dvuh_sluchaynyh_velichin_kompozitsiya_dvuh_zakonov_raspredeleniya.

14. Харкевич А.А. Спектры и анализ. М.: «ЛИБКОМ», 2009. 240 с.

15. Ширяев А.Н. Вероятность. М.: Наука, 1980. 576 с.

УДК 004.9

ОБҐРУНТУВАННЯ ШВИДКИХ АЛГОРИТМІВ КЛАСИФІКАЦІЇ НА МНОЖИНАХ BIG DATA ЗА КРИТЕРІЯМИ НАДІЙНОСТІ І ПРОДУКТИВНОСТІ

DOI 10.36994 / 2788-5518-2023-01-05-16

Одегов М.А., к.т.н., доц. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна. onick_64@ukr.net

Гаджисев М.М., д.т.н., проф. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна. gadjievmm@ukr.net

Бука́та Л.М. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна. ygrikluda@gmail.com

Глазунова М.В., к.ф-м.н. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна. lvglazun@gmail.com

Кочеткова М.В. Державний університет інтелектуальних технологій і зв'язку, Одеса, Україна. jubdvg@gmail.com

Анотація: У роботі розглянуто швидкі алгоритми класифікації, що можуть бути застосовані для розв'язку задач типу Big Data. Розглянуті алгоритми засновані на принципах M-means. При цьому всі можливі представники класів замінюються невеликою кількістю характеристик цих класів. Приведено теоретичне обґрунтування алгоритмів адаптивних метрик та адаптивних правил. Показано принципи реалізації даних алгоритмів для розв'язку досить широкої сфери практичних задач. Наведено приклад порівняння ряду алгоритмів класифікації за критеріями надійності та продуктивності.

Ключові слова: швидкі алгоритми класифікації, множини елементів, центри класів, метрики, розподілення ймовірностей, вирішальні правила, показники надійності та продуктивності

JUSTIFICATION OF FAST CLASSIFICATION ALGORITHMS ON BIG DATA SETS WITH RELIABILITY AND PERFORMANCE CRITERIA

Nick Odegov, Ph.D., Ass. Prof. State University of Intellectual Technologies and Communication, Odesa, Ukraine. onick_64@ukr.net

Matin Hadzhyiev, Dr. habil., Prof. State University of Intellectual Technologies and Communication, Odesa, Ukraine. gadjievmm@ukr.net

Liudmyla Bukata. State University of Intellectual Technologies and Communication, Odesa, Ukraine. ygrikluda@gmail.com

Liudmyla Glazunova, Ph.D., Ass. Prof. State University of Intellectual Technologies and Communication, Odesa, Ukraine. lvglazun@gmail.com

Marina Kochetkova. State University of Intellectual Technologies and Communication, Odesa, Ukraine. jubdvg@gmail.com

Abstract: Classification methods are among the simplest and "oldest" methods of artificial intelligence. This article discusses fast algorithms that can be used to solve Big Data problems. Big Data is a case when the methods and tools used do not allow solving the problem in a pleasant time. Therefore, when solving this type of problem, an important criterion is the performance of classification algorithms. Productivity in this sense refers to potential decision-making time. This time depends on the constructive dimension of the algorithm - the number of typical operations for making a decision. In addition, the execution time of the program depends on the typical operations of the algorithm. The article considers productive algorithms based on M-means principles. At the same time, all possible representatives of classes are replaced by a small number of characteristics of these classes. In the simplest form, these algorithms boil down to the fact that classes are replaced by class centers based on the results of training. Such centers are defined as vectors of average values over all one-dimensional projections of the factor space. Distance measures from these centers in Euclidean and other metric spaces are used to classify unknown objects. Methods of rapid adaptation of these characteristics in learning algorithms are considered. The theoretical justification of the algorithms of adaptive metrics and adaptive rules is given. It is shown that M-means algorithms can be effective in the case that the number of class representatives is extremely large, and the number of classes and the dimension of the factor space are relatively small. The advantages and disadvantages of the considered algorithms are noted. The scope of their practical application is outlined. The

principles of spreading these algorithms to a wide range of practical problems are also shown. An example of a comparison of a number of classification algorithms based on reliability and performance criteria is given. It was concluded that the algorithm of adaptive rules is the most effective for a significant number of problems.

Key words: fast classification algorithms, element sets, class centers, metrics, probability distribution, decision rules, reliability and performance indicators

Вступ

Поняття Big Data («великі дані») є невизначеним, умовним. У даній роботі ми дотримуємось ніби інтуїтивного розуміння цього поняття як ситуацію, коли наявні алгоритми або технічні засоби не дозволяють вирішити конкретну задачу за прийнятний час. Наприклад, дихотомія «свій - чужий» потребує часу вирішення у долі секунди. В іншій ситуації, наприклад при вирішенні проблеми: що вносить найбільший вклад у потепління на Землі «діяльність людини чи природні циклічні процеси» – прийнятним часом рішення може бути зовсім інший, більш тривалий час. В будь-якому випадку якість алгоритмів у випадку Big Data слід визначати за двома основними показниками: надійністю (частотою помилок класифікації) та продуктивністю (часом вирішення задачі).

Наша *перша парадигма* полягає в наступному: у випадку Big Data не існує універсального алгоритму, однаково ефективного для різних випадків. Тому розглянемо лише кілька варіантів таких, найбільш типових випадків.

Перший варіант: надзвичайно велика кількість класів, мінімальна кількість представників (елементів) у кожному класі, невелика кількість характеристик (розмірність фактор-простору) елементів. Типовий приклад: дактилоскопія – скільки людей на світі, стільки ж різних відбитків пальців.

Другий варіант: дуже велика кількість класів, можливо велика кількість елементів класів, відносно невелика кількість факторів. Приклад: встановлення особистості по фотографії. Тут розпізнаване обличчя відноситься до певного класу «двійників», а класів таких може бути багато.

Третій варіант: середня кількість класів, значна кількість елементів у кожному класі, середня кількість факторів. Приклад: розподіл громадян за неформальним критерієм якості життя, який ми спростимо до дихотомії: «задовільно - не задовільно». Ясна річ, що суб'єктивна оцінка якості життя може залежати не тільки від доходів. Як кажуть: «У когось борщ рідкий – у когось діаманти маленькі». Тому у сукупність факторів треба додати наявність житла, сімейний стан, стан медичного забезпечення та т.ін. Якщо населення вже визначилось на деякий час та розподілилось на «задоволених» та «незадоволених», то маємо базові набори представників класів типу А і типу В. У певний момент умовна держава щось змінює у правилах існування суспільства. Яка буде при цьому динаміка розподілу між «задоволеними» та «незадоволеними»?

Саме для третього варіанту і можуть бути ефективними швидкі алгоритми, що розглядаються у даній статті.

На даний час саме на задачах Data Mining в умовах Big Data зосереджуються значні зусилля світової наукової спільноти. В роботах [1,2] наведено огляд методів класифікації. При цьому в [2] є положення, що будь-який алгоритм для вирішення задач Big Data повинен бути «однопрохідним» тобто не застосовувати вкладених циклів у процедури прийняття рішень. Для радикального вирішення проблеми Р. Белмана «прокляття розмірності» в умовах Big Data пропонуються різні прийоми [3-8].

Одним з типових є розв'язок задач не на загальній множині елементів класів, а лише на вибіркових даних [3]. В алгоритмах типу К-найближчих сусідів пропонується обмежені множини тих самих «найближчих сусідів» визначати ще на етапі навчання алгоритму, як сукупність елементів, що близькі до границь класів [4].

Також, ефективним і достатньо природним є принцип скорочення розмірності фактор-простору за результатами попереднього навчання [5] або встановлення залежностей між факторами [6]. Втім, ще одним способом суттєвого зменшення обсягу обчислювань є заміна всіх множин елементів класів характерними представниками даних класів. Історично першим методом, що реалізує даний спосіб був метод M-means (у звичайному розумінні M центрів класів, але дослівно M засобів), який використовується як у задачах класифікації, так і у задачах кластеризації [7]. При цьому відмічають й недоліки даного методу, але розвиток алгоритмів його реалізації продовжується завдяки надзвичайної продуктивності цього методу [8]. У даній роботі

принцип M-means розуміється дещо ширше: на етапі навчання для класів визначається дуже обмежена кількість характеристик, які дозволяють швидко розв'язати задачу класифікації.

Метою досліджень є обґрунтування швидких алгоритмів класифікації, що засновані на принципах M-means Data Mining і відповідають наступним неформальним вимогам:

- дуже незначний час навчання;
- дуже мала кількість характеристик класів;
- дуже малий час прийняття рішень;
- достатня надійність для вирішення певних типів задач класифікації.

Виконання досліджень

1. Формалізація моделі класів

Класи з умовними номерами 1, 2, ..., M задаються базовими послідовностями елементів в одному й тому ж координатному просторі (фактор-просторі) розмірності D. Кожен елемент кожного класу визначається вектором значень факторів всього фактор-простору. Таким чином, базова модель задається сукупністю множин:

$$F_{m,n} = \{[f_{m,1}, f_{m,2}, \dots, f_{m,D}](n)\}, m = 1, 2, \dots, M, n = 1, 2, \dots, N_k, \quad (1)$$

де n – номер елементу у класі m .

Зауважимо, що кількість елементів у різних класах може значно відрізнятися, втім в усіх класах кількість елементів є дуже великою: умовно десятки тисяч, мільйонів або мільярдів в залежності від розмірності фактор-простору. Визначимо також додаткові умови та обмеження, як формальні, так і неформальні:

- розподілення значень факторів для всіх класів приблизно відповідають формальним умовам закону великих чисел [9], тобто є статистично та динамічно стійкими;
- накопичення даних у базових послідовностях виконується з урахуванням попередньо вирішеної задачі кластеризації, тобто класи співпадають з кластерами;
- всі фактори вважаються рівноправними з точки зору прикладної задачі;
- всі класи вважаються рівноправними з точки зору прикладної задачі.

Дві останні умови формалізують нашу *другу парадигму*: ніякий машинний алгоритм не може відповідати за прийняття рішень, що є прерогативою людини, оскільки людина має інтереси та потреби, а машина – ні. Наприклад, якщо вже встановлено на рівні законодавства, що розмір дорожнього збору залежить від об'єму двигуна, то навіщо крім цього фактора включати у процедури класифікації колір автомобіля, тип підвіски, діаметр коліс та ін.? Якщо класифікація «пропуск цілі чи хибна тривога» потребує як можливо зменшувати кількість саме пропусків цілей, то який формальний алгоритм може визначити вагові коефіцієнти помилок першого та другого роду? У таких задачах відповідальність залишається за людиною.

2. Задача дихотомії: алгоритм M-центрів

У найпростішому випадку класифікатор приймає рішення віднесення невизначеного об'єкту X до одного з класів: A або B. При цьому об'єкт X задано у тому ж самому фактор-просторі, що і елементи класів A і B. Рішення класифікатора можуть бути 4-х типів, які ми умовно позначимо:

- A=A – вірне рішення віднесення об'єкту X до класу A;
- B=B – вірне рішення віднесення об'єкту X до класу B;
- B/A – помилка 1-го роду: прийнято рішення віднесення об'єкту X до класу B, тоді як об'єкт відноситься до класу A;
- A/B – помилка 2-го роду: прийнято рішення віднесення об'єкту X до класу A, тоді як об'єкт відноситься до класу B.

В алгоритмах навчання дані варіанти рішень зручно асоціювати з умовними номерами:

$$(A=A) \Rightarrow 0; (B=B) \Rightarrow 1; (B/A) \Rightarrow 2; (A/B) \Rightarrow 3.$$

Відповідно, за результатами навчання можна приблизно визначити умовні частоти даних рішень. Зважаючи на велику кількість елементів класів та умову їх статистичної стійкості, в роботі не завжди будемо розрізняти поняття «ймовірності» та «частоти» (емпіричної ймовірності), до

яких будемо використовувати позначення P . Саме таким символом будемо позначати інтегральні функції розподілення, а позначкою p – диференціальні функції розподілення.

Таким чином, принцип рівноправності класів потребує хоча б приблизного виконання рівності: $P(B/A) \approx P(A/B)$. Зрозуміло, що при цьому обидві ці ймовірності треба мінімізувати.

У даному випадку самий швидкий алгоритм може бути побудований за принципами методу M -means (дослівно: M -засобів). Кількість таких «засобів» пропорційна кількості класів, а самими засобами можуть бути обмежені набори характеристик класів, вирішальних правил та т.ін.

Тут треба зробити термінологічне зауваження. В літературі по Data Mining [7] міцно закріпились терміни « K -найближчих сусідів» та « K -means». При цьому мова йде зовсім про різні по суті числа K . У першому випадку – про кількість елементів класів, у другому – про кількість класів. Тому у даній роботі ми все-ж таки будемо користуватись у другому випадку терміном « M -means» (саме від слова «Means»).

У найпростішому вигляді алгоритм M -means зводиться до алгоритму M -представників класів (лише по одному на кожний клас) [7]. Тоді множини елементів (1) замінюються представниками класів:

$$F_m = [f_{m,1}, f_{m,2}, \dots, f_{m,D}], m = 1, 2, \dots, M, n = 1, 2, \dots, N_k, \quad (2)$$

де значення факторів відносяться вже не до окремих елементів, а до їх представників.

У задачі дихотомії обидва класи замість всієї базової сукупності елементів замінюються лише двома «центрами класів», які ми також позначимо іменами класів. Відповідно, вирішальне правило формулюється так:

$$\rho(A, X) < \rho(B, X) \Rightarrow 0; \quad \rho(A, X) > \rho(B, X) \Rightarrow 1, \quad (3)$$

де ρ – обрана метрика у фактор-просторі.

Метрики можуть обиратись різними способами [7]. Втім, і поняття центрів класів також має бути пов'язане з типом метрики. Якщо обрана Евклідова метрика, то центри класів логічно визначати як вектори середніх значень по факторам, оскільки середнє мінімізує суму квадратів відстаней від нього. Якщо обрана метрика Хемінга, то логічно визначати центри класів як вектори медіан значень по факторам, оскільки медіана мінімізує суму модулів відстаней від неї.

Медіани є робастними статистиками, оскільки слабо чутливі до великих «викидів» у вибірках. Якщо задачі вирішуються на відносно малих масивах даних, то перевагу слід віддати метричним просторам Хемінга. Втім, для Big Data процедури навчання для визначення медіан вже ніяк не можна вважати «однопрохідними», оскільки дані алгоритми потребують ранжирування вибірок. Тобто, процес навчання може займати значний час. На відміну від цього, визначення середніх при нарощуванні обсягів елементів класів вирішується ітераційними алгоритмами за один крок. Наприклад, для класу A при додаванні нового, $(N+1)$ -го елементу нове середнє по 1-му фактору розраховується за елементарною формулою:

$$A_{N+1,1} = \frac{1}{N+1} \left(f_{A,1}(n+1) + \sum_{n=1}^N f_{A,1}(n) \right) = \frac{1}{N+1} f_{A,1}(n+1) + \frac{N}{N+1} A_{N,1}. \quad (4)$$

Розглянемо найпростіший випадок дихотомії: класифікація в одновірному просторі. При цьому допускаємо, що функції розподілу ймовірностей для класів A і B відрізняються лише значеннями середніх (рис. 1.а).

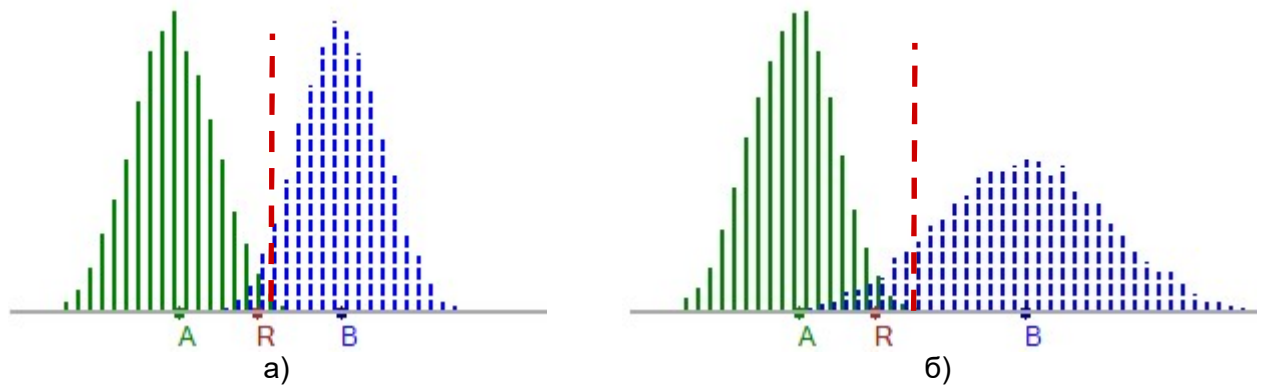


Рис. 1. Варіанти розподілів одного фактору для класів А і В

Такий алгоритм застосовується практично у задачі виявлення сигналів певного рівня на фоні шумів. Клас А – лише шум; Клас В – сигнал + шум. Втім, цим прикладом сфера застосування даного алгоритму і обмежується: навіть у 2-мірному просторі складно уявити ситуацію, коли два класи відрізняються лише середніми. Навіть у дуже простій ситуації, коли розподілення значень факторів відрізняються лише масштабом (мірою розсіювання відносно середнього), даний алгоритм буде давати очевидний перебік ймовірності помилок 2-го типу (рис. 1.б). Невідомі об'єкти будуть скоріше відноситись до класу А, ніж до класу В.

Також відомі інші недоліки даного алгоритму:

- чутливість до наявності факторів з сильною кореляцією: пара чи більше таких факторів будуть надавати найбільш суттєвий вплив на загальне рішення; цей недолік усувається на етапі попереднього аналізу даних – частина таких факторів або відбраковується, або дані фактори агрегуються у якусь сукупльну характеристику;

- велика чутливість до масштабів значень по факторам; для усунення цього недоліку використовується приведення факторів до єдиного масштабу.

Остання операція найпростіше і логічніше виконується, якщо функції розподілення значень факторів відрізняються лише масштабами: для цього метрики зважуються зворотно пропорційно стандартним відхиленням. В евклідовому просторі метрики щодо правил (3) уточнюються, як правило, у такому алгоритмі:

$$\begin{cases} \rho_{\kappa\delta}(A, X) = \sum_{d=1}^D \frac{1}{\sigma_d^2} (f_{A,d} - f_{X,d})^2; & \rho_e(A, X) = \sqrt{\rho_{\kappa\delta}(A, X)}; \\ \rho_{\kappa\delta}(B, X) = \sum_{d=1}^D \frac{1}{\sigma_d^2} (f_{B,d} - f_{X,d})^2; & \rho_e(B, X) = \sqrt{\rho_{\kappa\delta}(B, X)}; \end{cases} \quad (5)$$

де $\rho_{\kappa\delta}$ – квадратична метрика; ρ_e – евклідова метрика; σ_d – однакове для обох класів стандартне відхилення для фактору з номером d .

Таким чином, даний алгоритм поширюється на клас задач, де класи відрізняються середніми значеннями, але масштаби даних можуть бути різними для різних факторів.

Стосовно випадків Big Data зробимо необхідне зауваження, що крок обчислення саме Евклідової метрики (тобто, операція взяття кореню з квадратичної метрики) є не тільки зайвою, але й дуже шкідливою з точки зору продуктивності алгоритмів, оскільки така операція потребує значної кількості машинних тактів. Втім, ці метрики є еквівалентними у тому сенсі, що якщо: $\rho_{\kappa\delta}(A, X) = 0$, то $\rho_e(A, X) = 0$; якщо: $\rho_{\kappa\delta}(A, X) < \rho_{\kappa\delta}(B, X)$, то $\rho_e(A, X) < \rho_e(B, X)$. Доказ цього положення опускаємо за очевидністю.

Також, зауважимо, що при обробці Big Data досвідчений програміст ніколи не буде використовувати функції піднесення у другу ступень тієї чи іншої мови програмування типу X^2 або $\text{power}(X, 2)$. Як правило, він завжди використає замість цього простий вираз типу $X * X$.

Таким чином, результати навчання зводяться до зберігання у пам'яті лише 2D значень середніх по факторам класів та ще D значень стандартних відхилень. При цьому алгоритми навчання є такими, що оцінки цих значень сходяться зі швидкістю $1/N$, що відомо з факту збіжності емпіричних середніх незалежних іспитів до математичних очікувань [9].

3. Задача дихотомії: алгоритм адаптивних метрик

Розглянемо швидкий алгоритм для вирішення задачі дихотомії у випадку, що показано на рис. 1.6. Спочатку обмежимося випадком лише одного фактору. Теоретична задача: знайти оптимальну межу між класами A і B, яка буде відповідати принципу рівноправності класів у тому сенсі, що ймовірність помилок $P(A/B)$ та $P(B/A)$ буде однаковою.

У даному випадку встановимо вирішальну границю R між класами (рис. 2). Оптимальне значення R буде таким, що ймовірності помилок відповідатимуть рівності:

$$\frac{P(B < R)}{P(A < R)} = P(A/B) = P(B/A) = \frac{P(A > R)}{P(B > R)}. \quad (6)$$

Оскільки події $P(B < R)$ та $P(B > R)$ є несумісні та утворюють повну групу (аналогічно для подій $P(A < R)$ та $P(A > R)$), то рівності (6) можна представити у еквівалентній формі:

$$\frac{P(B < R)}{1 - P(A > R)} = \frac{P(A > R)}{1 - P(B < R)}. \quad (7)$$

Із співвідношення (7) витікає, що для виконання умови рівності помилок 1-го та 2-го роду достатньо рівності чисельників, оскільки рівність знаменників слідує з цього.

Нехай у даному фактор-просторі задана якась симетрична диференційна функція розподілення $p(x)$ з нульовим середнім та стандартним відхиленням σ_X . Функції розподілення класів A і B відрізняються від $p(x)$ зсувом та масштабом, тобто:

$$p_A(x) = p(q_A x - f_A); \quad p_B(x) = p(q_B x - f_B). \quad (8)$$

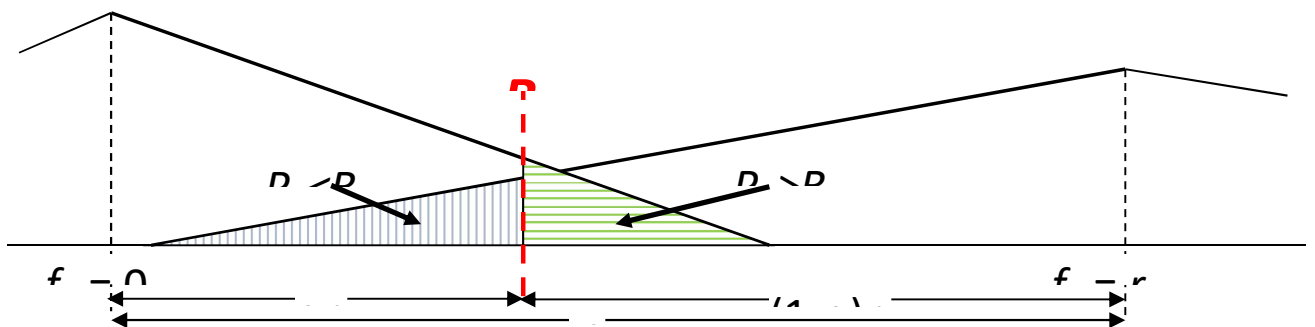


Рис. 2. Розподілення однакової форми, але з різними масштабами

Оскільки всі перетворення (8) є лише перетвореннями зсуву та масштабування, то без втрат спільності будемо вважати, що: $\sigma_X = 1$, $f_A = 0$, $f_B = r > 0$.

Можна довести, що ймовірність виходу випадкової величини X за межу « q сигм» від центру f залежить лише від виду функції розподілення та від числа q і не залежить від масштабу. Тоді із умов (7) витікає співвідношення:

$$q = \frac{\alpha r}{\sigma_A} = \frac{(1-\alpha)r}{\sigma_B} \Rightarrow \alpha r \sigma_B = \sigma_A - \alpha r \sigma_A \Rightarrow \alpha = \frac{\sigma_A}{\sigma_A + \sigma_B}; 1 - \alpha = \frac{\sigma_B}{\sigma_A + \sigma_B}. \quad (9)$$

Оптимальна межа буде визначатись як $R = \alpha r$, а вирішальне правило теж саме, як і у правилі (3), але відстані від центрів класів розраховуються наступним чином:

$$\rho_{\kappa\delta}(A, X) = \sum_{d=1}^D \frac{1}{\sigma_{A,d}^2} (f_{A,d} - f_{X,d})^2; \quad \rho_{\kappa\delta}(B, X) = \sum_{d=1}^D \frac{1}{\sigma_{B,d}^2} (f_{A,d} - f_{X,d})^2. \quad (10)$$

Порівняно з алгоритмом по п. 2 даний алгоритм призводить до необхідності додати лише значення стандартних відхилень для обох класів по всім факторам. Тобто для дихотомічного класифікатора треба зберігати в оперативній пам'яті 2D значень середніх та 2D значень стандартних відхилень.

Адаптація алгоритму за даними навчальних вибірок також може здійснюватися ітераційними процедурами типу (4). При цьому знов такі оцінки всіх параметрів виконуються з високою швидкістю порядку $1/N$.

Втім, порівняно з алгоритмом по п. 2 даний алгоритм може бути використаний до більш широкого класу практичних задач. Так, наприклад ріст людей може бути не тільки у середньому різнити для різних народів, але й відрізнятися значеннями стандартних відхилень.

Окрім цього, даний алгоритм також може застосовуватись у випадках, коли розподілення по факторам відрізняються не тільки зсувом та масштабом. З де якими уточненнями даний алгоритм можна застосовувати і для більш широкого класу практичних задач. Виходячи з нерівності Чебишева [9], можна вважати, що даний алгоритм буде давати невелику додаткову кількість помилок, якщо розподілення все-ж таки відрізняються за формою, але несуттєво.

4. Задача дихотомії: алгоритм адаптивних правил

У даному алгоритмі за результатами навчання заздалегідь встановлюються вирішальні границі між класами по окремим факторам згідно виразів (9). Нехай середні значення по 1-му фактору для класів А і В розміщено саме так, як на рис. 2, тобто в середньому фактор класу А менший за клас В ($f_{B,1} > f_{A,1}$). Тоді вирішальна границя буде в абсолютних значеннях знаходитись у точці

$$R_1 = f_{A,1} + \alpha r = f_{A,1} + \frac{\sigma_{A,1}(f_{B,1} - f_{A,1})}{\sigma_{A,1} + \sigma_{B,1}}. \quad (11)$$

Тоді для прийняття рішення по даному фактору можна використовувати бінарну метрику:

$$\begin{cases} f_{X,1} < R_1 \Rightarrow \rho_1(A, X) = 0, \rho_1(B, X) = 1 \\ f_{X,1} > R_1 \Rightarrow \rho_1(A, X) = 1, \rho_1(B, X) = 0 \end{cases} \quad (12)$$

по всій сукупності факторів визначаються відстані:

$$\rho(A, X) = \sum_{d=1}^D \rho_d(A, X), \quad \rho(B, X) = \sum_{d=1}^D \rho_d(B, X), \quad (13)$$

а рішення приймається за тим самим правилом (3).

Зауважимо, що операція порівняння у виразах (12) виконується за найшвидшою машинною процедурою. Дійсно, для рішення треба визначити лише знак різниці $(R_k - f_{X,1})$, тобто визначити значення біту знака: якщо 0, то різниця позитивна, якщо 1 – негативна.

Тому даний алгоритм не потребує навіть операцій множення і є найшвидшим з попередньо розглянутих. Для задачі дихотомії за результатами навчання визначається при цьому лише пара чисел для кожного фактору:

$$[f_{A,d}, R_d], \quad d = 1, 2, \dots, D.$$

Зрозуміло, що у випадку, якщо $f_{B,d} < f_{A,d}$ (центри класів міняються місцями порівняно з рис. 2), то змінюється і визначення вирішальної границі:

$$R_d = f_{B,d} + \alpha r = f_{B,1} + \frac{\sigma_{B,1}(f_{A,1} - f_{B,1})}{\sigma_{A,1} + \sigma_{B,1}} \quad (14)$$

та відповідно змінюються визначення метрик у виразах (12) на протилежні.

Зауважимо, що даний алгоритм дає однозначне рішення у випадку непарної кількості факторів. Також зауважимо, що визначені метрики (11-13) задають над простором класів метричний простір, тобто є суто метриками у розумінні аксіом:

$$\begin{aligned} \rho(A, A) &= 0 & (\text{аксіома тотожності}) \\ \rho(A, B) &\geq 0 & (\text{аксіома позитивності}) \\ \rho(A, B) &= \rho(B, A) & (\text{аксіома симетрії}) \\ \rho(A, C) &\leq \rho(A, B) + \rho(B, C) & (\text{аксіома трикутника}) \end{aligned}, \quad (15)$$

що доводиться тривіально.

Таким чином, для даного алгоритму також визначаються параметри, що залежать лише від середніх значень та значень стандартних відхилень по факторам.

5. Узагальнення сфери застосування швидких алгоритмів типу *M-means*

Випадок нестрогих нерівностей. У виразах вище для визначення правил класифікації навмисно всюди використано лише строгі нерівності. Якщо можливі значення факторів по всім класам мають дуже високу варіабельність, то ймовірність таких випадків близька до 0. Проте, якщо ці фактори можуть приймати дуже обмежену кількість значень наприклад, лише 10, то при невеликій розмірності фактор-простору ймовірність рівності метрик може бути високою. У даному випадку всі ці алгоритми корегуються: визначається особливий тип можливих рішень: відмова від класифікації, оскільки даний класифікатор не може прийняти будь-якого рішення. У таких випадках відмови класифікації або відправляються на «ручну» обробку, або використовується окремий алгоритм виключно для таких випадків. Наприклад, якщо алгоритм адаптивних правил дає відмову від класифікації, то застосовується алгоритм адаптивних метрик.

Випадок складних деформацій функцій розподілення ймовірностей. У даному випадку розподілення значень по факторам для різних класів відрізняються не тільки деформаціями типу зсуву та зміни масштабу, а й самі функції розподілення можуть мати дуже різну форму для різних класів. Втім, в окремих типах задач і такий випадок можна передбачити. Нехай є дві **різні** інтегральні функції розподілення $P_A(x)$ та $P_B(x)$. І нехай вони у метриці Чебишева [9] відрізняються на якусь невелику величину:

$$\max_{-\infty < x < \infty} |P_A(x) - P_B(x)| < \varepsilon,$$

тоді ймовірність помилок 1-го та 2-го роду порівняно з методом по п.3 зростає не більше ніж на ϵ . Дане положення доводиться за допомогою нерівності Чебишева.

Випадок значної кількості класів. Тут передбачається, що кількість класів $M > 2$. Оскільки всі алгоритми за п.п. 2-5 є дуже швидкими, то навіть для значної кількості класів можливо їх використовувати у найпростішому варіанті дерев рішень [7].

Нехай задано класи A, B, C, D ... Обираємо якийсь алгоритм з п.п. 2-5 (в залежності від особливостей розподілень факторів у класах). Далі використовуємо послідовне рішення задач дихотомії:

- відносно «невідомого» об'єкту X приймаємо рішення, до якого класу він ближчий – до A або B;
- допустимо, що рішення прийняте на користь класу A;
- тоді клас B виключається з подальшої обробки;
- вирішується задача дихотомії між класами A та C;
- один з цих класів виключається з подальшої обробки;
- і так далі до вичерпання всіх класів.

Очевидно, що кількість задач дихотомії буде складати $M - 1$, тобто алгоритм класифікації залишається «однопрохідним». Коректність даного алгоритму дерев рішень витікає з того, що всі задачі дихотомії вирішуються в єдиному метричному просторі. Те, що не треба повертатися до попередніх рішень саме цим і гарантовано: саме виконанням аксіоми трикутника у виразах (15).

При значній кількості класів найшвидший алгоритм адаптивних правил може мати недолік – необхідність зберігати у пам'яті попарні значення вирішальних границь R (заздалегідь, числа з плаваючою точкою типу float, double та т.ін.). Для вибору між вирішальними границями (11) або (14) треба зберігати ще пари ознак S (sign) співвідношення середніх по факторам ($A > B$ або $B > A$) – дані типу boolean. Для прикладу 5-ти класів структура матриць даних лише для одного фактору представлена у табл. 1.

У даному випадку для кожного фактора, таким чином, треба зберігати 20 значень. Для фактор-простору розмірності 10 треба буде зберігати за результатами навчання вже 200 значень параметрів вирішальних правил. Тут знову ж таки проявляється характерний ефект Big Data – «прокляття розмірності», оскільки для M класів у просторі D факторів треба буде зберігати $D \times M \times (M-1)$ даних.

Тобто об'єм оперативної пам'яті для зберігання цих параметрів збільшується у квадратичній залежності від кількості класів. Наприклад, якщо є відносно невелика кількість класів $M = 10000$, які задано у фактор-просторі розмірності $D = 10$, то треба буде зберігати 999900000 (майже мільярд!) значень параметрів.

Таблиця 1.
Приклад структури матриць даних для одного фактору та для 5 класів

Класи	A	B	C	D	A	B	C	D
	Вирішальні границі				Ознаки рішень			
B	R_{AB}				S_{AB}			
C	R_{AC}	R_{BC}			S_{AC}	S_{BC}		
D	R_{AD}	R_{BD}	R_{CD}		S_{AD}	S_{BD}	S_{CD}	
E	R_{AE}	R_{BE}	R_{CE}	R_{DE}	S_{AE}	S_{BE}	S_{CE}	S_{DE}

Вихід з даного положення – розрахунок вирішальних границь під час виконання програми. Тоді кількість даних в оперативній пам'яті буде аналогічна до кількості даних алгоритму адаптивних метрик. При вирішенні конкретних задач, таким чином, треба буде приймати компромісне рішення між швидкістю та розмірами масивів параметрів класифікаторів.

Випадок динамічних класів. Дуже велика кількість практичних задач вирішується для класів, характеристики яких змінюються з часом. Втім, завдяки незначній кількості характеристик класів, методи M-means можуть бути розповсюджені і на даний випадок. Відмінність застосування полягає лише у тому, що під час навчання алгоритмів встановлюються не фіксовані значення

параметрів розподілень, а параметри їх динамічних моделей виду $f_{m,d}(t)$ та $\sigma_{m,d}(t)$. У найпростішому вигляді, наприклад, для короткострокової екстраполяції, дані залежності можна представити лінійними функціями:

$$f_{m,d}(t) = f_{0,m,d} + f_{1,m,d} \cdot t, \quad \sigma_{m,d}(t) = \sigma_{0,m,d} + \sigma_{1,m,d} \cdot t. \quad (16)$$

Задача класифікації у даному випадку вирішується за допомогою прогнозованих значень характеристик класів. При цьому для моделей залежностей виду (16) обсяг пам'яті порівняно з випадком статичних класів підвищується лише вдвічі.

У практичних задачах залежності характеристик класів від часу можуть бути більш складними, ніж лінійні функції (16). Для таких випадків можна застосовувати відомі методи штучного інтелекту, наприклад, метод групового врахування аргументів [10]: такі методи дозволяють як мінімізувати кількість вільних параметрів моделей, але й також вибрати структури моделей аз критеріями екстраполяційної стійкості.

6. Порівняння алгоритмів класифікації за показниками надійності та продуктивності

У даному пункті наведемо лише один приклад порівняння характеристик надійності та продуктивності деяких алгоритмів класифікації. Порівняння здійснюється у задачі дихотомії, де фактор-простір має розмірність $D = 6$. Розглядається, таким чином 2 класи: А (умовно зелений) та В (умовно синій). Задача вирішується за допомогою імітаційної моделі (більш докладно розглянуто у статті [11]).

Окрім алгоритмів адаптивних метрик та адаптивних правил також розглядається відомий алгоритм К найближчих сусідів та алгоритм К випадкових сусідів. Останній відрізняється тим, що з серед представників класів обирається незначна кількість випадкових елементів. Рішення приймається шляхом обчислення відстаней від невизначеного об'єкту Х до всіх К випадкових елементів по класам А і В. Порівнюються або суми відстаней, або середнє значення відстаней.

Приклад апіорних розподілень значень факторів наведено у вигляді гістограм в розрізі факторів (рис. 3). Тут червоним показано в умовному масштабі гістограми помилок.

Як бачимо, задача вирішується у дуже складних умовах. Масштаби значень для «зеленого» та «синього» класів відрізняються суттєво: вдвічі! Окрім того, половина значень «синього» класу перекриває всі можливі значення «синього» класу по всім факторам.

Задача вирішувалась на множинах елементів $N=10000$ для кожного класу. Тестування відбувалось на 1000 «невдомих» об'єктів для кожного класу. Результати класифікації для даного прикладу дано у табл. 2.

Аналіз даних у табл. 2 дозволяє зробити такі висновки:

- найменшу загальну кількість помилок дає все-ж таки алгоритм К найближчих сусідів;
- найбільшу кількість помилок дає алгоритм К випадкових сусідів;
- алгоритм К найближчих сусідів, К випадкових сусідів центрів класів у даному випадку дають «перекіс» частоти помилок у бік помилок 2-го типу: перевага частіше віддається класу А, хоча тестовий об'єкт насправді відносився до класу В;
- алгоритм адаптивних метрик та адаптивних правил у найбільший ступені забезпечують виконання принципу «рівноправності» класів;
- загальна кількість помилок алгоритмів К найближчих сусідів, адаптивних метрик та адаптивних правил мають приблизно однаковий порядок;
- що стосується часу прийняття рішень, то тут, як кажуть, «коментарі зайві»: алгоритми, засновані на принципах М-means суттєво більш ефективні за критерієм продуктивності;
- найбільш швидким алгоритмом із розглянутих є алгоритм адаптивних правил.

Таблиця 2.

Результати визначення характеристик надійності та продуктивності різних алгоритмів класифікації

Метод класифікації	Кількість K сусідів	Час рішення, мс	Похибки B/A	Похибки A/B	Загальний процент помилок
K найближчих сусідів	100	155560	2 / 998	109 / 891	5,55
K випадкових сусідів	100	1618	5 / 995	484 / 516	24,45
Центрів класів	1	11	12 / 988	137 / 863	7,45
Адаптивних метрик	1	16	61 / 939	58 / 942	5,95
Адаптивних правил	1	1	59 / 941	59 / 941	5,70

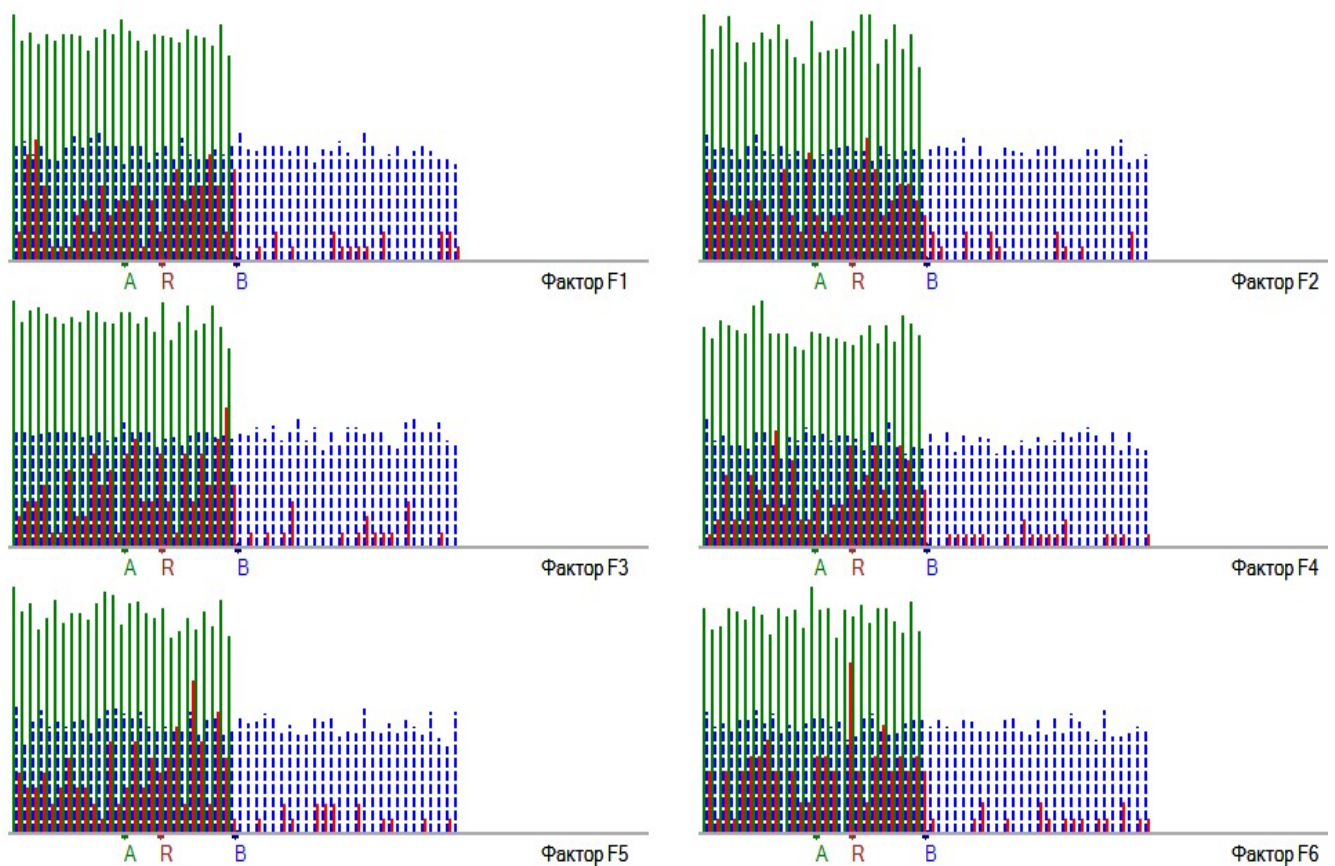


Рис. 3. Гістограми розподілу значень для 6-ти факторів

Різниця часу прийняття рішень всіма цими алгоритмами відрізняється, як бачимо, суттєво. Причини цього очевидні: алгоритм адаптивних метрик витрачає значно більше значення часу прийняття рішень, ніж адаптивних правил тому що використовує більш тривалі операції обчислення метрик – потребує операцій множення.

Алгоритми K найближчих сусідів та K випадкових сусідів використовують по суті ті ж самі метрики, що і алгоритм адаптивних метрик. Втім, кількість операцій підвищується приблизно пропорційно тій самій кількості K сусідів. Тому у табл. 2 і спостерігаються відповідні співвідношення: час прийняття рішень алгоритмом K випадкових сусідів приблизно у 100 разів більший, ніж час прийняття рішень алгоритмом адаптивних метрик. Відповідно, час прийняття рішень алгоритмом K найближчих сусідів вже у 10000 разів більший.

Висновки

1. Головний висновок полягає у тому, що різні методи (алгоритми) класифікації заздалегідь не є «поганими» чи «добrimi»: вирішення практичних задач потребує вибору найбільш ефективного алгоритму для даної конкретної задачі.
2. У задачах класифікації великих обсягів даних (Big Data) поряд з критерієм ефективності алгоритмів – надійності, слід обов'язково враховувати критерій продуктивності. Якщо по критерію надійності алгоритми приблизно рівні, то однозначно слід надавати перевагу більш продуктивним алгоритмам.
3. Для задач класифікації Big Data з надзвичайно великою кількістю представників класів, середньою кількістю класів, невеликою розмірністю фактор-простору значні переваги мають алгоритми, побудовані за принципами M-means. При цьому всі елементи класів замінюються незначною кількістю характеристик класів.
4. В простих випадках, оцінюючи ефективність за всіма критеріями, слід віддати перевагу алгоритмам адаптивних метрик або адаптивних правил. При цьому останній навіть при кількості класів порядку десятків тисяч може давати суттєвий вииграш у часі прийняття рішень, що видно вже з наведеного прикладу.
5. Принципи M-means дозволяють адаптувати алгоритми класифікації до дуже великої сфери практичних задач, у тому числі на випадок динамічних класів. Тому такі алгоритми мають перспективи широкого застосування для ефективного вирішення зазначеного типу задач, а продовження досліджень у цьому напрямку є актуальною науковою задачею.

Література

1. Pijush Kanti Dutta Pramanik, Saurabh Pal, Moutan Mukhopadhyay, Simar Preet. Big Data classification: techniques and tools. *Applications of Big Data in Healthcare*, Academic Press, 2021. P. 1-43.
2. Chenxia Jin, Fachao Li, Shijie Ma, Ying Wang. Sampling scheme-based classification rule mining method using decision tree in big data environment. *Knowledge-Based Systems*, Vol. 244, 2022.
3. Nooshin Hanafi, Hamid Saadatfar. A fast DBSCAN algorithm for big data based on efficient density calculation. *Expert Systems with Applications*, Vol. 203, 2022, 117501, ISSN 0957-4174.
4. Chaoyu Gong, Zhi-gang Su, Pei-hong Wang, Qian Wang, Yang You. Evidential instance selection for K-nearest neighbor classification of big data. *International Journal of Approximate Reasoning*, Vol. 138, 2021. Pages 123-144.
5. Ayman Taha, Ali S. Hadi, Bernard Cosgrave, Susan McKeever. A multiple association-based unsupervised feature selection algorithm for mixed data sets. *Expert Systems with Applications*, Vol. 212, 2023, 118718, ISSN 0957-4174..
6. Qian Tao, Chunqin Gu, Zhenyu Wang, Daoning Jiang. An intelligent clustering algorithm for high-dimensional multiview data in big data applications. *Neurocomputing*, Vol. 393, 2020, Pages 234-244, ISSN 0925-2312, комп'ютерних наук та кібернетики. Київ: КНУ, 2017. 150 с.
8. Abiodun M. Ikotun, Absalom E. Ezugwu, Laith Abualigah, Belal Abuhaija, Jia Heming. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, Vol. 622, 2023, Pages 178-210, ISSN 0020-0255..
9. Огірко О.І., Галайко Н.В. Теорія ймовірностей та математична статистика: навч. посібн. Львів: ЛьвДУВС, 2017. 292 с.
10. Ивахненко А.Г., Степашко В.С. Помехоустойчивость моделирования. Киев: Наукова думка, 1985. 216 с.
11. Одогов М.А., Гаджиев М.М., Буката Л.М., Глазунова Л.В., Кочеткова М.В. Порівняння алгоритмів класифікації Big Data методами імітаційного моделювання. *Інфокомунікаційні та комп'ютерні технології*. Київ, 2023. No 1 (стаття у даному випуску журналу).

References

1. Pijush Kanti Dutta Pramanik, Saurabh Pal, Moutan Mukhopadhyay, Simar Preet Singh. 1 - Big Data classification: techniques and tools. *Applications of Big Data in Healthcare*, Academic Press, 2021. P. 1-43.
2. Chenxia Jin, Fachao Li, Shijie Ma, Ying Wang. Sampling scheme-based classification rule mining method using decision tree in big data environment, *Knowledge-Based Systems*, Vol. 244, 2022.
3. Nooshin Hanafi, Hamid Saadatfar. A fast DBSCAN algorithm for big data based on efficient density calculation. *Expert Systems with Applications*, Vol. 203, 2022, 117501, ISSN 0957-4174.
4. Chaoyu Gong, Zhi-gang Su, Pei-hong Wang, Qian Wang, Yang You. Evidential instance selection for K-nearest neighbor classification of big data. *International Journal of Approximate Reasoning*, Vol. 138, 2021. Pages 123-144..
5. Ayman Taha, Ali S. Hadi, Bernard Cosgrave, Susan McKeever. A multiple association-based unsupervised feature selection algorithm for mixed data sets. *Expert Systems with Applications*, Vol. 212, 2023, 118718, ISSN 0957-4174..
6. Qian Tao, Chunqin Gu, Zhenyu Wang, Daoning Jiang. An intelligent clustering algorithm for high-dimensional multiview data in big data applications. *Neurocomputing*, Vol. 393, 2020, Pages 234-244, ISSN 0925-2312,

7. Marchenko O.O., Rossada T.V. Aktualni problemi Data Mining: navch. posibn. dlya studentiv fakultetu komp'yuternih nauk ta kibernetiki. Kiyiv: KNU, 2017. 150 s.
8. Abiodun M. Ikotun, Absalom E. Ezugwu, Laith Abualigah, Belal Abuhaija, Jia Heming. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, Vol. 622, 2023, Pages 178-210, ISSN 0020-0255.
9. Ogirko O.I., Galajko N.V. Teoriya jmovirnostej ta matematichna statistika: navch. posibn. Lviv: LvDUVS, 2017. 292 s.
10. Ivahnenko A.G., Stepashko V.S. Pomehoustojchivost modelirovaniya. Kiev: Naukova dumka, 1985. 216 s.
11. Odegov M.A., Hadzhyiev M.M., Bukata L.M., Glazunova L.V., Kochetkova M.V. Porivnyannya algoritmiv klasifikaciyi Big Data metodami imitacijnogo modelyuvannya. *Infokomunikacijni ta komp'yuterni tehnologiyi*. Kiyiv, 2023. No 1 (stattya u danomu vipusku zhurnalu).

УДК 004-658

МАТЕМАТИЧНА МОДЕЛЬ МОНІТОРИНГУ ПОДІЙ В СИСТЕМІ УПРАВЛІННЯ РОЗПОДІЛЕНИМИ БАЗАМИ ДАНИХ НА ОСНОВІ СЕНСОРІВ ПОДІЙ

DOI 10.36994 / 2788-5518-2023-01-05-17

Корнага Я.І., д.т.н., проф., Національний технічний університет України «Київський політехнічний університет імені Ігоря Сікорського», Київ, Україна, slovyan_k@ukr.net

Барабаш А.О. Національний технічний університет України «Київський політехнічний університет імені Ігоря Сікорського», Київ, Україна. andrew.barbsh@gmail.com

Анотація: У роботі розглянуто методи підвищення ефективності виявлення загроз в розподілених базах даних за допомогою системи моніторингу подій. Зазначено, що система працює на основі моделі моніторингу подій різномірних розподілених баз даних. Дана модель передбачає три етапи обробки подій, які базуються на відповідних методах. Виокремлено, що механізми, які покладені в основу функціонування згаданих методів, повинні бути приведені до єдиного формату даних для усунення можливої появи некоректної роботи в майбутніх обчисленнях. Метод обробки подій на сервері моніторингу дозволяє обробляти матриці подій та передавати на засоби контролю функцію на основі якої приймаються відповідні рішення щодо підвищення надійності.

В статті розроблено модифікований метод аналізу та моніторингу подій в нереляційних розподілених базах даних. Запропоновано параметри моніторингу подій для операцій пошуку в розподілених базах даних. Для підтвердження теоретичних результатів проведено математичне моделювання та експериментальне дослідження параметрів моніторингу та швидкості обробки даних на сервері моніторингу подій. Досліджено кількості виявлених інцидентів та кількості невірно ідентифікованих подій, що в свою чергу дозволяє забезпечувати захист інформації від несанкціонованих дій чи змін. Також, слід підкреслити, що запропонована в роботі модель описує детальний моніторинг запитів, транзакцій та збережених процедур в базі даних з інформацією щодо можливих інцидентів в режимі реального часу та спонукає до попереджень вторгнень. Крім того, особливості запропонованої моделі дозволяють досліджувати та відстежувати появу нових вразливих сегментів бази даних та оперативно реагувати на них, усуваючи загрози каналам доступу. В роботі проведено порівняння швидкості обробки даних на основі файлів та таблиць баз даних, що дозволяє в повній мірі виконати експерименти та підтвердити аналітичні дані експериментальними.

Ключові слова: аналіз подій, бази даних, моніторинг подій, розподілені бази даних, швидкість обробки.

MATHEMATICAL MODEL OF EVENTS MONITORING IN THE MANAGEMENT SYSTEM OF DISTRIBUTED DATABASES BASED ON EVENT SENSORS

Yaroslav Kornaga, Dr.habil., Prof. National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», Kyiv, Ukraine. slovyan_k@ukr.net

Andrii Barabash. National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», Kyiv, Ukraine., andrew.barbsh@gmail.com.

Abstract: Methods of increasing the effectiveness of threat detection in distributed databases using the event monitoring system are considered in the work. It is noted that the system works on the basis of the event monitoring model of heterogeneous distributed databases. This model involves three stages of event processing, which are based on appropriate methods. It is highlighted that the mechanisms underlying the functioning of the mentioned methods should be reduced to a single data format to eliminate the possible appearance of incorrect work in future calculations. The event processing method on the monitoring server allows for processing event matrices and transferring the function to control tools, on the basis of which appropriate decisions are made to improve reliability.

The article develops a modified method of analyzing and monitoring events in non-relational distributed databases. Event monitoring options for lookup operations in distributed databases are offered. To confirm the theoretical results, mathematical modeling and experimental research on monitoring parameters and data processing speed on the event monitoring server were carried out. The number of detected incidents and the number of incorrectly identified events were studied, which in turn allowed for the protection of information from unauthorized actions or changes. Also, it should be emphasized that the model proposed in the work describes

detailed monitoring of requests, transactions, and stored procedures in a database with information about possible incidents in real time and prompts intrusion warnings. In addition, the features of the proposed model make it possible to investigate and monitor the emergence of new vulnerable segments of the database and promptly respond to them, eliminating threats to access channels. The paper compares the speed of data processing based on files and database tables, which allows for fully performing experiments and confirming analytical data experimentally.

Keywords: event analysis, databases, event monitoring, distributed databases, processing speed.

Вступ

В сучасних різномірних розподілених системах управління базами даних (СУБД) та існуючих пошукових системах використовуються різноманітні способи, методи та засоби моніторингу подій, які дозволяють здійснювати керування базами даних у відповідності із вимогами що висуваються до систем. Механізми, що дозволяють одночасно налаштовувати параметри функціонування бази даних та управляти нею потребують приведення до єдиного формату обробки даних. При цьому, методи обробки подій в СУБД, що використовуються для операцій пошуку, не в повній мірі висвітлюють вплив даних подій на конкретні умови функціонування бази даних [1,2,3].

Виконання досліджень

З метою збільшення швидкості операцій в базах даних необхідно створити новий метод моніторингу подій, який дозволить удосконалити та уніфікувати обробку подій в різномірних розподілених баз даних та за допомоги додаткового сервера пришвидшить їх обробку, а також, надасть можливість нейтралізувати можливі загрози шляхом автоматичного втручання системи без обслуговуючого персоналу або за допомоги системного адміністратора.

Класична модель моніторингу подій базується на основі алгоритму найбільших статистичних аномалій (АНСА). Основу даного алгоритму складає формування вектору Бернуллі, який в подальшому обробляється сервером моніторингу подій. Таким чином, формується вектор $\bar{X} = \{x_1, x_2, \dots, x_n\}$, який порівнюється з пороговим вектором, після чого пересилається на сервер обробки моніторингу подій. В свою чергу, сервер приймає рішення щодо тієї чи іншої події. Складові вектору Бернуллі включають в себе тільки параметри подій. Проте, вектор не містить параметри впливу події на той чи інший елемент системи[3]. Тому потрібно створити систему моніторингу подій, яка б могла обробляти не тільки події, а й їх розподіл за відповідними властивостями бази даних та СУБД.

Запропонована математична модель моніторингу подій розподілених баз даних заснована на використанні сенсорів подій та включає в себе такі кроки:

I. Збір подій моніторингу.

1. Для кожної бази даних з розподіленою системою управління будуємо сенсори подій на основі тригерів.

2. Для кожної множини подій $\bar{E}_i = \{e_1, e_2, \dots, e_n\}$, що вимірюється сенсорами подій, будується порогова матриця за можливими наслідками кожної події із множини $\bar{C}_i = \{c_1, c_2, \dots, c_m\}$.

3. Сенсори подій проводять збір інформації для матриці подій $S_i = \bar{E}_i \times C_i$ з простору S і також пересилають їх до серверу моніторингу подій.

II. Метод обробки подій моніторингу на сервері моніторингу подій.

1. Після передачі матриці подій на сервер моніторингу подій проводяться операції аналізу відмінності між отриманою матрицею S_i та заданною пороговою матрицею S_0 , тобто $B_i = S_i - S_0$.

2. Проводиться формування експертних оцінок на основі формування вектору $\bar{W}_i = \{w_1, w_2, \dots, w_m\}$.

3. Після отримання вектору експертних оцінок $\bar{W}_i$, матриця B_i множиться на вектор експертних оцінок $\bar{W}_i$: $\bar{Z}_i = B_i \times \bar{W}_i$.

4. Отримані результати оцінюються за допомогою функції R , яка визначає від'ємні значення у векторах, та формує вектор відповіді $\bar{F}_i = \{f_1, f_2, \dots, f_n\}$ в якому записується результат обробки події.

$$\bar{F} = R(\bar{Z}_i) = R(B_i \times \bar{W}_i) = R((S_i - S_0) \times \bar{W}_i). \quad (1)$$

III. Обробка відповіді серверу моніторингу подій.

1. Вектор відповіді $\bar{F}_l = \{f_1, f_2, \dots, f_m\}$ передається в модуль контролю подій.
 2. В модулі контролю подій у випадку наявності подій, що можуть спричинити негативний вплив на бази даних розподіленої системи, генерується рішення про застосування певних заходів щодо нейтралізації негативних наслідків.

3. У випадку виникнення нештатних ситуацій, модуль контролю подій пересилає сигнал на блок формування попередження для адміністратора серверу моніторингу подій.

Архітектура засобів моніторингу подій розподілених баз даних на основі сенсорів подій.

Отримана система моніторингу подій, в автоматичному режимі, ідентифікує всі події, які відбуваються у розподілених базах даних. Під час роботи сенсор подій приєднується до тимчасово-виділеної пам'яті та циклічно проводить опитування з моніторингу подій на основі вибірки даних з пам'яті із деякою частотою[4]. Під час кожного циклу сенсор робить аналіз активних, у цей момент процесів, під час кожної сесії в базі даних та визначає події на основі правил, які встановлені адміністратором. Надалі інформація про події надсилається на сервер моніторингу подій де відбувається їх подальший аналіз та генеруються відповідні повідомлення.

На рис. 1 показано схему, що була отримана на основі розробленої моделі.

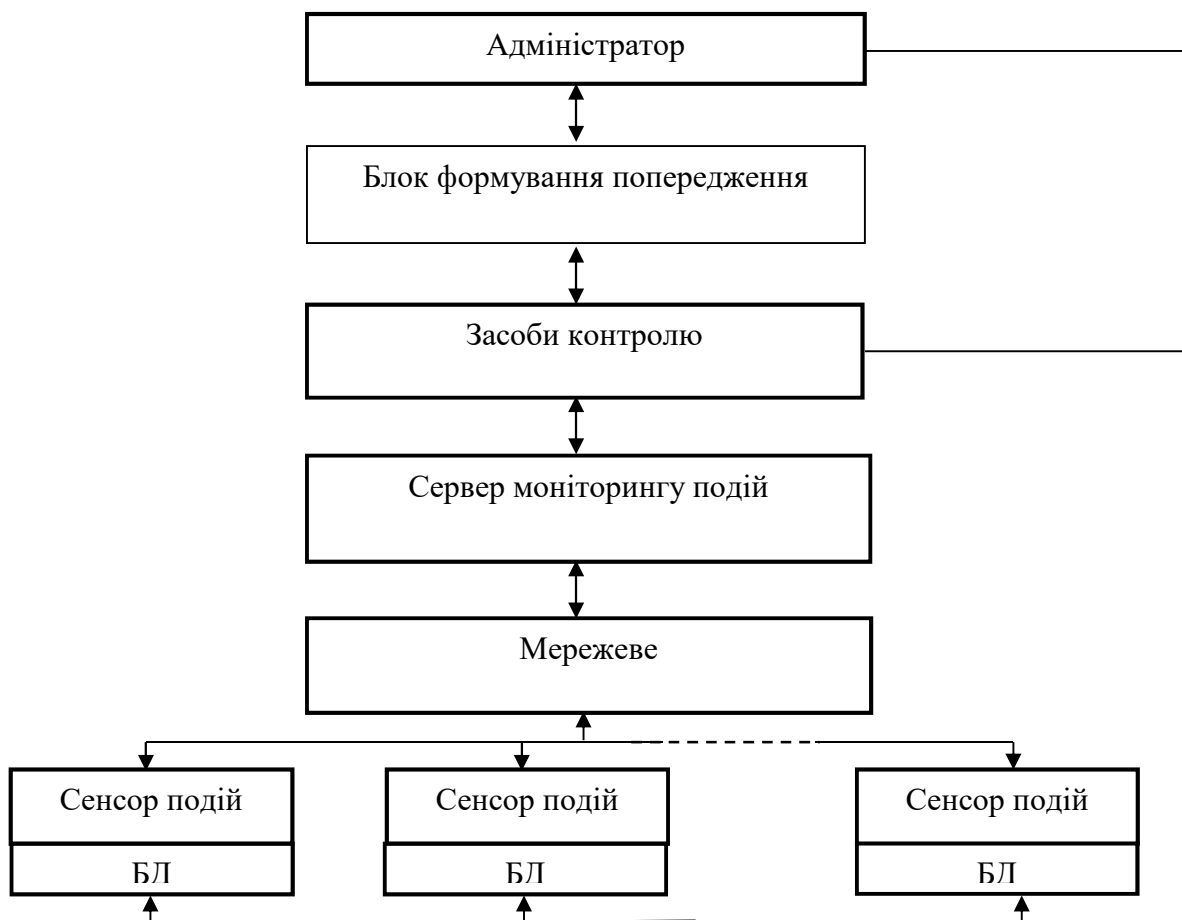


Рис.1. Архітектура моделі моніторингу подій на основі сенсорів подій

Система моніторингу подій у свою чергу також має можливість примусово припиняти сесії роботи у випадку певних несанкціонованих подій, що можуть бути вчинені, з боку користувачів. З іншого боку, для роботи сенсора подій необхідні невеликі обчислювальні ресурси бази даних та його робота практично не впливає на основні операції щодо обробки запитів. Попереджуючу функція сенсора можна реалізувати із використанням DDL-тригерів, що вибірково призупиняють виконання команд DDL та DCL на невеликі проміжки часу (до декількох мс), а це дозволяє сенсору подій вчасно реагувати щодо небезпечних дій [5].

На рис. 2 представлена детальна архітектура сенсора і серверу подій.

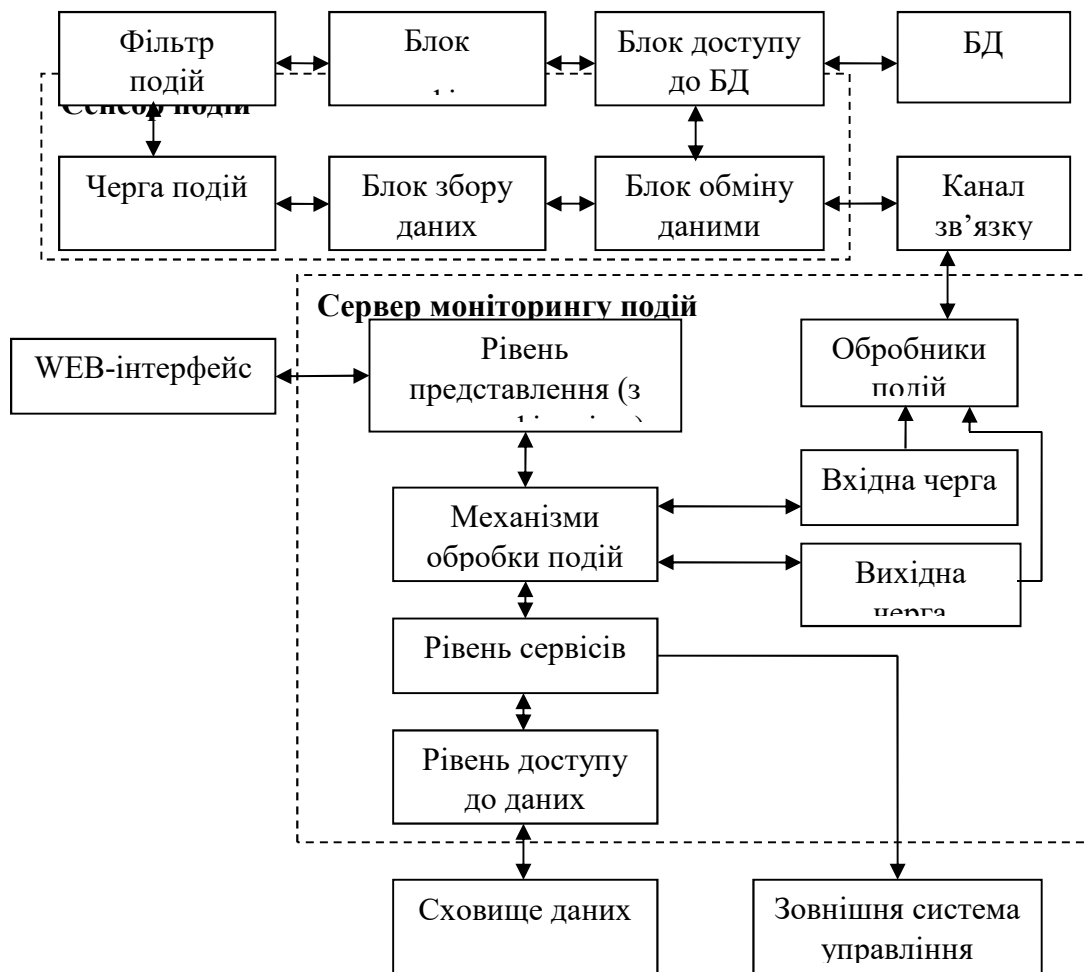


Рис.2.Зображення архітектури сенсора подій і сервера моніторингу подій

Визначений сервер моніторингу подій у представленій системі моніторингу подій здатний керувати декількома сенсорами подій із різних розподілених баз даних. Крім того, сервер підтримує опцію масштабування кількості сенсорів подій. Він достатньо легко інтегрується у структуру системи моніторингу СУБД. В свою чергу це дозволяє підвищити загальну ефективність управління надійністю обробки даних у розподілених базах даних.

Архітектура запропонованої системи забезпечує розподіл прав та повноважень між суб'єктами, що є однією із основних умов, які висуваються до систем найденної обробки різного типу даних[6]. Адміністратор даної системи є персоною, яка визначає основні правила політики надійності і безпеки та отримує повідомлення щодо інцидентів використовуючи блок формування попередження. Розглянута система дозволяє по-перше забезпечити потрібний рівень надійності щодо обробки даних у розподілених базах даних, а по-друге – неперервне виконання обробки запитів.

Методика оцінки ефективності системи моніторингу подій розподілених баз даних.

Щоб провести дослідження параметрів системи моніторингу подій у розподілених базах даних була розроблена методика, що дозволяє оцінювати ефективність роботи сервера моніторингу подій на основі використання сенсорів подій [7].

Операції у базах даних класифікуються за двома ознаками:

- Дозволені.
- Не дозволені.

Нехай G – кількість можливих подій у базі даних, K – множина правильно відпрацьованих подій в базі даних. Позначимо: $i \in G$ – множина подій, що не були знайдені системами

моніторингу подій, $o \in K$ – множина правильно відпрацьованих подій, що були неправильно ідентифіковані системою моніторингу подій як елементи, які порушують надійність.

Ймовірність коректного виявлення подій можна розраховувати наступним чином:

$$P_i = 1 - \frac{i}{G}. \quad (2)$$

Ймовірність неправильних ідентифікацій правильно відпрацьованих подій розраховується у вигляді:

$$P_k = 1 - \frac{i}{G-i+k}, \quad (3)$$

де $G - i + k$ – це кількість несанкціонованих подій, що були виявлені за допомогою системи моніторингу подій.

На основі цього ймовірність правильного виявлення несанкціонованих подій системою моніторингу подій можна розрахувати на основі формули:

$$P = 1 - \max\{P_i, P_k\} = 1 - \max\left\{\frac{i}{G}, \frac{k}{G-i+k}\right\}. \quad (4)$$

Визначений показник можна використовувати для комплексної оцінки ефективності щодо функціонування системи моніторингу подій у розподілених базах даних [8].

Крім того, ймовірність коректної роботи серверу моніторингу подій можна визначити наступним чином:

$$P = 1 - \max\{P_I, P_{II}\}, \quad (5)$$

де P_I – ймовірність помилки першого роду («пропущена подія» – відсутність реакції на подію), P_{II} – ймовірність помилки другого роду («хибна тривога» – помилкове спрацювання).

Експериментальне підтвердження, як це показано на рисунку 3, відбувалися у двох напрямках, а саме:

- Знаходження кількості спроб інцидентів.
- Загальна кількість помилок першого та другого роду, які знайдені в системі моніторингу подій.

Аналіз експериментальних досліджень показує, що сумарна ймовірність помилок першого та другого роду складає 6 – 12 %. Це в свою чергу дозволяє стверджувати, що ймовірність коректного функціонування серверу моніторингу складає 88 – 94 %.

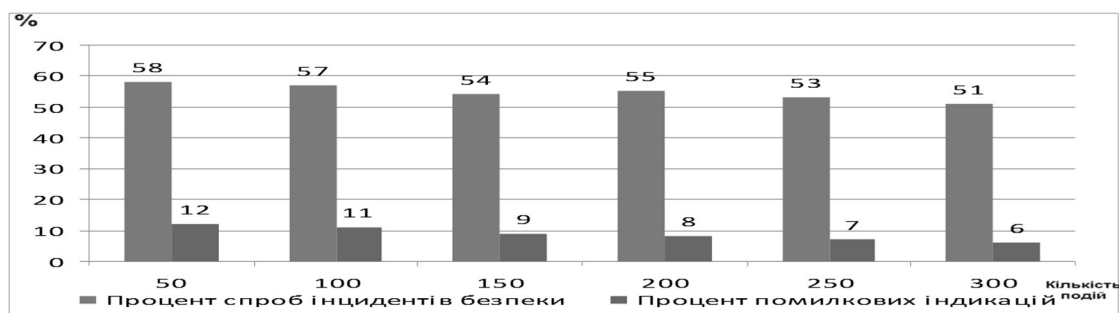


Рис. 3. Результати експериментальних досліджень серверу моніторингу подій

Висновки

Запропоновано математичну модель реалізації моніторингу подій баз даних, на основі якої виконується всебічний та комплексний моніторинг усіх дій з базою даних та забезпечується захист інформації від несанкціонованих дій легальних користувачів. Разом із цим, запропонована модель дозволяє описати детальний моніторинг запитів, транзакцій та збережених процедур в базі даних з повідомленнями щодо інцидентів в реальному режимі часу та попередженням вторгнень. Крім того, запропонована модель дозволяє дослідити та відстежити нові виявлені

вразливі місця бази даних й оперативно усунути ці вразливі канали доступу. Це є досить важливим в практиці застосування розподілених баз даних.

Література

1. Vadym Mukhin, Valerii Zavgorodnii, Yaroslav Kornaga, Ivan Krysak, Maxim Bazaliy, Oleg Mukhin. Program Code Protecting Mechanism Based on Obfuscation Tools. *IEEE International Conference on System Analysis&Intelligent Computing*. 2022. Pp 407-419.
2. Marcus Hirt, Marcus Lagergren. Oracle jrookit: The Definitive Guide. Packt Publishing, 2010. P.588.
3. Vadym Mukhin, Yaroslav Kornaga, Yurii Bazaka, Ievgen Krylov, Andrii Barabash, Alla Yakovleva, Oleg Mukhin. The Testing Mechanism for Software and Services Based on Mike Cohn's Testing Pyramid Modification. *11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*. 2021. Pp.589-595.
4. Teo Lachev. Applied Microsoft SQL Server 2008. Reporting Services. Prologika Press, 2008. P.768.
5. Коннор Мак Дональд. Oracle. Практичні рішення. 2006. С.560.
6. Babich Oksana, Vyshnyvskiy Viktor, Mukhin Vadym, Zamaruyeva Irina, Sheleg Michail, Kornaga Yaroslav. The Technique of Key Text Characteristics Analysis for Mass Media Text Nature Assessment. *International Journal of Modern Education&Computer Science*. 2022. Vol. 14, Issue 1, pp.1-16.
7. Мухін В.Є., Корнага Я.І. Механізми підвищення ефективності процедури моніторингу безпеки в розподілених базах даних. Вісник Національного технічного університету «Харківський політехнічний інститут». Збірник наукових праць. Серія: Інформатика та моделювання. Харків: НТУ «ХПІ». 2012. № 38., С.128-135.
8. Vadym Mukhin, Valerii Zavgorodnii, Oleg Barabash, Anna Zavgorodnya, Vitalii Statkevych. Method of Restoring Parameters of Information Objects in a Unified Information Space Based on Computer Networks. *International Journal of Computer Network & Information Security*. 2020.12 (2).

References

1. Vadym Mukhin, Valerii Zavgorodnii, Yaroslav Kornaga, Ivan Krysak, Maxim Bazaliy, Oleg Mukhin. Program Code Protecting Mechanism Based on Obfuscation Tools. *IEEE International Conference on System Analysis&Intelligent Computing*. 2022. Pp. 407-419.
2. Marcus Hirt, Marcus Lagergren. Oracle jrookit: The Definitive Guide. Packt Publishing, 2010. P.588.
3. Vadym Mukhin, Yaroslav Kornaga, Yurii Bazaka, Ievgen Krylov, Andrii Barabash, Alla Yakovleva, Oleg Mukhin. The Testing Mechanism for Software and Services Based on Mike Cohn's Testing Pyramid Modification. *11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*. 2021. Pp.589-595.
4. Teo Lachev. Applied Microsoft SQL Server 2008. Reporting Services. Prologika Press, 2008. P.768.
5. Konnor Mak Donald. Oracle. Praktychni rishennia [Oracle. Practical solutions]. 2006. C.560.
6. Babich Oksana, Vyshnyvskiy Viktor, Mukhin Vadym, Zamaruyeva Irina, Sheleg Michail, Kornaga Yaroslav. The Technique of Key Text Characteristics Analysis for Mass Media Text Nature Assessment. *International Journal of Modern Education&Computer Science*. 2022. Vol. 14, Issue 1, pp.1-16.
7. Mukhin V.I., Kornaga Y.I. Mekhanizmy pidvyshchennia efektyvnosti protsedury monitorynhu bezpeky v rozpodilenykh bazakh danykh [Mechanisms for increasing the efficiency of the security monitoring procedure in distributed databases] Visnyk Natsionalnoho tekhnichnoho universytetu «Kharkivskiy politekhnichnyi instytut». Zbirnyk naukovykh prats. Serii: Informatyka ta modeliuvannia. Kharkiv: NTU «KhPI». 2012. № 38., S.128-135.
8. Vadym Mukhin, Valerii Zavgorodnii, Oleg Barabash, Anna Zavgorodnya, Vitalii Statkevych. Method of Restoring Parameters of Information Objects in a Unified Information Space Based on Computer Networks. *International Journal of Computer Network & Information Security*. 2020. 12 (2).

УДК 004.021, 004.588, 004.78

РОЗВИТОК МОБІЛЬНОЇ ОСВІТИ ДЛЯ ПОКРАЩЕННЯ ЯКОСТІ НАВЧАННЯ

DOI 10.36994 / 2788-5518-2023-01-05-18

Середа А.В., Відкритий міжнародний університет розвитку людини «Україна», Інститут комп'ютерних технологій, Київ, Україна. <https://orcid.org/0000-0002-4779-5475>, andrew.sereda@gmail.com

Самарай В. П., к.т.н., проф., Відкритий міжнародний університет розвитку людини «Україна», Інститут комп'ютерних технологій, Київ, Україна. <https://orcid.org/0000-0003-4419-1366>, samaraj@ukr.net

Даценко І. П., к.т.н., доц., Відкритий міжнародний університет розвитку людини «Україна», Інститут комп'ютерних технологій, Київ, Україна. <https://orcid.org/0000-0002-0047-413X>, docik_ivan@ukr.net

Павленко В. І., к.ф.-м.н., доц., Відкритий міжнародний університет розвитку людини «Україна», Інститут комп'ютерних технологій, Київ, Україна. <https://orcid.org/0000-0002-3958-0415>, pavlenko.v@i.ua

Бровченко Є. М., Відкритий міжнародний університет розвитку людини «Україна», Інститут комп'ютерних технологій, Київ, Україна. <https://orcid.org/0000-0002-1416-0385>, yebrovchenko@hotmail.com

Таланчук П. М., д.т.н., проф., Відкритий міжнародний університет розвитку людини «Україна», Київ, Україна. <https://orcid.org/0000-0001-8748-6127>, talanshuk_pm@ukr.net

Анотація: у статті розглядається розвиток мобільної освіти як інноваційного інструменту для поліпшення якості навчання. Автори аналізують вплив мобільних технологій на освіту та пропонують шляхи використання мобільних пристроїв для досягнення цілей освітньої системи. Розкриваються переваги використання мобільних пристроїв у навчальному процесі, зокрема доступність навчання будь де та в будь-який час, гнучкість навчання, розширені можливості використання різноманітних форматів навчання (відеоуроки, інтерактивні ігри, тести тощо), підтримка комунікації між вчителем та учнем. Взаємодія користувача у освітньому процесі для поліпшення якості, розвиток та рекомендації на основі методу аналізу ієрархій(MAI). Проведене анкетування експертів, визначена основна мета – поліпшення якості навчання, визначені критерії та альтернативи. Математично доведено за допомогою програми Ексель, що мобільні технології покращують якість навчання. Підкреслюється необхідність підготовки вчителів та учнів до ефективного використання мобільної освіти щодо викликів потенційних негативних наслідків, які можуть виникнути при неконтрольованому використанні мобільних пристроїв у навчальному процесі. Розвиток мобільної освіти може суттєво покращити якість навчання, забезпечити більш ефективно та захоплююче навчання, а також знизити витрати на навчальні матеріали та обладнання.

Ключові слова: мобільні пристрої, освіта, захист, метод аналізу ієрархій, дистанційне навчання, якість навчання, покращення.

EVOLUTION OF MOBILE EDUCATION TO IMPROVE THE QUALITY OF LEARNING

Andrii Sereda, Open University of Human Development "Ukraine", Kyiv, Ukraine. <https://orcid.org/0000-0002-4779-5475>, andrew.sereda@gmail.com

Valery Samarai, Ph.D., Ass. Prof., Open International University of Human Development "Ukraine", Kyiv, Ukraine. <https://orcid.org/0000-0003-4419-1366>, samaraj@ukr.net

Ivan Datsenko, Ph.D., Open University of Human Development "Ukraine", Kyiv, Ukraine. <https://orcid.org/0000-0002-0047-413X>, docik_ivan@ukr.net

Volodymyr Pavlenko, Ph.D., Ass. Prof., Open University of Human Development "Ukraine", Kyiv, Ukraine. <https://orcid.org/0000-0002-3958-0415>, pavlenko.v@i.ua

Evgen Brovchenko, Open International University of Human Development "Ukraine", Kyiv, Ukraine. <https://orcid.org/0000-0002-1416-0385>, yebrovchenko@hotmail.com

Peter Talanchuk, Dr. habil., Prof., Open International University of Human Development "Ukraine", Kyiv, Ukraine. <https://orcid.org/0000-0001-8748-6127>, talanshuk_pm@ukr.net

Annotation: article considers the evolution of mobile education as an innovative tool for improving the quality of education. Authors analyze the impact of mobile technologies on education and suggest ways to use mobile

devices to achieve the goals of the educational system. Mobile devices used in the educational process is the advantage that disclose the availability of learning at any time and in any place, flexibility of learning, expanded opportunities for using various learning formats (video lessons, interactive games, tests, etc.), support for communication between teacher and student. User interaction in the educational process for quality improvement, development and recommendations based on the method of analytic hierarchy process (AHP). A survey of experts was conducted, to improve the quality of education was determined as the main goal, criteria and alternatives were determined. It has been mathematically proven using the Excel program that mobile technologies improve the quality of education. The need to prepare teachers and students for the effective use of mobile education regarding the challenges of potential negative consequences that may arise from the uncontrolled use of mobile devices in the educational process is emphasized. The development of mobile education can significantly improve the quality of education, provide more effective and exciting learning, and reduce the costs of educational materials and equipment.

Keywords: *mobile devices, education, protection, method of analysis of hierarchies, distance learning, quality of education, improvement.*

Вступ

У сучасному світі мобільні пристрої стали невід'ємною частиною нашого повсякденного життя, їх використання в освіті є актуальною та перспективною темою досліджень. Розвиток технологій та збільшення кількості мобільних пристроїв призвели до зростання популярності мобільної освіти, що надає учням та викладачам можливість доступу до навчального матеріалу та дистанційної комунікації в зручний час.

Мобільна освіта - набір технічних можливостей щодо поліпшення якості навчання і підвищення мотивації учнів[1]. У цій статті ми розглянемо принципи та підходи до розробки мобільної освіти, а також роль мобільних пристроїв у покращенні якості навчання та розвитку компетентності учнів.

Метою даної статті є сприяння розвитку мобільної освіти, розкриття нових можливостей використання мобільних технологій у навчальному процесі з метою підвищення його якості. Сподіваємось, що досвід та знання, набуті в ході дослідження розвитку мобільної освіти, будуть корисними для всіх учасників освітнього процесу.

В сучасному світі спостерігається стрімкий розвиток технологій та зростання популярності мобільних пристроїв. За даними статистики, кількість людей, які ними користуються, щорічно збільшується. Це розширює можливості використання мобільних технологій у навчальному процесі. Отже, актуальність дослідження полягає у необхідності розробки та впровадження нових технологій та ресурсів для мобільної освіти, що допоможе удосконалити процес навчання і якість освіти. Крім того, мобільна освіта може стати ефективним інструментом для розвитку економіки, зокрема, її інформаційного сектору. Розробка мобільних додатків та ресурсів для навчання може стати стимулом для розвитку інноваційних технологій та підвищення конкурентоспроможності країни. Отже, дане дослідження є актуальним для розвитку мобільної освіти як потужного фактору розвитку інноваційного сектору.

Виконання досліджень

Наукові дослідження в галузі мобільної освіти та її впливу на якість навчання проводилися протягом декількох останніх десятиліть. Аналіз наукових публікацій з цієї тематики дозволяє визначити основні напрямки досліджень, підходи до вивчення впливу мобільної освіти на навчання та результати проведених досліджень.

Ключовим напрямком досліджень є використання мобільних пристроїв у навчальному процесі. У статті "Mobile learning: Transforming education, engaging students, and improving outcomes" автори дослідження дійшли висновку, що використання мобільної технології у навчальному процесі може покращити якість навчання, підвищити мотивацію і активність студентів.

Іншим напрямком досліджень є вивчення впливу мобільної освіти на показники успішності студентів. У статті "Impact of mobile learning on students' learning behaviours and academic performance: A review of the literature" автори зазначили, що використання мобільних пристроїв в навчальному процесі позитивно впливає на академічну успішність студентів, забезпечуючи їм більш глибоке розуміння навчального матеріалу.

Проведені дослідження показали, що використання мобільних пристроїв стимулює розвиток самостійного навчання через вдосконалення навичок пошуку та аналізу інформації. У статті

"Mobile learning: A review of current research" автори дослідження підкреслюють, що мобільна освіта сприяє розвитку критичного мислення та підвищує мотивацію студентів до навчання. Дослідження показали, що використання мобільних пристроїв може стимулює інтерес студентів до матеріалу, забезпечуючи більш швидкий та зручний доступ до інформації[2].

Отже, необхідно забезпечити якісну підготовку викладачів та студентів до використання таких технологій, а також розробити оптимальні методики використання мобільних пристроїв у навчальному процесі.

Одним з можливих експертних математичних методів, який можна використати для дослідження ефективності застосування мобільних технологій в освіті є метод аналізу ієрархій (MAI). Використовується для прийняття рішень, коли є кілька критеріїв вибору і кожен критерій має декілька варіантів. Метод базується на ієрархічній структурі, в якій на вершині знаходиться головний критерій, а нижче розташовані критерії другого рівня, що впливають на головний критерій, і так далі до визначення конкретних альтернатив. Головне використання методу полягає у підтримці процесу прийняття рішень через ієрархічну композицію задач та рейтинг альтернативних варіантів. Цей метод дозволяє провести аналіз проблеми, де проблема ухвалення рішення виражена у вигляді ієрархічно впорядкованих елементів: альтернатив, системи зв'язків, критеріїв та мети рейтингу

Метод аналізу ієрархій (MAI) базується на розбитті проблеми на прості компоненти та послідовному обробленні суджень, які здійснює особа, що приймає рішення, шляхом їх парного порівняння. Результатом аналізу є відносний ступінь взаємодії між окремими елементами у побудованій ієрархії, який може бути виражений кількісно. На Рис. 1 наведено приклад розбиття проблеми поліпшення якості навчання.

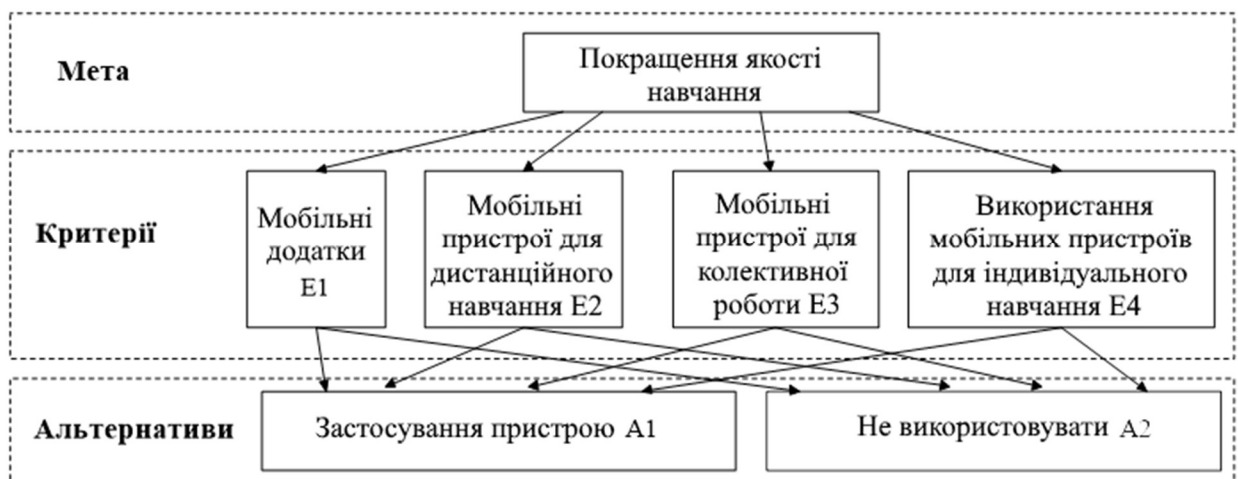


Рис. 1. Розбиття проблеми поліпшення якості навчання

Згідно з наданою ієрархією, потрібно визначити найбільш ймовірний сценарій поліпшення якості навчання (альтернативи A1, A2) на основі критеріїв: мобільні додатки (E1), мобільні пристрої для колективної роботи (E2), мобільні пристрої для дистанційного навчання (E3) та використання мобільних пристроїв для індивідуального навчання (E4). Після розбиття задачі на ієрархію визначається пріоритетність критеріїв шляхом порівняння їх попарно. Система порівняння суджень за методом аналізу ієрархій Т. Сааті [4] призводить до результату, який може бути представлений у вигляді зворотно-симетричної матриці, яка показана в табл. 1.

Таблиця 1. Зворотно-симетрична матриця.

Фактор	1 (j-й фактор)	...	N
1 (i-й фактор)	1		a (i,j)
...		1	
n	a (i,j)		1

Кожен елемент матриці a (i, j) визначається шляхом порівняння важливості фактору і відносно фактору j за шкалою інтенсивності від 1 до 9. Оцінки мають таке значення:

- 1 - обидва фактори мають рівну важливість;
- 3 - фактор i помірно домінує над фактором j;

- 5 - фактор і істотно переважає фактор j;
 7 - фактор і значно переважає фактор j;
 9 - фактор і має дуже сильну перевагу над фактором j;
 2, 4, 6, 8 - відповідні проміжні значення між цими категоріями.

Експертами можуть бути люди з великим досвідом використання мобільних технологій в освіті, або ж ті, хто має професійний досвід в галузі освіти та має компетентність для оцінки ефективності використання мобільних технологій у навчальному процесі. У залежності від мети дослідження та контексту, можуть використовуватися різні критерії для відбору експертів. Важливо, щоб група експертів була представлена різними категоріями учасників освітнього процесу та мала достатній рівень компетентності для здійснення оцінки ефективності застосування мобільних технологій в освіті.

У ролі експертів залучаються викладачі вищих навчальних закладів, які пройшли опитування на важливість критеріїв та відповідність альтернативам. Також можуть бути залучені студенти, які використовують мобільні технології в освітньому процесі.

Приклад анкети-опитування про ефективність використання мобільних технологій в освіті:

Як ви оцінюєте важливість використання мобільних технологій в освіті?

(Відповіді по шкалі від 1 до 5, де 1 - не важливо, 5 - дуже важливо)

Які з наступних критеріїв є найважливішими для вас при використанні мобільних технологій в освіті?

- Поліпшення якості навчання
- Збільшення мотивації студентів до навчання
- Поліпшення комунікації між викладачем та студентами
- Підвищення ефективності організації навчального процесу

Які типи мобільних пристроїв ви вважаєте найбільш ефективними для використання в освіті?

- Смартфон
- Планшет
- Ноутбук
- Система віртуальної реальності
- Інший (вказати)

Як ви оцінюєте відповідність наступних альтернатив за використанням мобільних технологій в освіті?

(Відповіді на шкалі від 1 до 5, де 1 - не відповідає, 5 - повністю відповідає)

- Використання мобільних додатків для навчання
- Використання мобільних додатків для комунікації з викладачем та студентами
- Використання мобільних пристроїв для доступу до електронних ресурсів та навчальних матеріалів

- Використання мобільних пристроїв для відеоконференцій з викладачем та студентами

Як ви оцінюєте доступність мобільних технологій для студентів?

(Відповіді на шкалі від 1 до 5, де 1 - не доступні, 5 – доступні)

Матриця попарних порівнянь суджень визначення пріоритетності критеріїв при виборі поліпшення якості освіти представлена на рис. 2

	Мета			
	E1	E2	E3	E4
E1	1	1/5	1/3	1/7
E2	5	1	3	1/3
E3	3	1/3	1	1/5
E4	7	3	5	1

Рис. 2. Матриця попарних порівнянь поліпшення якості освіти відповідно критеріям.

За даними матриці, при виборі поліпшення якості освіти, критерій E2 має істотну перевагу перед критерієм E1 та помірне домінування над критерієм E3. Критерій E3 демонструє помірне домінування над критерієм E1. Критерій E4, з свого боку, виявляє значну перевагу над критерієм E1, помірне домінування над критерієм E2 та істотну перевагу над критерієм E3.

За формулами (1-2) можна визначити пріоритет критеріїв для вибору альтернативи:

$$w_i = \sqrt[n]{\prod_{j=1}^n a_{ij}}, \quad (1)$$

$$w_{\text{норм}} = w_i / \sum_{i=1}^n w_i \quad (2)$$

де W позначає компоненту власного вектора матриці (середнє значення пріоритетності оцінок).

$W_{\text{норм}}$ вказує на нормований вектор матриці попарних порівнянь (вектор пріоритетів).

a_{ij} позначає значення окремих елементів матриці.

Оцінка узгодженості думок експертів здійснюється за формулами (3-5):

$$\lambda = \sum_{i=1}^n a_{ij} \cdot w_{\text{норм } i} \quad (3)$$

$$I_c = (\lambda - n) / (n - 1) \leq 0,2 \quad (4)$$

$$OC = I_c / I_{cc} \quad (5)$$

де λ – власне значення матриці;

I_c – індекс узгодженості;

I_{cc} – індексу узгодженості, середній (табл. 2);

OC - відношення узгодженості;

n – кількісна характеристика критеріїв.

Таблиця 2 Індекс узгодженості, середній.

Р										
озмір матриці										0
С										
ереднє значення індексу узгодженості (I_{cc})			,58	,90	,12	,24	,32	,41	,45	,49

На рис. 3 продемонстрований розрахунок пріоритетності критеріїв поліпшення якості освіти.

Мета									
	E1	E2	E3	E4	Wi	Wнорм		ICC=	0,9
E1	1,00	0,20	0,33	0,14	0,31	0,06		$\lambda=$	4,12
E2	5,00	1,00	3,00	0,33	1,50	0,26		$IC=(1-n)/(n-1)$	0,04
E3	3,00	0,33	1,00	0,20	0,67	0,12		$OC=IC/ICC$	0,04
E4	7,00	3,00	5,00	1,00	3,20	0,56			
Σ	16,00	4,53	9,33	1,68	5,68	1,00			

Рис. 3. Пріоритети критеріїв розрахунок поліпшення якості освіти

Згідно з обчисленнями, найважливішим критерієм, що визначає сценарій розвитку, є використання мобільних пристроїв для індивідуального навчання (E4). Значення індексу узгодженості менше 0,2 та відносної узгодженості менше 0,1 свідчать про згоду експертів.

Подальшим кроком є визначення пріоритетів альтернатив кожного з критеріїв (див. Рис. 4).

E1					E2				
	A1	A2	Wi	Wнорм		A1	A2	Wi	Wнорм
A1	1,00	5,00	2,24	0,83	A1	1,00	0,14	0,38	0,13
A2	0,20	1,00	0,45	0,17	A2	7,00	1,00	2,65	0,88
Σ	1,20	6,00	2,68	1,00	Σ	8,00	1,14	3,02	1,00

E3					E4				
	A1	A2	Wi	Wнорм		A1	A2	Wi	Wнорм
A1	1,00	3,00	1,73	0,75	A1	1,00	9,00	3,00	0,90
A2	0,33	1,00	0,58	0,25	A2	0,11	1,00	0,33	0,10
Σ	1,33	4,00	2,31	1,00	Σ	1,11	10,00	3,33	1,00

Рис. 4. Пріоритет альтернатив кожного з критеріїв

Вищий пріоритет за критерієм E1 «Мобільні додатки» має застосування пристрою (0,83), за критерієм E2 «Мобільні пристрої для колективної роботи» – не використовувати (0,88).

Відповідно до критерію E3 «Мобільні пристрої для дистанційного навчання» вищий пріоритет має застосування пристрою (0,75). Абсолютний пріоритет застосування пристрою має за критерієм E4 – «Використання мобільних пристроїв для індивідуального навчання».

Найвищий пріоритет за більшістю із критеріїв має застосування пристрою, проте за критерієм E2 – не використовувати. Отже, щоб врахувати значимість критеріїв, потрібно обчислити загальний пріоритет альтернатив.

Найбільш пріоритетна альтернатива визначається шляхом множення матриці пріоритетів альтернатив за кожним критерієм на матрицю пріоритетів критеріїв (див. Рис. 5).

	E1	E2	E3	E4		
Застосування пристрою	0,83	0,13	0,75	0,90	Пріоритетна альтернатива:	0,67
Не використовувати	0,17	0,88	0,25	0,10		0,33

Рис. 5. Розрахунок пріоритету поліпшення якості навчання

Пріоритетність альтернатив продемонструємо графічно (рис. 6).



Рис. 6. Значення оцінки поліпшення якості навчання

У результаті застосування методу аналізу ієрархій для експертної оцінки альтернатив, можна визначити їх пріоритетність з урахуванням обраних кількісних та якісних критеріїв, а також з урахуванням "людського фактора".

Подальше застосування мобільних технологій в освіті може бути спрямоване на наступні рекомендації[6]:

1. Розробка та впровадження спеціалізованих мобільних додатків: Розробка інтерактивних мобільних додатків, спрямованих на поліпшення процесу навчання і взаємодії між викладачем та студентами. Такі додатки можуть містити навчальні матеріали, завдання, тести, можливості комунікації та співпраці, що допоможуть підвищити якість навчання.

2. Розвиток мобільних платформ для дистанційного навчання: Використання мобільних технологій для дистанційного та електронного навчання є перспективним напрямом. Розробка платформ, які дають студентам доступ до онлайн-курсів, лекцій, вебінарів та інших навчальних ресурсів через мобільні пристрої, забезпечить більш гнучкий доступ до освіти.

3. Підтримка інтерактивного навчання:

Використання мобільних технологій може сприяти розвитку інтерактивного навчання, де студенти можуть брати активну участь у процесі навчання. Наприклад, за допомогою мобільних додатків студенти можуть виконувати завдання, розв'язувати задачі, взаємодіяти з викладачами та іншими студентами онлайн.

4. Постійне оновлення інформаційних ресурсів:

Освітні установи повинні підтримувати постійне оновлення навчальних матеріалів та ресурсів, до яких студенти матимуть доступ через мобільні пристрої.

5. Проведення навчальних програм та тренінгів для викладачів:

Важливо забезпечити належну підготовку викладачів щодо використання мобільних технологій в освітньому процесі. Організація навчальних програм та тренінгів допоможе викладачам оволодіти необхідними навичками та стратегіями, які сприятимуть ефективному використанню мобільних технологій у навчанні.

6. Створення сприятливої інфраструктури:

Для успішного впровадження мобільних технологій в освіті важливо забезпечити наявність необхідної інфраструктури. Це включає надійне підключення до Інтернету, наявність відповідного програмного забезпечення та апаратного забезпечення, яке підтримує роботу мобільних пристроїв.

7. Систематичний моніторинг та оцінка:

Необхідно встановити систему моніторингу та оцінки ефективності застосування мобільних технологій в освіті. Це дозволить збирати дані про результативність, задоволення студентів та викладачів, а також виявляти можливі проблеми та розробляти стратегії для подальшого вдосконалення.

8. Співпраця з індустрією та дослідницькими центрами:

Важливо активно співпрацювати з індустрією та дослідницькими центрами для розробки та впровадження нових інноваційних мобільних технологій в освіті. Це дозволить впроваджувати передові практики та забезпечувати стаке поліпшення якості навчання.

Висновки

На основі проведеного дослідження методом аналізу ієрархій (MAI) встановлено, що використання мобільних пристроїв для індивідуального навчання має значний потенціал та може позитивно впливати на освітній процес. Застосування мобільних технологій має сприяти покращенню якості навчання, підвищенню мотивації студентів, покращенню комунікації між викладачем та студентами, тобто удосконаленню організації навчального процесу.

Вагові коефіцієнти для кожного з цих критеріїв були визначені з використанням MAI і дозволяють встановити їх ранжування. Це дозволяє зосередити увагу на найважливіших аспектах використання мобільних технологій в освіті та прийняти обґрунтовані рішення щодо їх впровадження [5].

Рекомендації, отримані на основі MAI, підтверджують важливість розробки спеціалізованих мобільних додатків, підготовки викладачів, створення сприятливої інфраструктури.

Література:

1. Ву, Дж. Х., Теннісон, Р. Д., і Ся, Т. Л. (2010). Дослідження задоволеності студентів у середовищі змішаної системи електронного навчання. *Комп'ютери та освіта*, 55 (1), 155-164.
2. Шарплз М., Арнеділло-Санчес І., Мілрад М. та Вавула Г. (2009). Мобільне навчання: маленькі пристрої, великі проблеми. У М. Шарплз (ред.), *Міжнародний довідник з мобільного навчання* (стор. 1-15). Routledge.
3. Кукульська-Хюльме А. (2012). Як робоча сила вищої освіти повинна адаптуватися до прогресу в технологіях викладання та навчання?. *Інтернет і вища освіта*, 15 (4), 247-254.
4. Саати Т.Л. (1990). «Як прийняти рішення: процес аналітичної ієрархії». *Європейський журнал операційних досліджень*, 48(1), 9-26.
5. Тракслер, Дж. (2010). Дистанційне та мобільне навчання в країнах, що розвиваються. У М. Аллі (Ред.), *Мобільне навчання: Трансформація надання освіти та навчання* (стор. 81-98). Athabasca University Press.
6. Раньєрі М., Манка С. та Фіні А. (2012). Чому (і як) вчителі залучаються до інноваційних практик з ІКТ?. *Дослідницьке дослідження в контексті італійських середніх шкіл. Технологія, педагогіка та освіта*, 21 (3), 315-326.

Referenses

1. Wu, J. H., Tennyson, R. D., & Hsia, T. L. (2010). A study of student satisfaction in a blended e-learning system environment. *Computers & Education*, 55(1), 155-164.
2. Sharples, M., Arnedillo-Sánchez, I., Milrad, M., & Vavoula, G. (2009). Mobile learning: Small devices, big issues. In M. Sharples (Ed.), *The international handbook of mobile learning* (pp. 1-15). Routledge.
3. Kukulska-Hulme, A. (2012). How should the higher education workforce adapt to advancements in technology for teaching and learning?. *Internet and Higher Education*, 15(4), 247-254.
4. Saaty, T.L. (1990). "How to make a decision: The analytic hierarchy process". *European Journal of Operational Research*, 48(1), 9-26.
5. Traxler, J. (2010). Distance and mobile learning in developing countries. In M. Ally (Ed.), *Mobile learning: Transforming the delivery of education and training* (pp. 81-98). Athabasca University Press.
6. Ranieri, M., Manca, S., & Fini, A. (2012). Why (and how) do teachers engage in innovative practices with ICT?. An exploratory study in the context of Italian secondary schools. *Technology, Pedagogy and Education*, 21(3), 315-326.

УДК 004.94

ПЕРЕДБАЧЕННЯ ВАРТОСТІ ЗЕМЛІ СІЛЬСЬКОГОСПОДАРСЬКОГО ПРИЗНАЧЕННЯ ЗА ДОПОМОГОЮ АЛГОРИТМІВ МАШИННОГО НАВЧАННЯ

DOI 10.36994 / 2788-5518-2023-01-05-19

Бутенко М.Є. Відкритий міжнародний університет розвитку людини «Україна», Інститут комп'ютерних технологій. Київ, Україна. <https://orcid.org/0009-0009-9545-6572>, maksym.butenko.study@gmail.com

Павленко В. І., к.ф.-м.н., доц., Відкритий міжнародний університет розвитку людини «Україна», Інститут комп'ютерних технологій. Київ, Україна. <https://orcid.org/0000-0002-3958-0415>, pavlenko.v@i.ua

Анотація: Після запуску ринку землі в Україні питання визначення ринкової вартості землі стало більш нагальним. Використання методів машинного навчання дозволяє пришвидшити та спростити передбачення вартості. В даній роботі буде застосовано методи машинного навчання для передбачення вартості землі сільськогосподарського призначення в Україні. Існує декілька досліджень які описують цю задачу, але на прикладі інших країн. Набір даних для машинного навчання зібрано з мережі Інтернет і включає в себе 800 рядків (для деяких атрибутів є пропуски даних). Зроблено огляд проблем збору таких даних. В роботі передбачення вартості землі виконане за допомогою методів машинного навчання LightGBM, Fast Tree, Fast Forest. З початкового набору даних (D1) створено три додатковий: без даних про оренду землі (D2), без даних про нормативно-грошову оцінку (НГО) (D3) та без даних про оренду та НГО (D4). Використання методу LightGBM має найкращі результати передбачення, але у деяких випадках Fast Tree наближується до LightGBM по якості. Видалення НГО з набору даних для навчання покращує якість передбачення. Також, після додаткових розрахунків виявилось, що кореляція між ринковою вартістю та НГО відсутня. Також, на відміну від передбачення вартості нерухомості, використання методів градієнтного бустінгу надає кращі результати для передбачення вартості землі. Як і в інших статтях коректне передбачення визначено як передбачення, що має не більше 10% похибки. В залежності від набору даних та обраного метода машинного навчання якість передбачення складає від 35% до 92% коректних передбачень, що близько до результатів опублікованих в попередніх дослідженнях для інших країн. Оглянуто наявні обмеження дослідження та наведено можливі шляхи покращення якості передбачення вартості землі для наведених методів машинного навчання. Зокрема, подальше розширення набору даних (як по кількості рядків, так і по кількості атрибутів) повинно покращити якість передбачення.

Ключові слова: машинне навчання, передбачення вартості землі, прогнозне моделювання, методи LightGBM, Fast Tree, Fast Forest.

AGRICULTURAL LAND PRICE PREDICTION USING MACHINE LEARNING ALGORITHMS.

Maksym Butenko. Open International University of Human Development "Ukraine" Kyiv, Ukraine. <https://orcid.org/0009-0009-9545-6572>, maksym.butenko.study@gmail.com

Volodymyr Pavlenko, Ph.D., Ass. Prof., Open University of Human Development "Ukraine", Kyiv, Ukraine. <https://orcid.org/0000-0002-3958-0415>, pavlenko.v@i.ua

Abstract: The paper considers the prediction of land value using LightGBM, Fast Tree, and Fast Forest machine learning methods. The training dataset is collected from the Internet and contains of 800 rows (in some cases there are data gaps for certain attributes). An overview of the problems of collecting such data is made. From the full dataset (D1), three additional ones were created: without data on land rent price (D2), without data on normative monetary assessment (NMA) (D3) and without data on rent price and NMA (D4). Using the LightGBM method has the best prediction results, but in some cases Fast Tree's prediction quality is similar to LightGBM. Removing NMA data from the dataset improves the quality of prediction, because after calculations it turned out that there is no correlation between market value and NMA. Also, unlike real estate value prediction, gradient boosting methods provides better results. A correct prediction is defined as a prediction with no more than 10% error, which follow similar approaches in other researches. Depending on the dataset and the chosen method of machine learning, the quality of prediction ranges from 35% to 92% of correct predictions. The paper describes research's limitations and possible ways to improve the quality of land price predictions for reviewed machine learning methods.

Keywords: *machine learning, land price prediction, prediction modeling, LightGBM, Fast Tree, Fast Forest methods.*

Вступ

Після змін у законі щодо обігу земель сільськогосподарського призначення питання визначення ринкової вартості землі в Україні стало більш нагальним [1]. Розділяють два основні підходи до визначення вартості землі – підхід, що ґрунтується на думці експерта-оцінювач та підхід, що вимагає застосування економіко-математичних та геоінформаційних моделей і алгоритмів. Другий підхід вважається більш точним, то ж було би цікаво використати машинне навчання для передбачення вартості землі сільськогосподарського призначення. В роботі Jungsun Kim, Jaewoong Won, Hyeongsoon Kim та Joonghyeok Heo [2] розглянуто питання використання алгоритмів машинного навчання для передбачення вартості землі в місті Сеул, і в цій статті буде розглянуто аналогічний підхід з використанням даних щодо вартості землі сільськогосподарського призначення в Україні.

Виконання досліджень

В статті [2] розглянута оцінка вартості землі на прикладі вартості землі не сільськогосподарського призначення у великому місті з використанням методів комп'ютерного навчання. Машинне навчання може використовувати багато атрибутів, та використовувати гнучкі специфікації моделей, що керуються даними [3]. Автори описали підхід до збору даних, що виокремив вартість землі і споруд на ній та розраховували передбачену вартість землі. В роботі використані алгоритми Random Forest (випадковий ліс) та XGBoost (реалізація підсилювального (boosting) методу машинного навчання). Автори пропонують визначити коректне передбачення вартості як передбачення що має не більше 10% похибки (і визначають що такий підхід співпадає з підходами запропонованими у [4,5]). Результатом роботи в залежності від фільтрів накладених на початковий набір даних та використаного алгоритму було від 54% до 87% коректно передбачених випадків.

В статі [3] розглянуто використання машинного навчання для реалізації методу порівняння продажів для передбачення вартості нерухомості. Як і в статті [2] автори використали підсилювальний (boosting) метод машинного навчання, але на відміну від XGBoost використали методи LightGBM та CatBoost. В якості метрик дослідження автори вибрали MAPE (середня абсолютна похибка передбачення, виражена у відсотках), RMSPE (середньоквадратична похибка передбачення, виражена у відсотках), і загальний відсоток коректних передбачень з таким самим критерієм як у статті [2] – похибка менше 10% вважається коректним передбаченням. Результатом роботи в залежності від набору даних та використаного алгоритму було від 62% до 89% коректно передбачених випадків.

В статті [6] розглянуті і підсумовані проблеми збору даних для використання алгоритмів машинного навчання для аналізу великих наборів даних в сільському господарстві (наприклад, передбачення урожаю, випадків появи захворювань, розпізнавання якості посаджених культур, тощо), в тому числі збору і аналізу даних про землю. В цілому доступ до таких даних є частково ускладненим через відсутність однієї загальнодоступної бази даних, і зазвичай збір даних про земельні ділянки вимагає використання даних з різних джерел.

Об'єктом дослідження є вартість землі. Площа України складає 603700 кв. км (трохи більше 60 мільйонів га), з яких 42.7 мільйони га – це землі сільськогосподарського призначення. Дані про ринкові пропозиції вартості ділянок землі доступні в Інтернеті, але зазвичай така інформація або не репрезентативна (має невеликий об'єм даних і не описує генеральну сукупність усіх ділянок). Наприклад сайти повідомлень, сайти аукціонів – мають інформацію лише про наявні ділянки), або має невеликий набір доступних атрибутів даних (наприклад, інформація з Державного земельного кадастру чи інформація про право власності та речові права на земельну ділянку має інформацію про власників і у деяких випадках про нормативно-грошову оцінку, але не містить інформації про якість ґрунту ділянки, тощо). Проблема збору даних описана в [6], то ж ця проблема частково існує і в інших країнах. Також, з 24.02.2023 і до завершення воєнного стану в Україні обмежено доступ до Національної структури геопросторових даних та Державного земельного кадастру, що також ускладнює збір даних для машинного навчання.

Використовуючи дані доступні у мережі Інтернет зібрано набір даних з 800 ділянок з наступними атрибутами (табл.1):

Таблиця 1.
Опис атрибутів набору даних

Назва змінної (атрибут)	Опис змінної	Опис значень
Залежні змінні		
Вартість 1 Га ділянки	Вартість у Грн	Середнє значення 56500
Незалежні змінні: географічні дані		
Місцезнаходження ділянки	Назва області, району	Наприклад, Сумська область
Географічні координати ділянки	Координати усіх кутових точок ділянки у вигляді широта та довгота.	
Площа ділянки	у Га	Середнє значення 2.9 Га
Прогнозована родючість ділянки	В залежності від типу ґрунту можна визначити приблизну родючість ділянки, у центнерах з Га	Наприклад, прогнозована родючість пшениці 52 ц/Га
Тип ґрунту	Виключено з набору даних для навчання, використовується для розрахунку значення прогнозованої родючості	Наприклад, "чорноземи типові". Дані наявні для 249 ділянок.
Незалежні змінні: дані про використання землі		
Вартість ділянки	Виключено з набору даних для навчання, використовується для розрахунку значення вартості за 1 Га	
Вартість оренди ділянки	Вартість у Грн за 1 рік	Середнє значення 10700 грн за рік за ділянку
Термін оренди	Термін у роках	Середнє значення 15.5 років
НГО (нормативно-грошова оцінка)	Грошова оцінка у Грн	571 ділянка не має інформації про НГО

Для деяких атрибутів табл.1 (наприклад прогнозована родючість, тип ґрунту) у зібраному наборі даних існують пропуски (відсутність даних).

В табл.2 наведено розподіл ділянок по регіонах.

Таблиця 2.
Розподіл по регіонах у наборі даних

Область	Кількість ділянок
Сумська	307
Чернігівська	130
Полтавська	98
Вінницька	82
Херсонська	80
Хмельницька	32
Черкаська	25
Київська	14
Харківська	11
Миколаївська	6
Івано-Франківська	4
Волинська	3
Луганська	2
Тернопільська	2
Рівненська	2
Донецька	1
Житомирська	1

В табл.3 надано приклад пропущених даних.

Таблиця 3.

Приклад наявності пропущених даних у наборі даних

Наявність даних про НГО	НГО вказаний
Так	229
Ні	571

Використовуючи зібрані дані, розширимо набір даних наступною інформацією: відстань у км від кожного обласного центра України, кількість точок у координатах ділянок (щоб визначити, чи є ділянка трикутною, чотирикутною чи багатокутною).

Для передбачення вартості землі дані були розбиті на чотири набори даних. Визначимо набір даних D1 як повний набір усіх зібраних даних, D2 як набір даних без даних про оренду, D3 як набір даних без даних про НГО, D4 як набір даних без даних про оренду та НГО. Також, випадковим чином розіб'ємо повний набір даних на 2 додаткових набори даних – тренувальний та тестовий у відношенні 7:3.

Для передбачення вартості землі використаємо мову програмування C# та фреймворк машинного навчання ML.NET. В передбаченні будуть використані методи навчання LightGBM (цей метод, як і XGBoost, відноситься до методів градієнтного бустінгу (підвищення)), Fast Tree, Fast Forest. Результати навчання будуть порівняні за допомогою таких метрик як: коефіцієнт детермінації, середньо-квадратична помилка та відсоток коректних передбачень (які будуть вираховані схожим чином як в [2,3] – якщо передбачення відрізняється від реальної ціни не більше ніж на 10% - воно вважається коректним).

Тобто, передбачення вважається коректним, якщо:

$$\left| \frac{\text{передбачене значення вартості} - \text{реальне значення вартості}}{\text{реальне значення вартості}} \right| \leq 0.1$$

Застосувавши навчання на протязі 120 секунд (навчання виконувалося на процесорі Intel Core i7-12700H) на тренувальному наборі даних отримуємо результати навчання, які наведені у табл.4 та табл.5.

Таблиця 4.

Результат машинного навчання, коефіцієнт детермінації

Номер набору даних	Коефіцієнт детермінації (RSquared)		
	LightGBM	Fast Tree	Fast Forest
D1	0.8182	0.72	0.5983
D2	0.4766	0.475	0.5046
D3	0.8092	0.8027	0.596
D4	0.4691	0.539	0.5051

Таблиця 5.

Результат машинного навчання, середньо-квадратична помилка

Номер набору даних	Середньо-квадратична помилка (RMS-loss)		
	LightGBM	Fast Tree	Fast Forest
D1	9663	11177	14124
D2	16192	15787	15764
D3	9831	10535	14155
D4	16587	15673	15778

Детальний опис метрик (RSquared, RMS-loss) моделі доступний у [7]. Як видно з результатів, найкращим можна вважати використання методу LightGBM.

Виконаємо розрахунок важливості атрибута за допомогою метода визначення важливості індексу перестановки (Permutation Feature Importance, PFI). Для можливості порівняння значень між різними методами машинного навчання приведемо значення важливості до 1 та визначимо 7 найважливіших атрибутів для кожного з методів. Важливість атрибутів для набору D1 в залежності від методу наведено у табл.6.

Таблиця 6.
Важливість атрибутів для набору даних D1

Пріоритет важливості	Важливість атрибута		
	LightGBM	Fast Tree	Fast Forest
1	Орендна плата (0.474)	Орендна плата (0.406)	Відсоток орендної плати (0.286)
2	Площа (0.251)	Площа (0.299)	Відстань від м. Вінниця (0.194)
3	Відсоток орендної плати (0.158)	Відсоток орендної плати (0.223)	Орендна плата (0.286)
4	Відстань від м. Вінниця (0.03)	Орендна плата після сплати податків (0.016)	Координата широти (0.098)
5	Координата широти (0.02)	Відстань від м. Вінниця (0.014)	Відстань від м. Одеса (0.064)
6	Відстань від м. Миколаїв (0.013)	Відстань від м. Київ (0.008)	Відстань від м. Харків (0.048)
7	Відстань від м. Харків (0.011)	Відстань від м. Чернівці (0.005)	Відстань від м. Житомир (0.043)

Як показано в табл.6, атрибут орендної плати має сильний вплив на результат передбачення. Без даних про орендну плату (D2) отримаємо важливість атрибутів що наведені у табл.7.

Таблиця 7
Важливість атрибутів для набору даних D2

Пріоритет важливості	Важливість атрибута		
	LightGBM	Fast Tree	Fast Forest
1	Відстань від м. Вінниця (0.219)	Відстань від м. Вінниця (0.254)	Відстань від м. Вінниця (0.262)
2	Площа (0.109)	Координата широти (0.165)	Координата широти (0.180)
3	Координата широти (0.088)	Площа (0.117)	Відстань від м. Одеса (0.115)
4	Відстань від м. Одеса (0.071)	Відстань від м. Харків (0.066)	Площа (0.065)
5	Відстань від м. Чернівці (0.067)	Відстань від м. Миколаїв (0.063)	Відстань від м. Київ (0.062)
6	Відстань від м. Харків (0.061)	Відстань від м. Одеса (0.037)	Відстань від м. Житомир (0.051)
7	Відстань від м. Суми (0.06)	Відстань від м. Київ (0.032)	Відстань від м. Харків (0.05)

Як показано в табл.7, атрибут про НГО не є важливим, бо в початковому наборі даних ці дані присутні лише у 229 випадках. Можна також припустити, що дані про НГО часто розходяться з ринковою вартістю землі, то ж додавання НГО не покращує результат машинного навчання. Якщо перевірити це за допомогою функції кореляції по даним про ринкову вартість землі та НГО, то отримаємо що кореляція дорівнює -0.0622, тобто припущення підтверджується.

В табл.8 наведено відсоток коректних передбачень. В залежності від набору даних, вибраного метода машинного навчання отримаємо від 35 до 92 відсотків коректних передбачень, що в цілому збігається з результатами отриманими в [2,3]. На наборі даних D1 найкращим методом був LightGBM, у той час як найгіршим – Fast Forest. При видаленні даних про НГО, Fast Tree метод значно покращує результат передбачення, то ж цей метод на наявному наборі даних має більші вимоги до якості даних та їх підготовки до машинного навчання.

Таблиця 8.
Відсоток коректних передбачень

	Відсоток коректних передбачень											
	LightGBM				Fast Tree				Fast Forest			
	1	2	3	4	1	2	3	4	1	2	3	4
Тренувальний набір даних	1.2	5.7	1.7	2.6	4.1	3.2	2.5	3.5	6.9	0	7.6	0
Тестовий набір даних	7.5	5.4	3.2	5	3.7	7	7	2	7.5	3	6.6	2.9
Повний набір даних	7.1	5.6	8.2	3.3	1	1.3	7.8	3.1	7.1	1.2	7.3	0.8

Щоб додатково порівняти результат передбачення за допомогою методів машинного навчання спробуємо визначити певну базову модель, модель без використання методів машинного навчання. Наприклад, таку модель можна визначити як – передбачення вартості робиться як середнє значення вартості землі ділянок що знаходяться поруч (наприклад, у радіусі 40 км). Такий підхід називається порівняльним аналізом ринку (Comparative Market Analysis) та оглянутий, наприклад у [10].

Результат використання такого передбачення наведено у табл.9, то ж ця модель гірша за будь який результат роботи машинного навчання.

Таблиця 9.
Результат передбачення для базової моделі

Середньо-квадратична помилка (RMS-loss), у Грн	17556.87
Кількість коректних передбачень	219
Кількість не-коректних передбачень	581
Відсоток коректний передбачень	27.35

Під час огляду цієї моделі також були виконані розрахунки для менших значень радіуса – у такому випадку середньо-квадратична помилка підвищувалася на 40-50% (до ~25000), а при збільшенні радіуса – результат передбачення майже не змінювався. Також, у деяких випадках ділянок поруч в наборі даних може не бути, тобто використання базового методу має ряд обмежень.

У статтях [8, 9] відзначено, що для передбачення вартості нерухомості кращі результати повинні давати методи Fast Forest. Результати даного дослідження показують, що використання

методу градієнтного бустінгу надає кращі результати для передбачення вартості землі ніж методи Fast Tree, Fast Forest. Це твердження вимагає додаткового дослідження, але в дослідженні [2] також зазначалася перевага методів градієнтного бустінгу над Fast Forest для передбачення вартості землі. Також, варто відмітити важливість підготовки даних та чутливість Fast Tree до наявності даних з пропусками. На прикладі використання набору даних D3 – якість передбачення зросла після того як було видалено поле з інформацією про НГО.

Обмеженнями даного дослідження можна відзначити дві основні проблеми – обмеження накладені початковим набором даних (D1) та складності побудови моделі машинного навчання, що буде надавати стабільне і якісне передбачення ціни довільної ділянки.

Обмеження початкового набору даних, такі як наявність лише 800 ділянок, деякі пропуски для певних ділянок можна покращувати. Також, як зазначалося вище – до кінця воєнного стану в Україні доступ до частини даних в державних реєстрах частково обмежений, то ж після закінчення воєнного стану кількість даних доступних для аналізу зросте. Можливим розширенням набору даних може бути інформація що не стосується безпосередньо земельної ділянки (наприклад, кількість жителів в найближчому населеному пункті, чи середня заробітна плата у регіоні, тощо). Складності використання машинного навчання, такі як отримання різних результатів при використанні різних методів машинного навчання, отримання різних значень важливості атрибутів, необхідність навчання з використанням різних гіперпараметрів методу, тощо також була відзначена в [2].

Висновки

В роботі було розглянуто використання методів машинного навчання для передбачення вартості землі. Використано три методи машинного навчання – LightGBM, Fast Tree, Fast Forest на наборі даних, зібраного з мережі Інтернет, що включає інформацію про 800 ділянок. Використання LightGBM має кращі результати ніж використання Fast Tree, Fast Forest, але у випадку видалення атрибутів що має пропуски (наприклад НГО) результат використання Fast Tree наближується до LightGBM. Результати дослідження можна продовжувати – покращення якості початкових даних та розширення кількості атрибутів повинно покращити точність передбачення.

Література

1. Черевко Г., Грисьо. М. ОЦІНКА І ЦІНА ЗЕМЛІ В УКРАЇНІ ТА ДОСВІД КРАЇН ЄВРОПИ / Г. Черевко, М. Грисьо. // Наукові записки Національного університету «Острозька академія». Серія «Економіка». – 2023. – №28. – С. 4–13.
2. Machine-Learning-Based Prediction of Land Prices in Seoul, South Korea / K.Jungsun, W. Jaewoong, K. Hyeongsoon, H. Joonghyeok. // Sustainability. – 2021. – №13(23).
3. Yunjong K. Applying Comparable Sales Method to the Automated Estimation of Real Estate Prices / K. Yunjong, C. Seungwoo, Y. Mun Yong. // Sustainability. – 2020. – №12(14).
4. Crosby N. Property valuation variation and the 'margin of error' in the UK / N. Crosby, A. Lavers, J. Murdoch. // Journal of Property Research. – 1998. – №15. – С. 305–330.
5. Watkins C. The definition and identification of housing submarkets / Craig Watkins. // Environment and Planning. – 2001. – №33. – С. 2235 – 2253.
6. Challenges to Use Machine Learning in Agricultural Big Data: A Systematic Literature Review / A.Craverio, S. Pardo, S. Sepulveda, L. Munoz. // 2022. – №12.
7. Evaluate your ML.NET model with metrics [Електронний ресурс] // microsoft.com. – 2023. – Режим доступу до ресурсу: <https://learn.microsoft.com/en-us/dotnet/machine-learning/resources/metrics>.
8. A Fully Automated Adjustment of Ensemble Methods in Machine Learning for Modeling Complex Real Estate Systems / [J. Alfaro-Navarro, E. Cano, E. Alfaro-Cortes та ін.]. // Complexity. – 2020. – №2020. – С. 1–12.
9. Quang T. Housing Price Prediction via Improved Machine Learning Techniques / T. Quang, N. Minh, D. Hy. // Procedia Computer Science. – 2020. – №174. – С. 433–442.
10. What Is Comparative Market Analysis (CMA) in Real Estate? [Електронний ресурс] // Investopedia. – 2022. – Режим доступу до ресурсу: <https://www.investopedia.com/terms/c/comparative-market-analysis.asp>.

References

1. Cherevko H., Hryso. M. OTSINKA I TSINA ZEMLI V UKRAINI TA DOSVID KRAIN YEVIROPY / H. Cherevko, M. Hryso. // Naukovi zapysky Natsionalnoho universytetu «Ostrozka akademiia». Serii «Ekonomika». – 2023. – №28. – С. 4–13.2. Machine-Learning-Based Prediction of Land Prices in Seoul, South Korea / K.Jungsun, W. Jaewoong, K. Hyeongsoon, H. Joonghyeok. // Sustainability. – 2021. – №13(23). [in English]
3. Yunjong K. Applying Comparable Sales Method to the Automated Estimation of Real Estate Prices / K. Yunjong, C. Seungwoo, Y. Mun Yong. // Sustainability. – 2020. – №12(14). [in English]

4. Crosby N. Property valuation variation and the 'margin of error' in the UK / N. Crosby, A. Lavers, J. Murdoch. // *Journal of Property Research*. – 1998. – №15. – p. 305–330. [in English]
5. Watkins C. The definition and identification of housing submarkets / Craig Watkins. // *Environment and Planning*. – 2001. – №33. – p. 2235 – 2253. [in English]
6. Challenges to Use Machine Learning in Agricultural Big Data: A Systematic Literature Review / A.Cravero, S. Pardo, S. Sepulveda, L. Munoz. // 2022. – №12. [in English]
7. Evaluate your ML.NET model with metrics [Електронний ресурс] // microsoft.com. – 2023. – Режим доступу до ресурсу: <https://learn.microsoft.com/en-us/dotnet/machine-learning/resources/metrics>. [in English]
8. A Fully Automated Adjustment of Ensemble Methods in Machine Learning for Modeling Complex Real Estate Systems / [J. Alfaro-Navarro, E. Cano, E. Alfaro-Cortes та ін.]. // *Complexity*. – 2020. – №2020. – p. 1–12. [in English]
9. Quang T. Housing Price Prediction via Improved Machine Learning Techniques / T. Quang, N. Minh, D. Hy. // *Procedia Computer Science*. – 2020. – №174. – p. 433–442. [in English]
10. What Is Comparative Market Analysis (CMA) in Real Estate? [Digital resource] // Investopedia. – 2022. – Resource access: <https://www.investopedia.com/terms/c/comparative-market-analysis.asp>.

УДК 004.42:004.8

КОМБІНОВАНІ ПІДХОДИ ДО СТАТИЧНОГО АНАЛІЗУ КОДУ З ВИКОРИСТАННЯМ НЕЙРОННИХ МЕРЕЖ

DOI 10.36994 / 2788-5518-2023-01-05-20

Вохранов І.А. Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна. vokhranov@gmail.com

Булах Б.В., к.т.н., доц. Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна. bogdan.bulakh@gmail.com

Анотація: Ця стаття презентує огляд можливих підходів до застосування нейронних мереж в процесі статичного аналізу коду. У ній досліджується поточний стан речей у наявних підходах до покращення аналізу програм за допомогою методів машинного навчання із використанням пост-обробки сповіщень статичних аналізаторів, попередньої обробки коду, чи безпосереднього використання машинного навчання для аналізу вихідного коду. Окрім того, у статті досліджуються основні напрямки застосування підходів кожної категорії. Як класичні підходи, так і методи машинного навчання в аналізі програм мають чітко виражені сильні та слабкі сторони, які необхідно враховувати при їх практичному впровадженні. Однією з основних тез цього дослідження є те, що для побудови високоякісної системи вкрай важливо розуміти можливості комбінування даних підходів, використовуючи гнучкість, яку пропонують нейронні мережі, та зберігаючи при цьому достатній рівень надійності, що забезпечується класичними алгоритмами. Стаття покриває наступні три базові напрямки використання нейронних мереж для статичного аналізу коду. Перший напрямок – це коригування специфікацій: уточнення специфікацій, отриманих від класичного статичного аналізатора коду (усунення, кластеризація, ранжування попереджень або просто спрощення аналізу попереджень вручну, тощо). Другий напрямок – визначення специфікацій, прихованих у коді (вилучення та відбір властивостей, перетворення коду зі збереженням його поведінки, наприклад, з метою зробити його більш підходящим для класичних інструментів статичного аналізу). Третій напрямок – аналіз вихідного коду як "чорного ящика" з метою виявлення та виправлення дефектів (синтаксичних чи семантичних, або ж вразливостей) у коді, асистування при перевірці коду вручну, автоматичного форматування чи виявлення коду, що "погано пахне" (в цьому напрямку для аналізу використовується лише модель машинного навчання, тренування якої відбувається прямо на вихідному коді). Стаття окреслює коло можливих напрямків подальших досліджень, що фокусуються на розвитку та комбінуванні розглянутих підходів.

Ключові слова: статичний аналіз коду; нейронні мережі; машинне навчання; аналіз коду; виявлення дефектів; виправлення дефектів; коригування специфікацій; постобробка попереджень; визначення специфікацій; попередня обробка коду; навчання на вихідному коді.

COMBINED APPROACHES TO THE STATIC CODE ANALYSIS USING NEURAL NETWORKS

Illia Vokhranov. National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine. vokhranov@gmail.com

Bogdan Bulakh, Ph.D., Ass. Prof. National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine. bogdan.bulakh@gmail.com

Abstract: This article presents an overview of possible approaches to the application of neural networks in the process of static code analysis. It explores the current state of affairs in existing approaches to improving program analysis using machine learning methods, including postprocessing of static analysis alerts, preprocessing of source code, or direct use of machine learning for analyzing source code. Additionally, the article examines the main directions for applying approaches from each category. Both classical approaches and machine learning methods in program analysis possess distinct strengths and weaknesses that should be considered when implementing them in practice. One of the main theses of this research is that understanding the capabilities of combining these approaches, leveraging the flexibility offered by neural networks while maintaining a sufficient level of

reliability provided by classical algorithms, is crucial for building a high-quality system. This article covers the following three basic directions of the application of neural networks for the static source code analysis. The first direction is a specification tuning: a refinement of specifications produced by a 'classic' static code analyzer (a removal, clustering, ranking of warnings or just assistance in manual warning analysis, etc.). The second direction is a specification inference, to find specifications hidden in code (feature extraction, selection, or code transformation retaining its behaviour, e.g. to make it more suitable for the 'classic' static analysis tools). The third way is a black box analysis to discover and fix code defects (syntactic, semantic ones or vulnerabilities), to assist in manual code checking, to format the code automatically or to find code smells (in this direction only a machine learning model is used, its training is performed on the source code directly). The article outlines directions for the future research which will focus on the development and combining of the approaches covered here.

Keywords: static code analysis; neural networks; machine learning; code analysis; defect detection; defect correction; specification tuning; postprocessing of alerts; specification inference; preprocessing of source code; learning on source code.

Вступ

Забезпечення відсутності дефектів в процесі розробки програмного забезпечення є надзвичайно важливим для стабільної, безперебійної роботи програмних систем. Статичний аналіз програмного коду є одним із ключових інструментів пошуку дефектів у програмному коді. Під час пошуку помилок та вразливостей традиційні підходи статичного аналізу покладаються на математичні методи, які надають чіткий, детермінований результат з підтвердженням чи спростуванням певного твердження про властивість програми. Через це такі методи не можуть навчитися покладатися на шаблони, які присутні у вихідному коді, або ймовірно обробляти неоднозначну інформацію, яка часто трапляється в реальному коді.

Основною проблемою, що виникає при застосуванні класичних математичних методів у статичному аналізі, є недостатня їх гнучкість і адаптивність. Через детермінованість результатів така система перевірки має вищий показник хибно позитивних результатів (повідомляє користувачеві про можливі дефекти чи вразливості, коли вони відсутні). Водночас, через нездатність таких методів працювати з шаблонами в коді, складніші проблеми залишаються поза увагою.

Використання нейронних мереж у процесі статичного аналізу може бути потенційним розв'язанням проблем, присутніх у класичних підходах. На основі багатьох досліджень, проведених у цьому напрямку за останні роки, зрозуміло, що використання нейронних мереж в якості інструмента для аналізу коду має потенціал. Але, разом з тим, такий підхід має і певні труднощі та часто не здатний повністю замінити класичні підходи.

Виконання досліджень

Статичний аналіз коду загалом — це процес виявлення помилок і дефектів у вихідному коді програмного забезпечення. Він походить від одного з найстаріших і найефективніших методів виявлення вад — *перевірки коду вручну* (англ. *code review*). Принцип перевірки коду полягає в перегляді вихідного коду іншими розробниками (не автором) і наданні рекомендацій щодо його вдосконалення. Цей процес дозволяє групі розробників виявляти помилки та проблеми у вихідному коді, які можуть призвести до збоїв і вразливостей програми в майбутньому.

Основна проблема ручної перевірки коду полягає в тому, що вона досить дорога: розробникам доводиться витрачати багато часу та зусиль на перевірку будь-яких нових змін, внесених у код програми.

Статичний аналіз коду у вузкому трактуванні – це спрощена, автоматизована перевірка коду. Його застосування не вимагає виконання програми, не залежить від середовища розробки та може бути виконаним безпосередньо на етапі написання коду. Це набагато швидше, ніж перевірка вручну, і забезпечує повне охоплення коду. Своєчасне надання інформації про проблеми у вихідному коді суттєво знижує витрати часу та грошей у комерційній розробці. Національний Інститут Стандартів і Технологій (NIST) зазначає, що відносна вартість виправлення дефектів, виявлених на різних етапах розробки програмного забезпечення, кратно зростає з кожним наступним етапом розробки. Так, наприклад, оцінюється, що вартість виправлень коду, виконаних після публічного випуску програми, можуть досягати 30-кратної вартості порівняно з виправленнями на етапі її розробки [1].

Форми статичного аналізу та завдання, які він виконує, можуть бути досить різними, наприклад: виявлення вразливостей, допоміжні перевірки коду, де-дуплікація коду, оптимізація тощо. Інструменти статичного аналізу є високоефективними у забезпеченні певного рівня якості розробки, завдяки здатності автоматично виявляти типові помилки програмування (як ділення на нуль, вихід за межі масиву та багато інших). Під час цієї процедури, вихідний код піддається серії формальних тестів, які спираються на певну базу даних вразливостей та помилок, і, на основі цього, аналізатор надає список виявлених проблем. Однак, розв'язання цих проблем ще не гарантує абсолютного захисту від помилок у коді, оскільки кожна процедура статичного аналізу здатна визначати наявність проблеми у вихідному коді, але не може повністю гарантувати її відсутність.

Якщо інструменти статичного аналізу не можуть визначити, чи є об'єкт розгляду помилкою, вони видають попередження, яке сповіщає користувача про *потенційну* помилку в даному місці. Через ряд причин, таких як абстракції, що використовуються цими інструментами, та компроміс між точністю та масштабованістю аналізу, вони генерують досить велику кількість попереджень. Користувачеві потрібно вручну перевірити ці попередження та виділити справжні помилки серед інших – хибно позитивних результатів (помилкових спрацьовувань). Відповідно до дослідження, проведеного у [2], велика кількість помилкових попереджень та зусилля, необхідні для їх ручної перевірки, вважаються основними проблемами, пов'язаними з використанням інструментів статичного аналізу на практиці.

Останнім часом дослідники почали працювати в напрямку створення більш інтелектуальних систем аналізу коду, які могли б не тільки повідомляти про помилки на основі наявної бази знань, але й вчитися на помилках, зроблених у попередніх раундах аналізу, щоб підвищити свою точність у майбутніх. Використання алгоритмів машинного навчання в процесі статичного аналізу коду потенційно може бути підходом, який допоможе покращити можливості аналізу програмного забезпечення. Основною метою даної статті є аналіз можливих підходів до залучення методів машинного навчання для удосконалення традиційних механізмів статичного аналізу коду.

Аналіз програм на основі машинного навчання дає можливість виявляти проблеми на основі неоднозначної та часткової інформації. Вивчаючи загальні моделі програмного коду, такі методи можуть запропонувати гнучкі ймовірнісні результати для аналізу різних аспектів поведінки програми. При цьому, методи машинного навчання не обов'язково намагаються витіснити традиційні методи статичного аналізу. Натомість використання таких інструментів, як нейронні мережі, може бути цінним доповненням до загального набору методологій аналізу програмного забезпечення. Сфера використання машинного навчання в аналізі коду ще відносно молода, але наявні результати показують її потенціал.

Традиційні підходи статичного аналізу зазвичай проводять аналіз на основі чітких формальних методів і набору попередньо визначених інструкцій. Однак, визначення такого набору правил є виснажливим ручним завданням, яке досить важко масштабувати до широкого діапазону властивостей і програм. Формальні методи є незамінним інструментом під час роботи з програмами, де безпека має вирішальне значення, але водночас існує широкий спектр програм, які втрачають можливість отримати переваги від більш гнучких підходів аналізу програмного забезпечення. Аналіз програм на основі машинного навчання має на меті розв'язати цю проблему, але жертвує здатністю надавати гарантії. Перш за все, машинне навчання може допомогти впоратися з двома найпоширенішими джерелами неоднозначності: прихованими специфікаціями та неоднозначними контекстами виконання. Поєднання задач статичного аналізу, в класичному його розумінні, та методів машинного навчання, може мати досить різні форми. В рамках своєї роботи [3], М. Allamanis розділяє всі підходи із застосуванням методів машинного навчання в статичному аналізі програмного коду на 3 категорії: **коригування специфікацій** (англ. *specification tuning*), **визначення специфікацій** (англ. *specification inference*) та **аналіз “чорного ящика”** (англ. *black box analysis learning*). Перші дві категорії є комбінованими підходами, що поєднують в собі застосування машинного навчання і традиційних інструментів, а третя категорія базується лише на використанні машинного навчання. У даній роботі, дана класифікація була взята за основу розгляду підходів застосування нейронних мереж в статичному аналізі коду, і в наступних трьох підрозділах кожен із вказаних форм буде розібрано детальніше.

Коригування специфікацій — це підхід, у якому модель нейронної мережі використовується, щоб уточнити та покращити специфікації, отримані в результаті “класичного” статичного аналізу

коду. В літературі даний підхід також часто зустрічається як *“постобробка (англ. postprocessing) попереджень статичного аналізу”*. Метою коригування специфікацій є надання більш лаконічних і точних уявлень про очікувану поведінку програми, з точки зору її потенційних проблем та вразливостей.

Коли експерт має намір створити високонадійний аналіз програми та налаштовує занадто чутливі правила, це зазвичай призводить до високого рівня хибно позитивних результатів (помилкових попереджень про наявність проблем в коді). Виклик великої кількості помилкових попереджень призводить до ігнорування користувачами істинно позитивних результатів, що сильно знижує користь такого аналізу [3]. Через це, з метою розв'язання проблем хибно позитивних результатів, почали з'являтися дослідження із використанням методів машинного навчання для коригування (постобробки) аналізу програми шляхом вивчення того, які аспекти формального аналізу можуть бути відкинутими для підвищення фінальної точності [4].

Таким чином, коригування специфікацій полягає саме в обробці попереджень після їх генерації, а його основними цілями є зменшення кількості отримуваних попереджень та зусиль, необхідних для їх ручної перевірки. Зазвичай, застосування обробки попереджень статичного аналізу в даному підході приймає одну з наступних форм [4]:

- зменшення кількості попереджень перед наданням їх користувачеві;
- зменшення зусиль, необхідних для ручної перевірки, шляхом розширення попереджень додатковою інформацією;
- спрощення процесу ручної перевірки, шляхом асистування чи надання допоміжних інструментів під час перевірки попереджень.

Враховуючи всі перспективи, наявні дослідження пропонують широкий набір підходів до постобробки попереджень, методи реалізації яких досить різноманітні. В дослідженні публікацій сфери постобробки попереджень статичного аналізу, виконаному в роботі [2], автори розділяють всі підходи на шість категорій, які будуть розглянуті нижче.

З метою зменшення загальної кількості попереджень, які потрібно переглянути для виявлення загальних шаблонів чи безпосередніх причин дефектів, може застосовуватись кластеризація попереджень. Цей підхід полягає в об'єднанні схожих попереджень в декілька груп (кластерів) на основі подібності або кореляції між ними. Подібність попереджень може визначатись за такими критеріями, як тип дефекту, розташування в коді або серйозність попередження.

Для забезпечення кращого управління списком попереджень, його можна ранжувати відповідно до їх пріоритету чи важливості. Таке упорядкування може виконуватись на основі різних характеристик попереджень, таких як: розміщенні у вихідному коді, критичності впливу дефекту на систему, ймовірністю істинно позитивного результату, попередній історії виправлення помилок, зусиллями, необхідними для його усунення тощо. Такий підхід може допомогти користувачеві зосередитися спершу на найбільш релевантних попередженнях, ігноруючи менш актуальні.

Зменшення загальної кількості попереджень також може відбуватись шляхом усунення певної їх частини. За такого підходу, усі попередження класифікуються на дві категорії: дійсні та недійсні, після чого недійсні попередження прибираються. Таким чином, усуваються ті попередження, які вважаються зайвими, нерелевантними чи нецікавими з певних причин, наприклад: дублікати, конкретні типи дефектів, дефекти, що не мають критичного впливу тощо. Усунення попереджень може відбуватись, як на основі синтаксичного чи семантичного аналізу, так і на основі налаштувань самого користувача. Завдяки звуженню діапазону попереджень, які потребують опрацювання, даний підхід може допомогти зберегти час та ресурси користувача.

Менша кількість попереджень може досягатись шляхом зниження рівня хибно позитивних результатів. Для цього, отримані попередження перевіряються за допомогою додаткових методів, що спираються на додаткову інформацію в коді й здатні автоматично ідентифікувати та усунути хибно позитивні результати. Так помилкові попередження можуть бути усунуті без ручної перевірки, що, своєю чергою, збільшує довіру користувача до решти отриманих попереджень.

Покращення якості отриманих результатів може також відбуватись за допомогою динамічного аналізу. В такому випадку, результати інструментів статичного аналізу передаються на обробку інструментами динамічного аналізу для створення тестів, що підтверджують наявність помилки. Цей підхід дозволяє підвищити точність та комплексність виявлення дефектів, і зменшити кількість хибно позитивних та хибно негативних результатів.

Ручна перевірка попереджень може бути полегшена наданням додаткової інформації чи забезпеченням користувача інструментальною підтримкою. Така підтримка може приймати вигляд пояснень, рекомендацій, візуалізації, інтерактивних інтерфейсів тощо, і може здійснюватись за допомогою різних методів, таких як генерація природної мови, системи підтримки прийняття рішень, надання графічного інтерфейсу, або гейміфікації. Такий підхід може допомогти підвищити ефективність та результативність ручної обробки попереджень, а також зберегти час та зусилля користувача.

Визначення специфікацій — це підхід, у якому модель машинного навчання використовується для того, щоб краще виділяти специфікації, присутні в програмному коді. Метою визначення специфікацій є створення більш точного та повного представлення очікуваної поведінки програми. Вважається, що більшість коду відповідає деяким прихованим специфікаціям, і саме тому можуть бути корисними моделі машинного навчання, що здатні вивчати шаблони в програмному коді та виводити приховані форми цих специфікацій [3]. Таким чином, даний підхід передбачає застосування певних перетворень чи модифікацій до вихідного програмного коду перед використанням інструментів статичного аналізу, з метою покращення якості, або ефективності виявлення дефектів.

В задачах проведення попередньої обробки програмного коду, серед методів машинного навчання на сьогодні найбільш результативними є *рекурентні* та *графові нейронні мережі*.

Рекурентні нейронні мережі (англ. Recurrent neural networks, або RNN) дуже добре справляються з обробкою послідовностей, в тому числі таких, як природна мова чи вихідний програмний код. Вони не потребують заздалегідь заданих параметрів, а натомість здатні працювати з послідовностями токенів: необроблений вихідний код розбивається на невеликі одиниці (наприклад, “*int*”, “*func*”, “{” та “}”) і представляється в моделі у вигляді їх послідовності. RNN можуть захоплювати як синтаксичну, так і семантичну інформацію коду та використовувати її для виконання різноманітних завдань, наприклад, класифікації [5, 6].

Графові нейронні мережі (англ. Graph neural networks, GNN) є типом нейронних мереж, призначених для обробки даних з графовою структурою, таких як абстрактні синтаксичні дерева (англ. abstract syntax trees, AST) або графи управління потоком програми (англ. control flow graphs, CFG). Вони навчаються на основі вузлів та ребер графу, захоплюючи структурну та контекстуальну інформацію коду, використовуючи її для класифікації. GNN використовуються для моделювання структури та залежностей програмного коду, і можуть бути поєднані з інструментами статичного аналізу, щоб покращити точність аналізу програмного забезпечення. Наприклад, система ASTGNN [7] використовує графову нейронну мережу для моделювання AST програмного коду та виявлення помилок програмного забезпечення [3, 5].

Попередня обробка програмного коду за допомогою машинного навчання може приймати різні форми. В загальному, їх можна розділити на три основні групи: *вилучення властивостей* (англ. feature extraction), *відбір властивостей* (англ. feature selection), та *перетворення коду* (англ. code transformation).

Вилучення властивостей – вилучення відповідних функцій чи атрибутів із вихідного коду, на основі яких інструменти статичного аналізу матимуть можливість ефективніше та точніше визначати дефекти. Наприклад, методи машинного навчання можуть використовуватись для вилучення синтаксичних, семантичних або структурних властивостей коду за допомогою таких структур, як абстрактні синтаксичні дерева (AST), графи потоку даних (англ. data flow graphs, DFG), графи потоку керування (CFG) і т.д.

Відбір властивостей – вибір підмножини функцій, які є найбільш відповідними або інформативними для виявлення дефектів, та відкидання решти. Такий підхід може допомогти зменшити розмірність і складність обробки коду та покращити продуктивність інструментів статичного аналізу, а також сфокусувати їх на найбільш важливих ділянках коду. Наприклад, методи машинного навчання можуть використовуватись із застосуванням методів фільтрації (англ. filter methods) або обгортання (англ. wrapper methods) для вибору властивостей на основі певних критеріїв, таких як кореляція, взаємна інформація або точність класифікації.

Перетворення коду – модифікація вихідного коду без зміни його функціональності та поведінки, з метою зробити його більш зручним чи сумісним для інструментів статичного аналізу. Тобто, підвищення ефективності статичного аналізу відбувається внаслідок поліпшення представлення та якості коду. Наприклад, методи машинного навчання можуть

використовуватись для застосування методів рефакторингу коду, нормалізації коду, обфускації або синтезу коду, з метою надання коду більш придатної для обслуговування форми.

Аналіз “чорного ящика” — це підхід, при якому для аналізу програми використовується лише модель машинного навчання, тренування якої відбувається безпосередньо на вихідному коді. В такому підході, модель діє за принципом “чорного ящика”, генеруючи попередження в якості результатів аналізу програми, але при цьому ніколи явно не формулює конкретні специфікації.

Така форма аналізу програм є дуже гнучкою та може виходити за межі можливостей багатьох традиційних підходів. Однак, в обмін на свою гнучкість, аналіз “чорного ящика” жертвує ясністю та не надає жодних гарантій. Відсутність прозорості в процесі прийняття рішень моделі може ускладнити розуміння того, чому виникли певні попередження, або які кроки слід вжити для їх вирішення. А відсутність конкретних специфікацій означає відсутність гарантій, що отримані попередження точні, або що вони повністю покривають усі необхідні вимоги щодо потенційних проблем в коді. Тому, під час практичного застосування такого підходу, необхідно враховувати ці особливості, особливо у випадках, коли потрібен високий рівень впевненості в результатах аналізу. [3]

Як уже зазначалось, однією з особливостей застосування нейронних мереж в аналізі програмного коду є їх гнучкість та здатність виходити за рамки можливостей традиційних підходів. Завдяки цьому, вони мають досить широкий спектр можливостей практичного застосування, частину з яких розглянемо далі.

Автоматизоване виявлення дефектів (англ. automated defect detection). Основною задачею статичного аналізу є виявлення дефектів. Тому, одним із найбільш очевидних способів застосування методів машинного навчання в аналізі програмного коду є використання моделі безпосередньо для виявлення дефектів у ньому. Такий процес передбачає побудову класифікаторів, які здатні визначати області коду, що потенційно містять дефекти, спираючись на шаблони в програмному коді та використовуючи таку інформацію, як синтаксис, складність коду, семантику чи історію змін.

У контексті статичного аналізу програмного коду, під *дефектами* розуміються будь-які відхилення між очікуваною та реальною поведінкою програми. З точки зору комплексності проблем та направленості аналізу, в загальному, дефекти можуть бути розділені на три рівні (категорії): *синтаксичні дефекти*, *семантичні дефекти* та *вразливості*. Нижче, у таблиці 1, описано кожен із цих рівнів [6]. *Автоматизоване виправлення дефектів (англ. automated defect correction).* Наступним, більш амбіційним етапом після виявлення дефектів є їх автоматизоване виправлення. У такому процесі передбачається, що модель має автоматично визначити правки, необхідні для виправлення конкретної помилки, які надалі можуть бути застосовані з мінімальним втручанням, або й без втручання людини. У порівнянні з виявленням, виправлення є набагато складнішою задачею, яка стала реалістичною темою дослідження лише нещодавно, завдяки можливості використання останніх напрацювань сфери *обробки природних мов (англ. natural language processing, NLP)*. За такого підходу навчання моделей найчастіше виконується або на вибірках з парами неправильних та правильних елементів програм, де модель навчається перетворювати одні на інші, або на вибірках з правильними екземплярами, де модель навчається перетворювати програми, які мають від них відхилення [6]. Одним із прикладів, що демонструють перспективи застосування методів обробки природних мов до програмного коду, є нещодавній революційний проєкт у сфері NLP – *Chat GPT* [8], який іноді здатен демонструвати досить хороші результати у питаннях генерації та виправлення коду, хоч все ще і потребує значного ручного контролю результатів.

Асистування при перевірці коду вручну (англ. code review assistance). Одним із важливих інструментів у забезпеченні якості розробки програмного забезпечення є *перевірка коду вручну*, яка призначена не лише для пошуку помилок та покращення загальної якості програмного забезпечення, а й для постійної передачі знань між розробниками. У великих проєктах якісна перевірка коду зазвичай є складним завданням. Оскільки вона сильно залежить від індивідуальних можливостей рецензентів, часто залишається певний набір змін вихідного коду, що не був перевірений належним чином [9]. Підняття якості та швидкості ручної перевірки коду може досягатися шляхом використання методів машинного навчання, які будуть автоматично визначати фрагменти коду, що потребують особливої уваги та додаткової перевірки. Такий спосіб поєднання автоматизації з перевіркою коду вручну може покращити якість рецензування коду, зберігаючи при цьому можливість підтримувати передачу знань між розробниками [10].

Асистування при ручній перевірці коду часто може опиратись на підходи, що використовуються при автоматизованому виявленні дефектів. Однією з перспективних технік є використання машинного навчання в поєднанні з системою контролю версій і системою відстеження помилок. Передбачення дефектів на основі ітерацій змін коду може використовуватись для зосередження більшої уваги розробників на потенційно проблемних ділянках коду під час етапу ручної перевірки. Завдяки такому асистуванню може підтримуватись постійний базовий рівень якості перегляду коду, що буде зменшувати ймовірність непоміченого проходження помилок в кодову базу, і, як результат, – суттєво покращувати кінцеву якість програмного забезпечення з меншими експлуатаційними витратами [9; 10]. *Автоматичне оформлення / форматування коду (англ. automatic code stylizing / formatting)*. Підтримка сталості стилю коду є дуже важливим аспектом розробки програмного забезпечення. Такі елементи форматування, як відступи, найменування, коментування, порожні рядки та потік керування (порядок команд, інструкцій та викликів) мають досить сильний вплив на зрозумілість вихідного коду [11]. Автоматичні інструменти форматування можуть значно покращити якість оформлення коду та зменшити зусилля розробників необхідні для його підтримки. Такі інструменти особливо актуальні у великих проектах, де постійний контроль будь-яких змін коду є дуже складним. Здебільшого, системи автоматичного форматування коду побудовані на основі правил, проте, вони також можуть будуватись на основі машинного навчання [12; 13].

Таблиця 1.
Категорії дефектів програмного коду

Категорія дефекту	Де проявляється	Приклади
Синтаксичні дефекти (порушення граматики мови)	На етапах компіляції та запуску коду	Відсутні чи неправильно розміщені дужки, оператори, помилки в назвах ключових слів або змінних
Семантичні дефекти (помилки логіки)	Під час виконання програма не відповідає передбаченій поведінці (висока складність виявлення)	Використання змінних, структур даних, алгоритмів не за призначенням
Вразливості	Семантичні дефекти, які загрожують безпеці, цілісності або доступності системи чи її даних	Переповнення буфера, переповнення цілих чисел, міжсайтовий скриптинг, “завислі” вказівники

Вибір стилю форматування коду часто є суб'єктивним питанням, яке може викликати дискусії між розробниками. Через це, інструменти форматування повинні підтримувати досить широкі можливості налаштувань, щоб бути здатними задовольняти потреби різних команд розробки. Робота з системами на основі правил часто потребує чимало зусиль. Оскільки кожна специфікація передбачає форматування для конкретного стилю, кожен новий стиль вимагає складних модифікацій або створення нових наборів правил. Навіть незначні зміни у граматиці можуть потребувати адаптації правил форматування. При цьому наявність величезної кількості правил у таких системах є досить типовою ситуацією на практиці.

Механізми форматування, засновані на машинному навчанні, зазвичай працюють у два етапи: навчання та форматування. Навчання моделі відбувається безпосередньо на базі корпусу

коду із бажаним форматуванням. В результаті даного етапу, формується статистична модель, яка представляє стиль форматування виділених із навчальних даних. На етапі форматування модель будує прогнози щодо того, як автор корпусу відформатував би певний контекст, повертаючи директиву форматування [13].

Виявлення “коду, що погано пахне” (англ. code smell detection). В процесі підтримки та покращення програмного забезпечення, розробники повинні регулярно вносити в нього зміни, щоб відповідати новим вимогам, поліпшувати функціональність та виправляти виявлені дефекти. Проте, обмежений час та інші фактори можуть сильно ускладнювати можливість підтримки якісної архітектури та знаходження оптимальних рішень. В результаті цього, часто порушується початковий дизайн системи та виникає так званий “код, що (погано) пахне”. “Запах” коду був визначений М. Фаулером та К. Беком [14] як ознаки, що вказують на неоптимальний дизайн або варіанти реалізації у вихідному коді. Вони представляють одну із найсерйозніших форм технічної заборгованості (англ. *tech debt*) й не лише знижують загальну зрозумілість коду, а й роблять його більш схильним до появи помилок. Таким чином, “код, що пахне” є усталеною характеристикою вихідного коду, що вказує на потенційну наявність більших проблем. Їх усунення здійснюється за допомогою рефакторингу [15; 16; 17].

Задача виявлення “коду, що пахне” є дуже подібною до задачі виявлення дефектів, а тому, аналогічно до неї, має схожий набір викликів, пов'язаний із використанням інструментів виявлення на основі правил. Перш за все, важливим питанням є коректність налаштування правил. Розробники можуть мати різні погляди на допустимість певних властивостей у коді, тому метрики, що використовуються для виявлення запахів коду різними інструментами, можуть суттєво різнитися. Крім того, навіть якщо метрики однакові, їх порогові значення можуть змінюватися. Це може призводити до збільшення або зменшення кількості виявлених “запахів”. Окрім цього, не менш важливою проблемою є точність результатів. Як і при виявленні дефектів, системи, що базуються на правилах, часто схильні до хибно позитивних результатів при виявленні “запахів” коду. З усіх цих причин, нещодавньою тенденцією є поява нових підходів до виявлення запахів коду на основі методів машинного навчання. Основною перевагою таких підходів є можливість навчати модель використовуючи достатньо велику кількість даних, наявних в тому ж, або в інших подібних проектах, замість того, щоб опиратись на попередньо визначені порогові значення метрик [15].

На рис. 1 представлено загальну структуру категоризації підходів до застосування машинного навчання в статичному аналізі коду.

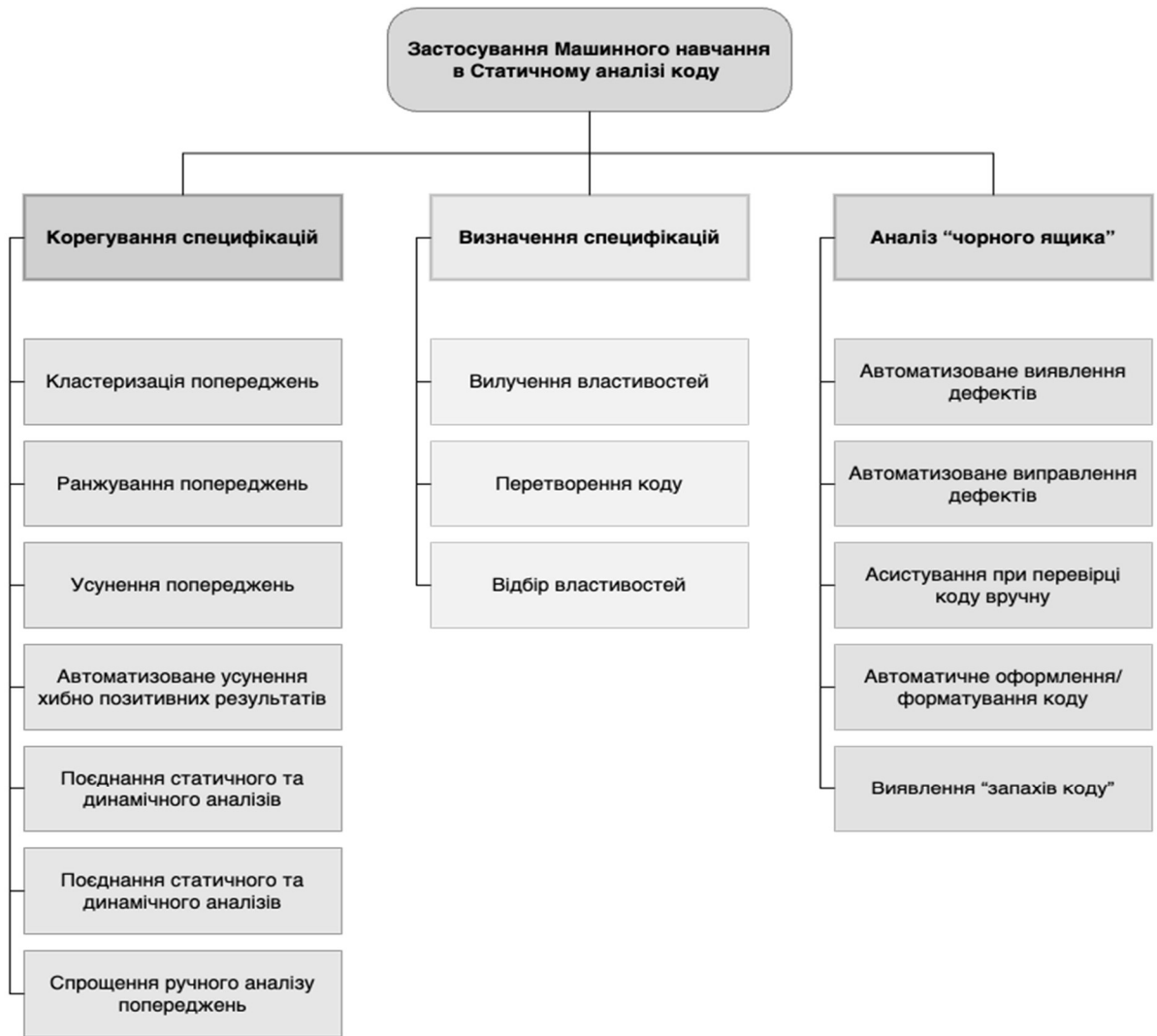


Рис. 1. Підходи до застосування машинного навчання в статичному аналізі

Враховуючи всі особливості розглянутих підходів до використання нейронних мереж в статичному аналізі коду, можна сказати, що *коригування специфікацій* та *аналіз "чорного ящика"* виглядають як більш перспективні рішення з точки зору практичного застосування. Підходи із категорії *визначення специфікацій*, на поточний момент, не виглядають як ті, що можуть запропонувати певні унікальні переваги, порівняно з іншими. При цьому, вони не уникають більшості викликів та ускладнень, пов'язаних із навчанням нейронних мереж безпосередньо на вихідному коді та не надто спрощують складність побудови статичних аналізаторів з огляду на необхідність конфігурації наборів правил. Звісно, успішні реалізації підходів з цієї категорії потенційно можливі, проте порядок застосування нейронних мереж та класичних алгоритмів статичного аналізу в них не виглядає оптимальним, що також підтверджується досить малою кількістю дослідницьких робіт у цьому напрямку.

Як *коригування специфікацій*, так і *аналіз "чорного ящика"* мають власні переваги та недоліки. В той час, як перша категорія орієнтована на піднятті точності результатів зі збереженням надійності класичних підходів, друга – спрямована на отримання більшої гнучкості комплексного аналізу, допускаючи при цьому зниження рівня зрозумілості та надійності. Завдяки тому, що підходи з категорії *коригування специфікацій* орієнтуються на вдосконаленні статичного аналізу з використанням наявних аналізаторів на базі правил, вони мають досить хороший потенціал для отримання нових інструментів аналізу, які матимуть вищу точність, не поступаючись при цьому їм у надійності. Особлива увага у подальших дослідженнях має приділятися застосуванню ансамблевого навчання (англ. *ensemble learning*) на базі кількох статичних аналізаторів. У такій формі, ансамблеве навчання передбачає поєднання результатів кількох інструментів статичного аналізу та використання моделі нейронної мережі для

постобробки і порівняння результатів кожного із них. Такий підхід, в перспективі, повинен покращувати точність аналізу завдяки “взаємному” підтвердженню аналізаторами повідомлень про потенційні дефекти, знижуючи таким чином вплив помилок окремих інструментів.

Висновки

Активна розробка програмного забезпечення – це неперервні зміни програмного коду. Для підвищення його цінності, розробникам необхідно вносити новий чи покращувати наявний функціонал, виправляти виявлені помилки та вдосконалювати архітектуру рішень. Усе це несе потенційну небезпеку появи нових дефектів у програмній системі. При роботі з великою кодовою базою майже неможливо знайти та усунути всі помилки, проте критично важливо зводити їх присутність до певного прийнятного мінімуму. Головне питання, яке виникає під час усунення несправностей вихідного коду, полягає в тому, як зробити це найбільш ефективно.

Враховуючи те, що з кожним наступним етапом розробки витрати на внесення виправлень зростають, логічним висновком є те, що основні зусилля для виявлення дефектів повинні прикладатись якомога раніше. Саме тому статичний аналіз коду є одним із найефективніших інструментів покращення процесу виявлення несправностей, оскільки здатен працювати вже на етапі появи перших зразків коду.

Класичні підходи статичного аналізу базуються на алгоритмах зіставлення шаблонів. Це означає, що вони сканують вихідний код, намагаючись виявити помилкові шаблони, порівнюючи їх із тими, що присутні в їхній базі даних. Недоліком такого підходу є його нездатність навчатися та адаптувати свою поведінку, коли виявлення не вдається, що може призводити до високого рівня хибно позитивних результатів і, своєю чергою, знецінювати корисність інструменту для розробників.

Використання таких підходів машинного навчання, як нейронні мережі, має хороші перспективи підвищення гнучкості систем та розширення їх можливостей. Підтримка методів машинного навчання в таких системах має забезпечувати їм здатність навчатися на попередніх виявленнях та зменшувати кількість хибно позитивних результатів, отримуючи високий рівень істинно позитивних результатів. При цьому, важливо відзначити, що використання нейронних мереж у статичному аналізі далеко не завжди означає відмову від класичних підходів, які добре зарекомендували себе у багатьох задачах. Відповідно до проведеного дослідження, методи машинного навчання можуть як повністю замінити традиційні підходи (*аналіз “чорного ящика”*), так і використовуватись у комбінації з ними, для забезпечення більшої точності та ефективності (*коригування та визначення специфікацій*). У будь-якому випадку, побудова більш досконалих систем статичного аналізу змушує шукати нові рішення в області поєднання гнучких за природою нейронних мереж та детермінованих та надійних класичних алгоритмів.

Література

1. Тессі, Г. (2002). Економічні наслідки неадекватної інфраструктури для тестування програмного забезпечення. Технічний звіт. NIST, 309 стор.
2. Муске Т., Серебреник А. (2022). Огляд підходів до постобробки сповіщень статичного аналізу. ACM Computing Surveys (CSUR), № 55, стор. 1 - 39, 10.1145/3494521
3. Алламніс, М. (2022). Графічні нейронні мережі в програмному аналізі. Wu, L., Cui, P., Pei, J., Zhao, L. (eds) Graph Neural Networks: Foundations, Frontiers, and Applications. Springer, Сінгапур, стор. 483 - 497. 10.1007/978-981-16-6054-2_22
4. Муске Т. Б. (2020). Постобробка тривог статичного аналізу. [Phd Thesis 1 (Research TU/e / Graduate TU/e), Mathematics and Computer Science]. Технологічний університет Ейндховена, 174 с.
5. Еррамредді, С., Мордал, А., Кок, У., Вей, С., Фостер, Дж., Карпуат, М., Портер, А. (2023). Емпірична оцінка підходів машинного навчання для сортування звітів інструментів статичного аналізу. Емпірична інженерія програмного забезпечення, вип. 28. 10.1007/s10664-022-10253-z
6. Мар'янов Т., Пашенко І., Массаччі Ф. (2022). Машинне навчання для виявлення вразливостей у вихідному коді: що працює, а чого ще немає. IEEE Security & Privacy, том. 20, № 5, С. 60-76. 10.1109/MSEC.2022.3176058
7. Чжан, З., Лу, С., Хуан, З., Чжао, З. (2022). ASGNN: графічні нейронні мережі з адаптивною структурою. ArXiv, 10.48550/arXiv.2210.01002
8. ChatGPT. OpenAI. Отримано з: <https://openai.com/blog/chatgpt>
9. Лал Х., Пахва Г. (2017). Аналіз аналізу коду програмної системи з використанням методів машинного навчання. 11-та Міжнародна конференція з інтелектуальних систем і управління (ISCO), Коїмбатор, Індія, стор. 8 - 13. 10.1109/ISCO.2017.7855962
10. Мадера, М., і Томон, Р. (2017). Випадкове дослідження моделі машинного навчання для експертної системи перевірки коду в програмній інженерії. Федеративна конференція з комп'ютерних наук та інформаційних систем (FedCSIS), стор. 1357-1363.

11. Міара Р., Масселман Дж., Наварро Дж., Шнейдерман Б. (1983). Програмний відступ і зрозумілість. *Комун. ACM*, № 11, вип. 26, стор. 861-867. 10.1145/182.358437
12. Рейс, С. (2007). Автоматична стилізація коду. Матеріали 22-ї міжнародної конференції IEEE/ACM з автоматизованої розробки програмного забезпечення, 10.1145/1321631.1321645
13. Парр Т., Вінджу Дж. (2016). На шляху до універсального форматувальника коду через машинне навчання. Матеріали міжнародної конференції ACM SIGPLAN 2016 року з розробки мови програмного забезпечення, 10.1145/2997364.2997383
14. Фаулер М., Бек К., Брант Дж., Опдайк В., Робертс Д. (1999). *Refactoring: Improving the Design of Existing Code*, 1st ed. Addison-Wesley Professional, 464 стор. ISBN 978-0201485677
15. Фонтана, Ф., Ментюля, М., Заноні, М., Маріно, А. (2016). Порівняння та експериментування методів машинного навчання для виявлення запаху коду. *Емпірична інженерія програмного забезпечення*, № 21, стор. 1143-1191. 10.1007/s10664-015-9378-4
16. Азім, М., Паломба, Ф., Ши, Л., Ван, К. (2019). Методи машинного навчання для виявлення запаху коду: систематичний огляд літератури та мета-аналіз. *Інф. програмне забезпечення техн.*, С. 115-138. 10.1016/J.INFSOF.2018.12.009
17. Деванган С., Рао Р., Мішра А., Гупта М. (2022). Виявлення запаху коду за допомогою алгоритмів ансамблевого машинного навчання. *Прикладні науки*, 12(20), 10.3390/app122010321

References

1. Tasey, G. (2002). The economic impacts of inadequate infrastructure for software testing. *Technical Report. NIST*, 309 p.
2. Muske, T., Serebrenik, A. (2022). Survey of Approaches for Postprocessing of Static Analysis Alarms. *ACM Computing Surveys (CSUR)*, No 55, pp. 1 - 39, 10.1145/3494521
3. Allamanis, M. (2022). Graph Neural Networks in Program Analysis. *Wu, L., Cui, P., Pei, J., Zhao, L. (eds) Graph Neural Networks: Foundations, Frontiers, and Applications*. Springer, Singapore, pp. 483 - 497. 10.1007/978-981-16-6054-2_22
4. Muske, T. B. (2020). Postprocessing of static analysis alarms. *[Phd Thesis 1 (Research TU/e / Graduation TU/e), Mathematics and Computer Science]*. Eindhoven University of Technology, 174 p.
5. Yerramreddy, S., Mordahl, A., Koc, U., Wei, S., Foster, J., Carpuat, M., Porter, A. (2023). An empirical assessment of machine learning approaches for triaging reports of static analysis tools. *Empirical Software Engineering*, vol. 28. 10.1007/s10664-022-10253-z
6. Marjanov, T., Pashchenko, I., Massacci, F. (2022). Machine Learning for Source Code Vulnerability Detection: What Works and What Isn't There Yet. *IEEE Security & Privacy*, vol. 20, No. 5, pp. 60-76. 10.1109/MSEC.2022.3176058
7. Zhang, Z., Lu, S., Huang, Z., Zhao, Z. (2022). ASGNN: Graph Neural Networks with Adaptive Structure. *ArXiv*, 10.48550/arXiv.2210.01002
8. ChatGPT. OpenAI. Retrieved from: <https://openai.com/blog/chatgpt>
9. Lal, H., Pahwa, G. (2017). Code review analysis of software system using machine learning techniques. *11th International Conference on Intelligent Systems and Control (ISCO)*, Coimbatore, India, pp. 8 - 13. 10.1109/ISCO.2017.7855962
10. Madera, M., & Tomon, R. (2017). A case study on machine learning model for code review expert system in software engineering. *Federated Conference on Computer Science and Information Systems (FedCSIS)*, pp. 1357-1363.
11. Miara, R., Musselman, J., Navarro, J., Shneiderman, B. (1983). Program indentation and comprehensibility. *Commun. ACM*, No. 11, vol. 26, pp. 861-867. 10.1145/182.358437
12. Reiss, S. (2007). Automatic code stylizing. *Proceedings of the 22nd IEEE/ACM International Conference on Automated Software Engineering*, 10.1145/1321631.1321645
13. Parr, T., Vinju, J. (2016). Towards a universal code formatter through machine learning. *Proceedings of the 2016 ACM SIGPLAN International Conference on Software Language Engineering*, 10.1145/2997364.2997383
14. Fowler, M., Beck, K., Brant, J., Opdyke, W., Roberts, D. (1999). *Refactoring: Improving the Design of Existing Code*, 1st ed. Addison-Wesley Professional, 464 p. ISBN 978-0201485677
15. Fontana, F., Mäntylä, M., Zanoni, M., Marino, A. (2016). Comparing and experimenting machine learning techniques for code smell detection. *Empirical Software Engineering*, No. 21, pp. 1143-1191. 10.1007/s10664-015-9378-4
16. Azeem, M., Palomba, F., Shi, L., Wang, Q. (2019). Machine learning techniques for code smell detection: A systematic literature review and meta-analysis. *Inf. Softw. Technol.*, pp. 115-138. 10.1016/J.INFSOF.2018.12.009
17. Dewangan, S., Rao, R., Mishra, A., Gupta, M. (2022). Code Smell Detection Using Ensemble Machine Learning Algorithms. *Applied Sciences*, 12(20), 10.3390/app122010321

УДК 004.056.5: 004.056.2

ЗАХИСТ НЕСТРУКТУРОВАНОЇ ІНФОРМАЦІЇ НА МОБІЛЬНОМУ ПРИСТРОЮ

DOI 10.36994 / 2788-5518-2023-01-05-21

Бровченко Є. М., Відкритий міжнародний університет розвитку людини «Україна», Інститут комп'ютерних технологій, Київ, Україна. <https://orcid.org/0000-0002-1416-0385>, yebrovchenko@hotmail.com

Самарай В. П., к.т.н., проф., Відкритий міжнародний університет розвитку людини «Україна», Інститут комп'ютерних технологій, Київ, Україна. <https://orcid.org/0000-0003-4419-1366>, samaraj@ukr.net

Даценко І. П., к.т.н., доц., Відкритий міжнародний університет розвитку людини «Україна», Інститут комп'ютерних технологій, Київ, Україна. <https://orcid.org/0000-0002-0047-413X>, docik_ivan@ukr.net

Павленко В. І., к.ф.-м.н., доц., Відкритий міжнародний університет розвитку людини «Україна», Інститут комп'ютерних технологій, Київ, Україна. <https://orcid.org/0000-0002-3958-0415>, pavlenko.v@i.ua

Серєда А.В., Відкритий міжнародний університет розвитку людини «Україна», Інститут комп'ютерних технологій, Київ, Україна. <https://orcid.org/0000-0002-4779-5475>, andrew.sereda@gmail.com

Анотація: У статті розглянуто проблеми використання мобільного пристрою в умовах активного використання та захист неструктурованої інформації. Основна увага приділена використанню мобільного пристрою як частини інформаційної системи та можливі заходи щодо захисту. Розвиток та динаміка використання сучасних інформаційних технологій у багатьох сферах життя вимагає від інженерів та розробників нових рішень, необхідно закласти міцний фундамент для подальшого розвитку та можливість ефективного удосконалення чи підтримки. Мобільний пристрій займає своє важливе місце у житті сучасної людини. Тому треба розглядати мобільний пристрій як не від'ємну складову сучасних інформаційних систем. Була проведена робота по вивченню проблеми в реальних умовах використання та взаємодії. Треба зазначити що питання безпеки є комплексним: починаючи від етапу створення та контроль на етапі експлуатації, також дуже важливим є інформаційна обізнаність та відповідальність кінцевого користувача. Кібербезпека виходить на принципово новий рівень. Існують певні готові інженерно-архітектурні рішення що допомагають інтегрувати мобільний пристрій в інформаційну систему, але ці рішення інколи не виконують поставлений завдань. Аналіз прикладної області та результати дослідження говорять про те що захист інформації в цілому є важливим та є попит на ефективні рішення. Не слід забувати також про економічну складову. Вироблено ряд рекомендацій щодо можливих заходів по безпеці.

Ключові слова: мобільний пристрій, захист інформації, неструктурована інформація, кібербезпека

PROTECTION OF UNSTRUCTURED INFORMATION ON A MOBILE DEVICE

Evgen Brovchenko, Open International University of Human Development "Ukraine", Kyiv, Ukraine. <https://orcid.org/0000-0002-1416-0385>, yebrovchenko@hotmail.com

Valery Samarai, Ph.D., Ass. Prof., Open International University of Human Development "Ukraine", Kyiv, Ukraine. <https://orcid.org/0000-0003-4419-1366>, samaraj@ukr.net

Ivan Datsenko, Ph.D., Open University of Human Development "Ukraine", Kyiv, Ukraine. <https://orcid.org/0000-0002-0047-413X>, docik_ivan@ukr.net

Volodymyr Pavlenko, Ph.D., Ass. Prof., Open University of Human Development "Ukraine", Kyiv, Ukraine. <https://orcid.org/0000-0002-3958-0415>, pavlenko.v@i.ua

Andrew Sereda, Open University of Human Development "Ukraine", Kyiv, Ukraine. <https://orcid.org/0000-0002-4779-5475>, andrew.sereda@gmail.com

Annotation: Information technology (IT) and computer systems are becoming an integral part of modern man's life, irreplaceable assistants. The development of IT happens at lightning speed and has a great impact on our society, every day we could observe new technical and engineering inventions. Moreover, it can be both a certain integration into the work process - automation of the daily routine, and part of leisure - communication and entertainment, etc. The entire complex of technologies and systems improve and facilitate the performance of everyday tasks, open new opportunities and remove borders. It is difficult to imagine a modern person without a

mobile phone or tablet, which has almost continuous and unlimited access to the Internet or the worldwide global network. Today, a mobile device has a certain sacred meaning, it can contain a lot of different information or data. The nature and importance of this information may vary, but protection is a priority. And considering the developed dynamics and the uncertainty of some processes, engineers and developers face a difficult and important task to protect information systems and ensure reliable operation. To some extent, the mobile device can be considered the most problematic component. This device has a few advantages and features - this is exactly what makes it vulnerable to potential attackers and user actions can cause a few problems. This article is devoted to the problems of using mobile devices and protecting unstructured information. Protecting unstructured information on a mobile device is an important task from a cybersecurity perspective. Unstructured information can be in the form of text documents, emails, photos, videos, voice recordings, and other content stored on mobile devices. The main problems include loss or theft of a mobile device; operating system and software vulnerabilities; insufficient level of encryption and insufficient level of complexity of passwords; neglect of cyber hygiene. Existing engineering and architectural solutions do not always fulfill the task and may have problems with integration. Do not forget about the economic component. The analysis of the application area and the research results indicate that information protection in general is important and there is a demand for quick and effective solutions regarding the security of information systems in general and mobile devices. Only an integrated approach can provide an effective solution.

The future of protecting unstructured information on mobile devices will be based on a combination of technological innovations and increased user awareness of cybersecurity issues. Some areas of development include the use of artificial intelligence (AI) and machine learning: AI can help develop advanced defense mechanisms that automatically detect anomalous or suspicious activity and block potential threats; biometric authentication, improvements in biometric authentication technologies, such as facial recognition, fingerprints, can provide more reliable protection of unstructured information; application of blockchain technologies: blockchain can play an important role in ensuring information security, as it allows data to be stored distributed and protected from changes; development of quantum-resistant encryption algorithms: with the development of quantum computers, there is a need for new encryption algorithms that can withstand the power of quantum hacking; privacy by design (Privacy by Design): designing mobile devices and applications with user privacy in mind will help protect unstructured information on mobile devices; user education and awareness: increasing users' awareness of cyber security and teaching them effective methods of protecting information on mobile devices will help reduce the risk of security incidents; development of the legislation and regulatory framework, improvement of the legislation and regulatory framework in the field of cyber security will contribute to the creation of a reliable environment for the protection of unstructured information. Securing unstructured information on mobile devices will rely on a combination of technological advances, user education, and collaboration between various parties such as developers, service providers, and lawmakers. Ensuring data security and privacy in the digital world remains a priority as the number of mobile devices and their use continues to grow.

The modern user of mobile devices has certain requirements related to ergonomics and the level of information awareness. The development of the industry requires fast, reliable, and effective solutions that fall not only in the economic plane, but also in the plane of cyber security. The use of mobile devices can adapt to the fast-paced modern security needs of an organization and can be easily configured to meet specific requirements. A mobile device is convenient and affordable for many users, it is usually close to the user and can be used as an additional security element.

Key words: *mobile device, information protection, unstructured information, cyber security*

Вступ

Інформаційні технології (ІТ) та обчислювальні системи стали невід'ємною частиною життя сучасної людини, незамінними помічниками. Розвиток ІТ відбувається так швидко, що майже кожен день ми маємо можливість спостерігати нові технічні та інженерні винаходи. Причому це може бути як певна інтеграція у робочий процес – автоматизація повсякденної рутини, так і частина дозвілля – спілкування та розваги, тощо. Весь комплекс технологій та систем поліпшують та спрощують виконання повсякденних задач, відкривають нові можливості та прибирають кордони. Важко уявити собі сучасну людину без мобільного телефону чи планшета, що має майже безперервний та безлімітний доступ до Інтернету чи всесвітньо глобальної мережі. Сьогодні мобільний пристрій має певний сакральний сенс, він може містити багато різноманітної інформації чи даних. Природа та важливість цієї інформації може бути різною, але захист має пріоритетне значення. А враховуючи динаміку розвитку та невизначеність деяких процесів, перед інженерами та розробниками стоїть важке та важливе завдання по захисту інформаційних систем та забезпеченню надійної роботи. До певної міри мобільний пристрій можна вважати найбільш проблематичною складовою. Цей пристрій має ряд переваг та особливостей – власне це й роблять його уразливим перед потенційними зловмисниками та й дії користувача можуть

спричинити ряд проблем. Данна стаття присвячена проблематиці використання мобільних пристроїв та захисту неструктурованої інформації. Захист неструктурованої інформації на мобільному пристрої є важливим завданням з точки зору кібербезпеки. Неструктурована інформація може бути у вигляді текстових документів, електронних листів, фотографій, відео, записів голосу та іншого контенту, який зберігається на мобільних пристроях. До основних проблем можна віднести: втрата або крадіжка мобільного пристрою; вразливості операційної системи та програмного забезпечення; недостатній рівень шифрування та недостатній рівень складності паролів; нехтування кібергігієною. Існуючі інженерно-архітектурні рішення не завжди виконують поставлену задачу та можуть мати проблеми з інтеграцією. Не слід забувати й про економічну складову. Аналіз прикладної області та результати дослідження говорять про те що захист інформації в цілому є важливим та є попит на швидкі та ефективні рішення щодо безпеки інформаційних систем в цілому та мобільного пристрою зокрема. Лише комплексний підхід може забезпечити ефективне рішення.

Виконання досліджень

Як відомо обчислювальна система - це сукупність апаратних і програмних компонентів, які взаємодіють для вирішення обчислювальних завдань. Обчислювальні системи можуть включати різноманітні пристрої, такі як комп'ютери, мобільні телефони, планшети, сервери, ноутбуки, суперкомп'ютери. Слід розуміти що перелік цих складових може змінюватись, причому деякі пристрої можуть зникати, а про деякі ми сьогодні навіть і не здогадуємось. Система складається з апаратного та програмного забезпечення. Апаратна частина - це фізичні компоненти комп'ютерної системи, які виконують обчислювальні операції, операції по збереженню та керуванню. Програмне забезпечення - це набір програм, що контролюють та керують роботою апаратної частини комп'ютера. Воно може включати в себе операційні системи, драйвери пристроїв, програми-редактори, веб-браузери, мультимедійні прошивки та інші програми. Обчислювальні системи можуть бути використані для різноманітних завдань, таких як обробка даних, створення документів, візуалізація інформації, аналіз даних, ігри та інше. Вони можуть бути використані в різних галузях, таких як наука, бізнес, медицина, техніка, транспорт, тощо. Рівень інтеграції обчислювальних систем у повсякденне життя з часом буде лише збільшуватись та набувати нових рис.

Отже що таке мобільний пристрій? Мобільний пристрій - це портативний електронний пристрій, який має можливість підключення до мережі зв'язку та доступу до різноманітних сервісів та додатків. У сучасному світі мобільні пристрої стали невід'ємною частиною життя людей, і це змінило спосіб, яким ми спілкуємося, взаємодіємо, працюємо та розважаємося. Однією з основних функцій мобільного пристрою є забезпечення зв'язку, що дозволяє користувачам здійснювати дзвінки та відеодзвінки, надсилати та отримувати повідомлення, електронну пошту та інші форми комунікації. Мобільні пристрої також забезпечують доступ до Інтернету, що дає користувачам можливість шукати інформацію, переглядати відео, грати в ігри, слухати музику та виконувати інші дії. Мобільні пристрої також стали необхідним інструментом для роботи та навчання, дозволяючи користувачам працювати з електронною поштою, текстовими документами, електронними таблицями та іншими програмами. Вони можуть бути використані для спостереження за здоров'ям та фітнесом, ведення щоденника, планування та організації розкладу та інших аспектів життя. Однією з основних переваг мобільних пристроїв є їх портативність та доступність, що дозволяє людям бути зв'язаними та отримувати інформацію в будь-якому місці та в будь-який час. У цілому, мобільні пристрої стали невід'ємною частиною життя сучасної людини, і вони продовжують еволюціонувати та вдосконалюватися. Завдяки їхній функціональності та зручності використання, мобільні пристрої дозволяють людям бути більш продуктивними та організованими в їхньому житті та роботі.

Можна відділити такі основні типи мобільних пристроїв:

- Смартфоне - найпоширенішим типом мобільних пристроїв. Вони поєднують в собі функціональність мобільного телефону і комп'ютера, забезпечуючи широкий спектр можливостей, таких як здійснення дзвінків, відправлення повідомлень, серфінг в Інтернеті, встановлення додатків тощо.

- Планшет - це портативні пристрої з екраном сенсорного керування, більшими, ніж у смартфонів, зазвичай від 7 до 12 дюймів. Вони надають можливість виконувати багато завдань, подібних до комп'ютера, таких як перегляд веб-сторінок, перегляд відео, читання електронних книг, встановлення додатків і т. д.

▪ Ноутбук – персональний комп'ютер зі складним корпусом, що містить клавіатуру та дисплей. Вони надають широкі можливості, які забезпечують роботу з текстом, виконання завдань, перегляд мультимедійного контенту, доступ до Інтернету.

▪ Ультрабук - це тонкий, легкий і потужний ноутбук з довгим часом автономної роботи. Вони надають високу продуктивність і можливості, подібні до ноутбуків, але з меншими розмірами та вагою.

▪ Пристрій що носить (wearables) – різноманітні гаджети, які можуть бути одягнуті на тіло, наприклад, смарт-годинники, фітнес-браслет.

Існують основні операційні системи: Android, iOS, Windows/Mobile, BlackBerry, KaiOS. Різноманіття мобільних пристроїв з одного боку дає можливість обрати саме те що треба користувачу, з іншого боку ускладнює роботу проєктувальників та розробників.

Особливу увагу слід приділити Android Enterprise. Отже Android Enterprise (раніше відомий як Android for Work) - це програма та набір функцій, розроблених компанією Google для підприємств, які використовують мобільні пристрої на базі операційної системи Android. Android Enterprise надає розширені можливості керування і захисту даних, що дозволяє організаціям легко і безпечно інтегрувати та керувати пристроями Android в корпоративному середовищі.

Деякі ключові функції та можливості Android Enterprise включають:

1. Робочі профілі: дозволяє розділити особисті дані і програми від корпоративних на пристрої з Android. Це дозволяє зберігати особисті дані окремо від корпоративних даних і додатків.

2. Управління пристроями: забезпечує централізоване управління пристроями, що дозволяє адміністраторам віддалено керувати налаштуваннями, політиками безпеки, встановлювати додатки та проводити оновлення.

3. Політики безпеки: дозволяє встановлювати політики безпеки на пристроях, такі як вимоги до паролів, шифрування даних, обмеження функцій пристрою та інші заходи безпеки.

4. Магазин додатків для підприємств: надає можливість створювати власний магазин додатків для підприємств, де адміністратори можуть контролювати доступні додатки та встановлювати їх на пристрої.

5. Інтеграція з обліковими записами Google Workspace: компанії, які використовують Google Workspace (раніше G Suite), можуть легко інтегрувати свої корпоративні облікові записи з Android Enterprise і керувати доступом до облікових записів та сервісів.

6. Захист даних: можна використовувати різні методи захисту даних, такі як шифрування пристроїв, вимоги до паролів, обмеження прав доступу та інші політики безпеки, щоб забезпечити конфіденційність і цілісність даних.

7. Інтеграція з іншими системами: може інтегруватися з різними корпоративними системами, такими як системи керування ідентифікацією (IAM), системи електронної пошти та календаря, хмарні сервіси для зберігання даних та багато іншого.

8. Завдання та автоматизація: створювання різних завдань та автоматизування процесів на пристроях, що дозволяє підвищити продуктивність та ефективність користувачів.

Неструктурована інформація - це інформація, яка не зберігається у структурованому форматі бази даних або в іншому структурованому форматі. Наприклад, це можуть бути тексти, документи, електронні листи, фотографії, відео та інші файли, які зберігаються на різних пристроях, таких як комп'ютери, мобільні телефони, планшети та інші пристрої з можливістю зберігання даних. Неструктурована інформація зазвичай не має чіткої структури, тому вона може бути складнішою для організації та обробки, порівняно з інформацією, яка зберігається у структурованих форматах. Проте, неструктурована інформація містить в собі багато важливих даних, таких як особисті інформаційні дані, комерційні та бізнес-документи, що можуть мати значення для користувачів та організацій, тому захист неструктурованої інформації є дуже важливою та пріоритетною задачею.

Характеристики неструктурованої інформації:

▪ Текстовий контент: мобільний пристрій містить значну кількість текстової інформації, яка може бути неструктурованою. Це можуть бути нотатки, записи, електронні книги, блоги, коментарі, соціальні медіа, статті та інше. Цей текст може бути у різних форматах, включаючи неперервний потік тексту, списки, маркери, таблиці або навіть скановані документи.

▪ Мультимедійний контент: мобільний пристрій дозволяють зберігати і отримувати доступ до різноманітного мультимедійного контенту, який також може бути неструктурованим. Це

включає фотографії, відео, аудіозаписи, графіку, анімацію та інші типи медіа. Цей контент може бути в неструктурованій формі без додаткової метаданих або організації.

- Соціальні медіа та комунікація: мобільний пристрій широко використовуються для соціальних мереж, месенджерів та інших комунікаційних платформ. Інформація, що обмінюється через ці канали, може бути неструктурованою, включаючи повідомлення, коментарі, пости, фотографії, відео та інші медіа.

- Веб-сторінки та блоги: мобільний пристрій дозволяють переглядати веб-сторінки та блоги. Ця інформація також може бути неструктурованою, залежно від оформлення сторінки, наявності текстових блоків, заголовків, списків, посилань тощо.

- Документи і файлова система: мобільний пристрій можуть містити різні документи у форматах, таких як PDF.

Неструктурована інформація на мобільних пристроях є широко поширеним явищем і включає текстовий контент, мультимедійний контент, соціальні медіа, веб-сторінки та блоги, а також документи і файлову систему. Ця інформація може бути в різних форматах і організаціях, що може створювати виклики в управлінні, пошуку та використанні її. Управління неструктурованою інформацією на мобільних пристроях є важливим завданням. Користувачі повинні мати ефективні інструменти для організації, пошуку і доступу до цієї інформації. Технології обробки мови, комп'ютерного зору, машинного навчання та інші інновації в галузі штучного інтелекту можуть впроваджуватися для автоматизації розпізнавання та організації неструктурованої інформації. Досягнення в області розпізнавання тексту, обробки мови та аналізу великих обсягів даних роблять неструктуровану інформацію більш доступною та корисною для користувачів мобільних пристроїв. Продовжується розвиток інструментів та технологій, що допомагають в організації і використанні цієї інформації на мобільних пристроях. Загалом, неструктурована інформація на мобільних пристроях є важливим аспектом сучасного інформаційного оточення. Розвиток і використання інструментів для управління цією інформацією відкриває нові можливості для користувачів, забезпечуючи зручний доступ до ресурсів.

Для захисту мобільно пристрою в цілому та інформації що зберігається на ньому зокрема досить ефективною є використання біометричної ідентифікації. Біометрія - це використання фізичних або поведінкових характеристик людини для ідентифікації, аутентифікації та авторизації. На сучасних мобільних пристроях все більше використовуються біометричні методи для захисту неструктурованої інформації, що зберігається на них. Використання біометрії дозволяє підвищити безпеку і зручність доступу до цієї інформації.

Технології біометричної ідентифікації на мобільних пристроях:

- Сканування відбитків пальців: сканер вбудований в сенсорний екран або на задній панелі пристрою. Відбиток пальця порівнюється з збереженими біометричними даними, що дозволяє впізнати користувача.

- Розпізнавання обличчя: камера пристрою використовується для захоплення обличчя користувача, яке порівнюється з попередньо збереженими біометричними даними.

- Розпізнавання відбитків очей: сітківка може бути використані для ідентифікації користувача. Спеціальні камери або сенсори захоплюють деталі очей і порівнюють їх з шаблонами біометричних даних.

- Розпізнавання голосу: вбудований мікрофон записує голос користувача, який порівнюється з попередньо збереженими голосовими шаблонами.

Переваги використання біометрії на мобільних пристроях:

- Вищий рівень безпеки. Біометричні дані є унікальними для кожної особи і майже неможливо підробити або підмінити. Це дозволяє забезпечити високий рівень ідентифікації та аутентифікації користувача.

- Зручність використання. Використання біометрії на мобільних пристроях забезпечує зручний та швидкий доступ до інформації.

- Ізольованість зберігання. Біометричні дані зазвичай зберігаються в зашифрованому вигляді на самому пристрої. Це зменшує ризик їх витоку або несанкціонованого доступу до них.

- Можливості багатофакторної аутентифікації. Біометричні методи можуть бути поєднані з іншими методами аутентифікації, такими як пін-коди або паролі, для створення більш надійного і багатоетапного процесу ідентифікації.

- **Необхідність фізичного присутності.** Використання біометрії вимагає фізичної присутності особи для ідентифікації. Це ускладнює можливість несанкціонованого доступу до пристрою або даних.

Наступним кроком можна вважати захист інформації на рівні операційної системи чи мобільно додатку. MAM (Mobile Application Management) є скороченням від англійського терміну "Mobile Application Management". Він відноситься до підходу до управління мобільними додатками в організації. MAM використовується для керування та контролю мобільними додатками, які встановлюються на пристрої користувача. Це може бути застосунок для внутрішнього використання в організації або публічний додаток, який користувачі завантажують з магазинів додатків. Основна мета MAM полягає в управлінні додатками на рівні самого додатку, незалежно від пристрою. Вона надає можливості забезпечити безпеку, контроль та конфігурацію додатків, а також здійснювати моніторинг, оновлення та віддалене управління додатками на мобільних пристроях. Деякі можливості MAM можуть включати: можливість встановлення та виконання політик, що стосуються використання додатків, таких як обмеження функцій або доступу до даних; здатність забезпечувати безпеку додатків шляхом використання шифрування, контролю доступу та ідентифікації користувачів; можливість розповсюдження додатків на пристрої користувачів через магазини додатків або корпоративні канали, а також віддалене оновлення додатків; здатність збирати дані про використання додатків, продуктивність та поведінку користувачів для аналізу та прийняття відповідних рішень. MDM (Mobile Device Management) є скороченням від англійського терміну "Mobile Device Management". Це підходу до управління мобільними пристроями в організації. MDM використовується для централізованого керування, контролю та захисту мобільних пристроїв, таких як смартфони та планшети, що використовуються в організації. Він надає можливості забезпечення безпеки, налаштування, моніторингу та віддаленого управління мобільними пристроями з одного централізованого інтерфейсу. Основна мета MDM полягає в забезпеченні безпеки, конфігурації та керування мобільними пристроями в організації. Деякі можливості MDM включають: можливість реєстрації мобільних пристроїв в системі MDM та їх прив'язка до конкретних користувачів або груп; здатність встановлення та забезпечення політики безпеки на мобільних пристроях, включаючи налаштування паролів, шифрування даних, віддалене видалення даних тощо; можливість віддалено налаштовувати пристрої та встановлювати додатки, контролювати доступ до магазинів додатків та обмежувати встановлення небезпечних або небажаних додатків; здатність збирати дані про стан та використання мобільних пристроїв, включаючи інформацію про батарею, сигнал мережі, використання пам'яті.

Не слід забувати й про користувача інформаційних систем. Користувач подекуди може бути найслабкішим та вразливішим елементом. Обізнаність користувача щодо безпеки при користуванні мобільним пристроєм є важливим аспектом для забезпечення захисту особистих даних і запобігання кіберзлочинності. Всі ці моменти слід забезпечувати, контролювати та поліпшувати. Можна виділити основні аспекти: захист паролів: користувачі повинні встановлювати міцні паролі або використовувати біометричні методи аутентифікації для захисту своїх пристроїв; важливо використовувати унікальні та складні паролі і не ділитися ними з іншими особами; оновлення програмного забезпечення регулярно, оновлення операційної системи та програмного забезпечення на пристрої є важливим, оскільки виробники випускають оновлення та виправлення для заповнення потенційних слабких місць; встановлення довірених додатків, користувачі повинні завантажувати додатки тільки з офіційних магазинів додатків, таких як Google Play або App Store, і перевіряти рейтинги та відгуки інших користувачів перед встановленням; уважність при відкритті посилань та відправці особистої інформації, користувачі повинні бути обережними при натисканні на посилання в недовіреному повідомленні або електронній пошті та уникати надсилання особистих чутливих даних через ненадійні канали комунікації; використання антивірусного програмного забезпечення, встановлення добре відомого антивірусного програмного забезпечення на мобільний пристрій може допомогти виявити і запобігти інфікуванню шкідливим програмним забезпеченням.

Висновок

Майбутнє захисту неструктурованої інформації на мобільних пристроях буде засноване на поєднанні технологічних нововведень та підвищенні обізнаності користувачів з питань кібербезпеки. Деякі напрямки розвитку включають застосування штучного інтелекту (AI) та машинного навчання: AI може допомогти в розробці розширених механізмів захисту, що

автоматично виявляють аномальні або підозрілі дії та блокують потенційні загрози; біометрична автентифікація, покращення технологій біометричної автентифікації, таких як розпізнавання обличчя, відбитків пальців, можуть забезпечити більш надійний захист неструктурованої інформації; застосування блокчейн-технологій: блокчейн може відігравати важливу роль у забезпеченні безпеки інформації, оскільки він дозволяє зберігати дані розподілено та захищено від змін; розвиток квантово-стійких алгоритмів шифрування: з розвитком квантових комп'ютерів створюється потреба в нових алгоритмах шифрування, які можуть витримати потужність квантового злому; приватність за замовчуванням (Privacy by Design): проектування мобільних пристроїв та застосунків з врахуванням приватності користувачів сприятиме захисту неструктурованої інформації на мобільних пристроях; освіта та обізнаність користувачів: підвищення обізнаності користувачів щодо кібербезпеки та навчання їх ефективним методам захисту інформації на мобільних пристроях допоможе зменшити ризик виникнення інцидентів, пов'язаних з безпекою; розвиток законодавства та нормативної бази, вдосконалення законодавства та нормативної бази у галузі кібербезпеки сприятиме створенню надійного середовища для захисту неструктурованої інформації. Захист неструктурованої інформації на мобільних пристроях спиратиметься на комбінації технологічного прогресу, навчання користувачів та співпраці між різними сторонами, такими як розробники, постачальники послуг та законодавці. Забезпечення безпеки та приватності даних в цифровому світі залишається пріоритетом, оскільки кількість мобільних пристроїв та їх використання продовжують зростати.

Сучасний користувач мобільних пристроїв має певні вимоги що стосуються на ергономіку та рівень інформаційної обізнаності. Розвиток галузі в цілому вимагає швидких, надійних та ефективних рішень що потрапляють не лише в економічну площину а й площину кібербезпеки. Використання мобільних пристроїв може адаптуватися до швидких сучасних потреб безпеки організації та може бути легко налаштований відповідно до конкретних вимог. Мобільний пристрій є зручним і доступним для багатьох користувачів, він зазвичай знаходиться поруч з користувачем і може використовуватися як додатковий елемент безпеки.

Література

1. Бровченко, Є. М., Самарай, В. П., Даценко, І. П., Павленко, В. І. (2023). Мобільний пристрій як частина адаптивної кейс-менеджмент системи. Інфокомунікаційні та комп'ютерні технології, 2(04), С. 157–163. <https://doi.org/10.36994/2788-5518-2022-02-04-18>.
2. Choi, J. S., & Kim, H. S. (2019). An Efficient Encryption Method for Securing Unstructured Data on Mobile Devices. IEEE Access, 7, 124616-124626.
3. Krcmar, Helmut, ed. Mobile Computing: Concepts, Methodologies, Tools, and Applications. IGI Global, 2018.
4. Liu, Z., Huang, L., Gao, Y., Wang, Z., & Yang, X. (2017). Research on cloud computing security and privacy protection model. Cluster Computing, 20(3), 1893-1901.
5. Madden J., Madden B. Enterprise Mobility Management: Everything you need to know about MDM, MAM, and BYOD. – Jack Madden, 2013. 176 p.
6. Sarker, S., & Chatterjee, S. (2018). Big data analytics: A literature review paper. Journal of Business Research, 93, 240-253.
7. Yang, Y., Luo, X., & Wu, J. (2019). Protecting Sensitive Data in Mobile Devices with Security-Enhanced Linux and Encryption Technologies. Mobile Networks and Applications, 24(5), 1648-1660.
8. The official Android Enterprise website: <https://www.android.com/enterprise/>

References

1. Evgen Brovchenko, Valery Samarai, Ivan Datsenko, Volodymyr Pavlenko. (2023). MOBILE DEVICE AS PART OF AN ADAPTIVE CASE MANAGEMENT SYSTEM. INFOCOMMUNICATION AND COMPUTER TECHNOLOGIES, 2(04), P. 157–163. <https://doi.org/10.36994/2788-5518-2022-02-04-18>. [in Ukrainian]
2. Choi, J. S., & Kim, H. S. (2019). An Efficient Encryption Method for Securing Unstructured Data on Mobile Devices. IEEE Access, 7, 124616-124626.
3. Krcmar, Helmut, ed. Mobile Computing: Concepts, Methodologies, Tools, and Applications. IGI Global, 2018.
4. Liu, Z., Huang, L., Gao, Y., Wang, Z., & Yang, X. (2017). Research on cloud computing security and privacy protection model. Cluster Computing, 20(3), 1893-1901.
5. Madden J., Madden B. Enterprise Mobility Management: Everything you need to know about MDM, MAM, and BYOD. – Jack Madden, 2013. 176 p.
6. Sarker, S., & Chatterjee, S. (2018). Big data analytics: A literature review paper. Journal of Business Research, 93, 240-253.
7. Yang, Y., Luo, X., & Wu, J. (2019). Protecting Sensitive Data in Mobile Devices with Security-Enhanced Linux and Encryption Technologies. Mobile Networks and Applications, 24(5), 1648-1660.
8. The official Android Enterprise website: <https://www.android.com/enterprise/>

УДК 534.21: 537.226.86

ХВИЛІ ЗСУВУ В ШАРУВАТО-ПЕРІОДИЧНІЙ СТРУКТУРІ «П'ЄЗОЕЛЕКТРИК - ЩІЛИНА»

DOI 10.36994 / 2788-5518-2023-01-05-22

Левченко В.В., к.ф.-м. н., доц. Державний університет телекомунікацій, Київ, Україна, e-mail: ylvv@ukr.net

Павленко В. І., к.ф.-м.н., доц., Відкритий міжнародний університет розвитку людини «Україна», Інститут комп'ютерних технологій, Київ, Україна. <https://orcid.org/0000-0002-3958-0415>, pavlenko.v@i.ua

Таланчук П.М., д.т.н., проф., Відкритий міжнародний університет розвитку людини «Україна», Київ, Україна. <https://orcid.org/0000-0001-8748-6127>, talanshuk_pm@ukr.net

Анотація. Дано постановку і розв'язано задачу про поширення об'ємних акустоелектричних пружних хвиль зсуву у регулярно - шаруватому середовищі п'єзоелектрик-щілина. Отримано дисперсійні співвідношення, які дозволяють дослідити можливість поширення хвильових полів з довільним геометриєю періоду структури. Проведено порівняння спектрів дисперсійних кривих для об'ємних акустоелектричних хвиль у регулярно – шаруватому просторі.. Отримані дисперсійні співвідношення аналізувалися також чисельно, а результати наведено у вигляді графіків. Наведено графіки меж границь зон пропускання для різних видів геометрії породжуючого пакету шарів.

Ключові слова: акустоелектричні об'ємні хвилі зсуву, дисперсійні співвідношення, регулярно шарувате середовище, п'єзоелектрик-щілина.

BULK SHEAR WAVES IN A LAYER--PERIODIC STRUCTURE "PIEZOELECTRIC - SLOT"

Volodymyr Levchenko, Ph.D. n., Assoc. State University of Infrastructure and Technology, Kyiv, Ukraine, e-mail: ylvv@ukr.net

Volodymyr Pavlenko., Ph.D. n., Assoc. Open University of Human Development "Ukraine", Kyiv, Ukraine. <https://orcid.org/0000-0002-3958-0415>, pavlenko.v@i.ua

Peter Talanchuk, Dr. habil., Prof., Open International University of Human Development "Ukraine", Kyiv, Ukraine. <https://orcid.org/0000-0001-8748-6127>, talanshuk_pm@ukr.net

Abstract. A large number of scientific works are devoted to the problems of propagation of waves of different physical nature in periodic and heterogeneous media. The main problem considered in the article was the study of the shear wave in the layered-periodic structure "piezoelectric - gap". The statement is given and the problem of the propagation of volumetric acoustoelectric elastic shear waves in a regularly layered piezoelectric-slit medium is solved. Dispersion relations were obtained, which allow to investigate the possibility of propagation of wave fields with an arbitrary geometry of the period of the structure. A comparison of the spectra of the dispersion curves for volumetric acoustoelectric waves in a regularly layered space was carried out. The obtained dispersion relations were also analyzed numerically, and the results are presented in the form of graphs. Graphs of the limits of the boundaries of the transmission zones for various types of geometry of the generating layer package are given. As a result of numerical experiments, the possibility of the existence of transmission zones for bulk acoustic-electric waves in regularly layered media of the "piezoelectric - slot" type was established.

Key words: acoustoelectric bulk shear waves, dispersion relations, regularly layered medium, piezoelectric- slot.

Вступ

Досить широке застосування в різних галузях науки і техніки знайшли і шарувато-періодичні структури з електропружних матеріалів, що зумовлює потребу в детальному моделюванні та вивченні закономірностей поширення акустоелектричних хвиль у шаруватих середовищах з п'єзоефектом. У рамках концепцій механіки суцільного середовища для побудови та аналізу дисперсійних рівнянь, що описують поширення об'ємних, поверхневих та хвильоводних мод різної

поляризації в шарувато-періодичних середовищах різних типів, було запропоновано низку підходів [1–3]. У цій роботі на основі підходу [1] вивчаються об'ємні акустоелектричні хвилі зсуву, що поширюються в регулярно-шаруватій структурі « п'єзоелектрик -вакуумний зазор». Дисперсія таких хвиль визначається зв'язком п'єзоелектричних шарів електричними полями через проміжок [4]. який віднесемо до декартової системи координат $Oxyz$

Постановка задачі.

Розглянемо регулярно-шаруватий простір, який утворений періодичним повторенням вздовж осі Ox шару п'єзоелектрика товщини h_p , фізико-механічні властивості якого описуються матеріальними співвідношеннями кристалографічного класу $6mm$ і вакуумного зазору (щілини) товщини h_v . Направимо вісь Oz декартової системи координат вздовж осі симетрії шостого порядку. Розповсюдження вздовж осі Oy зсувної хвилі в п'єзоелектричних шарах описується лінеаризованою системою рівнянь електропружності [3]

$$\rho_p \partial_t^2 u_p = c_{44,p}^* \Delta u_p, \Delta \psi_p = 0, \quad (1)$$

а в щілині рівнянням [5]

$$\Delta \phi_0 = 0, \quad (2)$$

де введені позначення $c_{44,p}$, ρ_p , $\varepsilon_{11,p}$, $e_{15,p}$ – фізико-механічні параметри п'єзоелектричного шару, $c_{44,p}^* = c_{44,p} + e_{15,p}^2 / \varepsilon_{11,p}$, $\psi_p = \phi_p - \left(\frac{e_{15,p}}{\varepsilon_{11,p}} \right) u_p$, а ϕ_p й ϕ_0 – електростатичні потенціали відповідно у п'єзоелектриці та вакуумі.

Розв'язання задачі.

Розв'язання системи рівнянь (1)–(2) у кожному з шарів шукатимемо у вигляді

$$u_p = B_{2n-1}^{(1)} \sin \Omega_p (x - x_{n,p}^*) + B_{2n-1}^{(2)} \cos \Omega_p (x - x_{n,p}^*), \quad \psi_p = D_{2n-1}^{(1)} \operatorname{sh} k (x - x_{n,p}^*) + D_{2n-1}^{(2)} \operatorname{ch} k (x - x_{n,p}^*), \quad x_{n-1,v}^* < x < x_{n,p}^*;$$

$$\phi_0 = D_{2n}^{(1)} \operatorname{sh} k (x - x_{n,v}^*) + D_{2n}^{(2)} \operatorname{ch} k (x - x_{n,v}^*), \quad x_{n,p}^* < x < x_{n,v}^*; \quad n = 0, \pm 1, \pm 2, \dots \quad (3)$$

де $x_{n,p}^* = nh + h_p$, $x_{n,v}^* = nh + h$, $h = h_v + h_p$, $\Omega_p = (k_p^2 - k^2)^{1/2}$, $k_p^2 = \omega^2 / \tilde{c}_p^2$, $\tilde{c}_p^2 = c_{44,p}^* / \rho_p$, k – хвильове число, ω – кругова частота. У формулах (3) множник $\exp(iky - i\omega t)$ опущений.

На межах розділу властивостей аналізованого середовища $x = x_{n,p}^*$ і $x = x_{n,v}^*$ повинні виконуватися умови [5]

$$c_{44,p}^* \partial_x u_p + e_{15,p} \partial_x \psi_p = 0, \quad \partial_x \psi_p = \partial_x \phi_0, \quad \psi_p - \frac{e_{15,p}}{\varepsilon_{11,p}} u_p = \phi_0. \quad (4)$$

Ці умови виражають відсутність зсувних напруг, безперервність електричних потенціалів та нормальних компонентів електричної індукції.

Підставляючи рішення (3) у граничні умови (4), вихідне завдання зведемо до нескінченної системи алгебраїчних рівнянь щодо невідомих $\vec{B}_j$ та $\vec{D}_i$

$$\begin{aligned} N(\varepsilon_{11,p}; k_p^*) \vec{D}_{2n-1} + N_u(e_{15,p}; \theta_p) \vec{B}_{2n-1} &= N(\varepsilon_0; 0) \vec{D}_{2n}, \\ N(\varepsilon_0; k_v^*) \vec{D}_{2n} &= N_u(e_{15,p}; 0) \vec{B}_{2n+1} + N(\varepsilon_{11,p}; 0) \vec{D}_{2n+1}, \\ \vec{M}(a_p; \theta_p) \vec{B}_{2n-1} + \vec{M}_u(e_{15,p}; k_p^*) \vec{D}_{2n-1} &= 0, \\ \vec{M}(a_p; 0) \vec{B}_{2n+1} + \vec{M}_u(e_{15,p}; 0) \vec{D}_{2n+1} &= 0, \end{aligned} \quad (5)$$

$n = 0, \pm 1, \pm 2, \dots$

У системі (5) введено матриці

$$N(\varepsilon_{11,p}, k_p^*) = \begin{bmatrix} \varepsilon_{11,p} k \operatorname{ch} k_p^* & -\varepsilon_{11,p} k \operatorname{sh} k_p^* \\ -\operatorname{sh} k_p^* & \operatorname{ch} k_p^* \end{bmatrix},$$

$$N(\varepsilon_0, k_v^*) = \begin{bmatrix} \varepsilon_0 k \operatorname{ch} k_v^* & -\varepsilon_0 k \operatorname{sh} k_v^* \\ -\operatorname{sh} k_v^* & \operatorname{ch} k_v^* \end{bmatrix},$$

$$N(e_{15,p}, \theta_p) = \begin{bmatrix} 0 & 0 \\ -e_{15,p} k \sin \theta_p & e_{15,p} k \cos \theta_p \end{bmatrix},$$

вектор-рядки $\vec{M}(a_p; \theta_p) = [a_p \cos \theta_p; a_p \sin \theta_p]$, $\vec{M}_u(e_{15,p}; k_p^*) = [e_{15,p} k \operatorname{ch} k_p^*; -e_{15,p} k \operatorname{sh} k_p^*]$, вектор-стовпці $\vec{B}_i = \operatorname{col}(B_i^{(1)}, B_i^{(2)})$, $\vec{D}_i = \operatorname{col}(D_i^{(1)}, D_i^{(2)})$ та позначення $k_p^* = k h_p$, $\theta_p = \Omega_p h_p$, $a_p = c_{44,p}^* \Omega_p$, $k_v^* = k h_v$.

Виконавши ряд перетворень, систему рівнянь (5) можна звести до системи щодо невідомих $\vec{D}_i$

$$\begin{aligned} \vec{N}(e_{11,p}; k_p^*; \theta_p) \vec{D}_{2n-1} &= N(\varepsilon_0; 0) \vec{D}_{2n}, \\ N(\varepsilon_0; k_v^*) \vec{D}_{2n} &= \vec{N}(e_{11,p}; 0; 0) \vec{D}_{2n+1}, \quad n = 0, \pm 1, \pm 2, \dots, \end{aligned} \quad (6)$$

де елементи матриці $\vec{N}(e_{11,p}; k_p^*; \theta_p)$ мають вигляд

$$\begin{aligned} \vec{N}^{11} &= \varepsilon_{11,p} k \operatorname{ch} k_p^*, \quad \vec{N}^{12} = -\varepsilon_{11,p} k \operatorname{sh} k_p^*, \quad \vec{N}^{21} = -\operatorname{sh} k_p^* - \frac{e_{15,p}}{\varepsilon_{11,p}} R^{11} \sin \theta_p + \frac{e_{15,p}}{\varepsilon_{11,p}} R^{21} \cos \theta_p, \\ \vec{N}^{22} &= \operatorname{ch} k_p^* + \frac{e_{15,p}}{\varepsilon_{11,p}} R^{22} \cos \theta_p; \quad R^{11} = -\frac{k e_{15,p}}{a_p}, \quad R^{21} = \frac{-k e_{15,p} \operatorname{ch} k_p^* - a_p R^{11} \cos \theta_p}{a_p \sin \theta_p}, \quad R^{22} = \frac{k e_{15,p} \operatorname{sh} k_p^*}{a_p \sin \theta_p}. \end{aligned}$$

Використовуючи запропонований у [1] підхід, загальне рішення системи рівнянь (6) представимо у вигляді

$$\begin{aligned} \vec{D}_{2n-1} &= \sum_{j=1}^2 \chi_j^n \vec{N}^{-1}(e_{11,p}; 0; 0) N(\varepsilon_0; k_v^*) N^{-1}(\varepsilon_0; 0) \vec{Y}_j, \\ \vec{D}_{2n} &= \sum_{j=1}^2 \chi_j^n N^{-1}(\varepsilon_0; 0) \vec{Y}_j, \quad n = 0, \pm 1, \pm 2, \dots, \end{aligned} \quad (7)$$

де χ_j^n і $\vec{Y}_j$ ($j = 1, 2$) – характеристичні числа та відповідні їм власні вектори передавальної матриці

$$N = \vec{N}(e_{11,p}; k_p^*; \theta_p) \vec{N}^{-1}(e_{11,p}; 0; 0) N(\varepsilon_0; k_v^*) N^{-1}(\varepsilon_0; 0). \quad (8)$$

На основі системи рівнянь (5) отримаємо рішення для $\vec{B}_{2n-1}$ у вигляді

$$\vec{B}_{2n-1} = \sum_{j=1}^2 (\vec{R}_1 + \vec{R}_2) \chi_j^n \vec{N}^{-1}(e_{11,p}; 0; 0) N(\varepsilon_0; k_v^*) N^{-1}(\varepsilon_0; 0) \vec{Y}_j,$$

де введені вектор-стовпці $\vec{R}_1 = \operatorname{col}(R^{11}; R^{22})$, $\vec{R}_2 = \operatorname{col}(R^{21}; 0)$

Безпосередні обчислення показали, що детермінант матриці $\det N = 1$. Тоді згідно [1], умова існування об'ємних електропружних хвиль зсуву має вигляд

$$|b_{pv}| \leq 1, \quad b_{pv} = \operatorname{spur} N / 2, \quad (9)$$

а рівняння

$$b_{pv} = \pm 1 \quad (10)$$

задають межі цих зон.

У граничному випадку, коли товщина щілини $h_v = 0$, дисперсійне рівняння (10) для меж зон пропускання набуде вигляду

$$a_p \sin \theta_p (\operatorname{ch} k_p^* \mp 1) + \frac{k(e_{15,p})^2}{\varepsilon_{11,p}} (\cos \theta_p \mp 1) = 0. \quad (11)$$

Знак «мінус» в (11) відповідає межах, де $b_{pv} = 1$. Незавжди побачити, що $\cos \theta_p = \pm 1$ визначатиме коріння останнього рівняння.

Аналіз чисельних результатів та висновки.

Нижче наведено результати чисельного аналізу отриманих дисперсійних рівнянь для меж зон пропускання у разі коли шар п'єзоелектрика має параметри CdS. На рис.1 по осі абсцис відкладено безрозмірне хвильове число $\bar{k} = k h_p$, а по осі ординат – безрозмірна частота $\bar{\omega} = \omega h_p / c_p$ ($c_p^2 = c_{44,p} / \rho_p$). В силу того, що межі зон пропускання лежать досить близько одна до одної, на малюнку представлений лише фрагмент другої зони пропускання, який відображає основну закономірність зміни положення меж зон за різних товщин вакуумного зазору.

При чисельному аналізі товщина шару CdS залишалася постійною $\bar{h}_p = h_p/h_p = 1$, а товщина щілини $\bar{h}_v = h_v/h_p$ була такою: $\bar{h}_v = 0$ (крива 1), $\bar{h}_v = 0,05$ (крива 2), $\bar{h}_v = 0,5$ (крива 3). Зазначені криві відповідають тим межах зон, де $b_{pv}=1$. У межах масштабу малюнку межі зон, які визначаються рівнянням $b_{pv}=-1$ та відповідають різним значенням $\bar{h}_v$, практично збігаються і тому позначені однією кривою 4.

Як очевидно з малюнку, кордон $b_{pv}=1$ зі збільшенням $\bar{h}_v$ наближається до кордону $b_{pv}=-1$.

На рис.2 в координатах $(\bar{k}, \bar{h}_v)$ для другої зони пропускання і $\bar{\omega} = 3,1$ нанесена суцільною кривою $b_{pv}=-1$, і штрихова крива, відповідно $b_{pv}=1$. Наведені результати дозволяють простежити зміну ширини зони пропускання зі зростанням товщини вакуумного зазору.

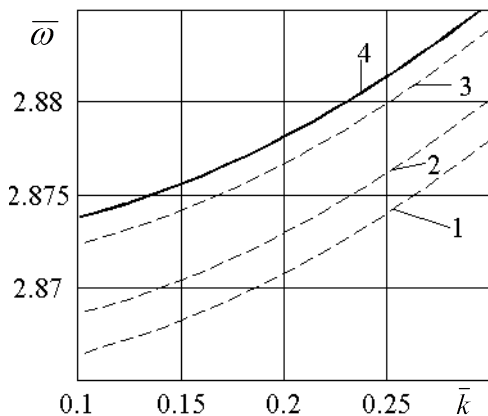


Рис. 1. Межі зон пропускання

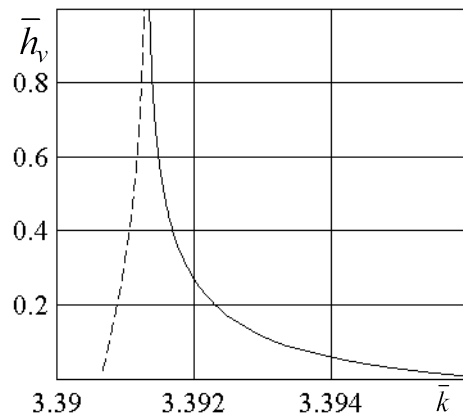


Рис. 2.. Зона пропускання.

Висновки

Таким чином, в результаті чисельних експериментів встановлено можливість існування зон пропускання для об'ємних акустoeлектричних хвиль у регулярно-шарових середовищах типу «п'єзoeлектрик – щілина». Ці зони дуже вузькі в порівнянні з зонами запирання і локалізовані в околиці кривих $\cos \theta_p = \pm 1$. Зі збільшенням товщини вакуумного зазору $\bar{h}_v$ ширина зони пропускання стає дедалі меншою й межі зон вироджується в лінію. Встановлено, що зі зростанням $\bar{k}$ та $\bar{\omega}$ ширина зон незначно збільшується. Отримані дисперсійні співвідношення проаналізовано чисельно, а результати представлено у вигляді графіків.

Література

1. Шульга Н.А. Основи механіки шаруватих середовищ періодичної структури. - К.: Наук. думка, 1981. 200 с.
2. Podlipenets A. N., Shulga N. A., Zinchuk L. P. Hamiltonian system formalism in the theory of transparent wave propagation in multilayered piezoelectrics // Zeitschrift für Angewandte Mathematik und Mechanik. 1996. V. 76, S. 5. P. 395-396.
3. Зінчук Л. П., Левченко В. Ст, Шульга Н. А. Поширення об'ємних електроупругих хвиль зсуву в регулярно-шаруватому середовищі типу метал - п'єзoeлектрик // Мат. методи та фіз. - хутро. поля. 1989. Вип. 30. С.4-8.
4. Rushitsky Ya.Ya. On a nonlinearly elastic Stoneley wave// Prikl. Mechanics. — 2014. —50, No. 6 – P. 39
5. Levchenko V.V. Propagation of magnetoelastic shear waves through a regularly laminated medium with metalized interfaces //Int. Appl. Mech. - 2004. - 40, No. 1. – P.97-25.

References

1. Shulga N.A. Fundamentals of mechanics of layered media of periodic structure. - K.: Science. dumka, 1981. 200 p.
2. Podlipenets A. N., Shulga N. A., Zinchuk L. P. Hamiltonian system formalism in the theory of transparent wave propagation in multilayered piezoelectrics // Zeitschrift für Angewandte Mathematik und Mechanik. 1996. V. 76, S. 5. P. 395-396.
3. L. P. Zinchuk, V. St. Levchenko, and N. A. Shulga, Propagation of bulk electroelastic shear waves in a regularly layered medium of the metal-piezoelectric type // Mat. methods and physical - fur fields 1989. Issue 30. P.4-8.
4. Rushitsky Ya. Ya. On a nonlinearly elastic Stoneley wave// Prikl. Mechanics. — 2014. —50, No. 6 – P. 39
5. Levchenko V.V. Propagation of magnetoelastic shear waves through a regularly laminated medium with metalized interfaces //Int. Appl. Mech. - 2004. - 40, No. 1. – P.97-25.

ПРАВИЛА ОФОРМЛЕННЯ СТАТЕЙ

Статті подаються в обсязі 7-12 сторінок формату А4.

Текст в редакторі Word, шрифт –Arial , 11 пт, інтервал - 1, розміри полів: ліве - 2,5 см., праве - 1 см. Абзаци- 0,75 см. Формули у внутрішньому редакторі MS Office Word.

Порядок оформлення: УДК 11 пт, назва статті заголовним 12 пт, ПІБ-все жирно ; назва університету, місто, країна, e-mail ,10 п. –все курсивом.

Анотації курсивом 10 п. Слово «Анотація» жирно.

В статті мають бути розділи: вступ, дослідження або розробка, висновки. Заголовки 11 п. жирно.

Підпис рисунків і таблиць 10 п.

В кінці наводиться список літератури, 9 п, не більше 15 джерел. . Оформлення літератури за стандартом ДСТУ 8302:2015.

ARTICLE DESIGN RULES

Articles are submitted in the amount of 7-12 pages of A4 format.

Text in Word editor, font - Arial, 11 pt, spacing - 1, field sizes: left - 2.5 cm, right - 1 cm. Paragraphs - 0.75 cm. Formulas in the internal editor MS Office Word.

Procedure: UDC 11 pt, title of the article heading 12 pt, full name, everything is bold; name of university, city, country, e-mail, 10 points - all in italics.

Annotations in italics 10 n. The word "Abstract" is bold.

The article should have sections: introduction, research or development, conclusions. Headings 11 sts bold.

Caption of figures and tables 10 p.

At the end is a list of references, 9 items, no more than 15 sources. . Registration of literature according to the standard DSTU 8302: 2015.

Тираж 50 пр. Зам. №

Сайт: visn-icct.uu.edu.ua/

E-mail: infctt@ukr.net

Адреса редакції: м. Київ, вул. Львівська, 23, тел. (044) 424-62-74

Відповідальний редактор *Семенко А.І.*

Комп'ютерна верстка *Войтенко В.М.*

Дизайн обкладинки *Бабинець Н.А.*

Видавець і виготовлювач університет «Україна»

03115, м. Київ, вул. Львівська, 23, тел./факс (044) 424-24-14, 424-56-26

E-mail: ukraina.vdk@email.ua

Свідоцтво суб'єкта видавничої справи ДК №405 від 06.04.01