



НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «КИЄВО-МОГИЛЯНСЬКА АКАДЕМІЯ»
TEMPUS FUGIT, ACADEMIA SEMDITERNA

16 Міжнародна науково-практична конференція

***«Теоретичні та прикладні аспекти побудови
програмних систем»***

***«Theoretical and Applied Aspects of Program Systems
Development»***

TAAPSD 2025

***Праці конференції
Proceeding***

**23-24 листопада 2025 року
КИЇВ**

**23-24 November 2025
KYIV**

УДК 004.4

Друкується за рішенням Вченої ради
факультету інформатики Національного університету «Києво-Могилянська
академія»
(протокол №1 від 09.01.2025 р.)

Теоретичні та прикладні аспекти побудови програмних систем: праці
16 міжнародної науково-практичної конференції, Київ, 24-25 листопада 2025 р.
/за заг. ред.: М.М.Глибовця, Т.В.Панченка та ін. /
Київ: Видавництво НаУКМА, 2025, --- с.
У збірнику представлено тези доповідей науково-практичної
конференції «Теоретичні та прикладні аспекти побудови програмних систем»
(ТАAPSD 2025), яка проходила в Києві з 24 по 25 листопада 2025

© НаУКМА, 2025

ОРГАНІЗАТОРИ КОНФЕРЕНЦІЇ:

**Міністерство освіти і науки України
Національна академія наук України
Факультет інформатики Національного університету
«Києво-Могилянська академія»
Факультет комп'ютерних наук та кібернетики
Київського національного університету ім. Тараса Шевченка
Інститут програмних систем НАН України**

Міжнародний програмний комітет:

П. Андон (Україна) – Голова
А. Анісімов (Україна) – Співголова
М. Глибовець (Україна) – Співголова
І. Сініцин (Україна) – Співголова
С. Горлач (Німеччина)
К. Хенгскен (Німеччина)
С. Єршов (Україна)
М. Нікітченко (Україна)
М. Медведєв (Республіка Азербайджан)
І. Didmanidze (Грузія)

Організаційний комітет:

А. Глибовець (Україна) – Голова
Т. Панченко (Україна) – Співголова
В. Шевченко (Україна) – Співголова
О. Коник (Україна)
В. Колінько (Україна)
О. Марченко (Україна)
Г. Малашонок (Україна)
С. Гороховський (Україна)
О. Жежерун (Україна)
О. Олецкий (Україна)
О. Смиш (Україна)
О. Кирієнко (Україна)

ЗАГАЛЬНА ІНФОРМАЦІЯ

Форма проведення конференції: очно-заочна інтернет-конференція з дублюванням відео всіх доповідей через мережу Інтернет.
Секційні засідання відбулися 24-25 листопада.
Доповіді проходили в онлайн-режимі.

=====

Організаційний комітет ТАAPSD спільно з MacPaw визначив найкращі доповіді конференції:

MASOVO ПАРАЛЕЛЬНА РЕАЛІЗАЦІЯ MORPHONAS / MASSIVELY PARALLEL MORPHONAS IMPLEMENTATION

Медвідь С.О. / Medvid S.

КОНТРОЛЬ ІНТЕНСИВНОСТІ ЕМОЦІЙ В УКРАЇНСЬКОМУ СИНТЕЗИ МОВЛЕННЯ ЧЕРЕЗ КРОС-МОВНЕ ПЕРЕНОСЕННЯ ЗНАНЬ ТА АДАПТЕРИ LORA / EMOTION INTENSITY CONTROL IN UKRAINIAN TTS THROUGH CROSS-LINGUAL TRANSFER KNOWLEDGE AND LORA ADAPTERS

Іващенко Д.С. / Ivashchenko D.

ГЕНЕРАТИВНИЙ ФРЕЙМВОРК ДЛЯ ПОБУДОВИ ВІЗУАЛЬНО-ТЕКСТОВИХ ДАТАСЕТІВ НА ОСНОВІ ОНТОЛОГІЙ / ONTOLOGY-DRIVEN FRAMEWORK FOR VQA DATASET GENERATION

Чоловський С.О., Здирко В.В. / Cholovskyi S., Zdyrko V.

=====

З М І С Т

СЕКЦІЯ 1. КОМП'ЮТЕРНІ НАУКИ

- ІНТЕЛЕКТУАЛЬНІ ПРОГРАМНІ СИСТЕМИ ПІДТРИМКИ ІІ / INTELLIGENT AI SOFTWARE SUPPORT SYSTEMS** / Глибовець М.М., доктор фізико-математичних наук, професор кафедри інформатики факультету інформатики Національного університету «Києво-Могилянська академія» / Національний університет «Києво-Могилянська академія» / 04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики / E-mail: glib@ukma.edu.ua 11
- РОЗШИРЕННЯ КРИТЕРІЮ ТЕСТОВОГО ПОКРИТТЯ ДЛЯ ВИЯВЛЕННЯ СЛАБКОЇ ВАЛІДАЦІЇ ВХІДНИХ ДАНИХ / EXTENDING TEST COVERAGE CRITERION TO REVEAL WEAK INPUT VALIDATION** / Колчин О.В., старший науковий співробітник, кандидат фіз.-мат. наук, старший науковий співробітник / Руденко Д.Б., аспірантка / Інститут кібернетики імені В.М. Глушкова НАН України / 0380 МСП, просп. Академіка Глушкова, 40 / E-mail: kolchin_av@yahoo.com, dasharudenko1902@gmail.com 13
- КОНЦЕПЦІЯ СЕРВІСУ ОФЛАЙН МАП НА ОСНОВІ CESIUMJS** / Ткаченко В.О., студент аспірантської програми / Національний університет «Києво-Могилянська академія» / 04655, м. Київ, вулиця Григорія Сковороди, 2 / E-mail: vo.tkachenko@ukma.edu.ua 15
- SYSTEM AND APPROACH FOR AUTOMATED PHOTO-LIKE FACIAL IMAGE GENERATION FOR USE IN TESTING, TRAINING OR EVALUATION** / Артюшенко Б.А., к.т.н., ст. викл (сумісник) Кафедра Інформатики Факультет Інформатики / National University of Kyiv-Mohyla Academy 04655, Kyiv, Hryhorii Skovoroda St 2. E-mail: b.artiushenko@ukma.edu.ua 17
- ПРО ЗАСТОСУВАННЯ ЦИФРОВИХ ДВІЙНИКІВ В ІНТЕЛЕКТУАЛІЗОВАНИХ МЕРЕЖАХ ПУБЛІЧНИХ СЕРВІСІВ / ON THE APPLICATION OF DIGITAL TWINS IN INTELLECTUALIZED NETWORKS OF PUBLIC SERVICES** / Олецкий О.В., кандидат технічних наук, доцент, доцент кафедри мультимедійних систем факультету інформатики / Моголівський В.О., аспірант програми «Комп'ютерні науки» факультету інформатики, асистент кафедри мультимедійних систем факультету інформатики / Національний університет «Києво-Могилянська Академія» / 04655, Київ, вул. Григорія Сковороди, 2 / E-mail: oletsky@ukma.edu.ua, v.moholivskiy@ukma.edu.ua 19
- NOWCASTING ЯК СУЧАСНИЙ ПІДХІД ДО ОЦІНЮВАННЯ ВВП УКРАЇНИ: ПОРІВНЯННЯ З ТРАДИЦІЙНИМИ МОДЕЛЯМИ ПРОГНОЗУВАННЯ / NOWCASTING AS A MODERN APPROACH TO ESTIMATING UKRAINE'S GDP: COMPARISON WITH TRADITIONAL FORECASTING MODELS** / Болотов Є.С., студент 2-го курсу магістратури за спеціальністю F1 «Прикладна математика» / Дрінь С.С., кандидат фізико-математичних наук, старший викладач / Національний університет «Києво-Могилянська Академія» / 04655, Київ, вул. Григорія Сковороди, 2 / E-mail: yehor.bolotov@ukma.edu.ua, svitlana.drin@ukma.edu.ua 21
- ФРАКТАЛЬНО-ДИФУЗІЙНІ ГЕНЕРАТИВНІ МОДЕЛІ: ІЄРАРХІЧНИЙ ПІДХІД ДО СИНТЕЗУ ЗОБРАЖЕНЬ / FRACTAL-DIFFUSION GENERATIVE MODELS: A HIERARCHICAL APPROACH TO IMAGE SYNTHESIS** / Шалімов А.В., магістрант спеціальності 113 «Прикладна математика» / Авраменко О.В., доктор фізико-математичних наук, професор, професор кафедри математики / Національний університет «Києво-Могилянська Академія» / 04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра математики / E-mail: a.shalimov@ukma.edu.ua, o.avramenko@ukma.edu.ua 23
- ІНТЕГРАЦІЯ МЕТОДУ АНАЛІЗУ ІЄРАРХІЙ ІЗ НЕЧІТКОЮ ЛОГІКОЮ ДЛЯ ВИРІШЕННЯ БАГАТОКРИТЕРІАЛЬНИХ ЗАДАЧ / INTEGRATION OF THE ANALYTICAL HIERARCHY METHOD WITH FUZZY LOGIC FOR SOLVING MULTI-CRITERION PROBLEMS** / Тригуб О.С., доцент кафедри інформатики НаУКМА, кандидат фізико-математичних наук, доцент / Національний університет «Києво-Могилянська

Академія» / 04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики / E-mail: oleksandr.tryhub@ukma.edu.ua25

ТОПОЛОГІЧНА СКЛАДНІСТЬ ВУЛИЧНИХ МЕРЕЖ УКРАЇНСЬКИХ МІСТ: ЕНТРОПІЙНИЙ АНАЛІЗ / Смиш О.Р., доктор філософії (PhD) з комп'ютерних наук, старший викладач кафедри мультимедійних систем факультету інформатики / Національний університет «Києво-Могилянська академія» / вул. Г. Сковороди, 2, м. Київ, 04070, тел.: +380 044 425 77 53 / E-mail: o.smysh@ukma.edu.ua27

МОДЕЛЮВАННЯ ВНУТРІШНІХ ГРАВІТАЦІЙНО-КАПІЛЯРНИХ ХВИЛЬ ЗА ДОПОМОГОЮ PINN / PINN MODELING OF INTERFACIAL GRAVITY-CAPILLARY WAVES / Авраменко О.В., доктор фізико-математичних наук, професор, професор кафедри математики / Сотников М.О., магістрант спеціальності 113 «Прикладна математика» / Національний університет «Києво-Могилянська Академія» / 04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра математики / E-mail: o.avramenko@ukma.edu.ua, m.sotnikov@ukma.edu.ua29

ML-ПІДХІД ДО ІДЕНТИФІКАЦІЇ ЕМОЦІЙ У ЗВУКОВИХ ЗАПИСАХ НА БАЗІ MFCC / MACHINE LEARNING APPROACH TO EMOTION IDENTIFICATION BASED ON MFCC / Тимошевський Д.С., студент 2-го курсу магістратури за спеціальністю F1 «Прикладна математика» / Національний університет «Києво-Могилянська Академія» / 04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра математики / E-mail: d.tymoshevskiy@ukma.edu.ua31

РОЗРОБЛЕННЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ МОДЕЛЮВАННЯ ПРОЦЕСУ КРИСТАЛІЗАЦІЇ ЗЛИВКУ З ВИКОРИСТАННЯМ НЕЧІТКОЇ ЛОГІКИ / DEVELOPMENT OF AN INTELLIGENT SYSTEM FOR MODELING THE CRYSTALLIZATION PROCESS OF A MELT USING FUZZY LOGIC / Павлюченко В.О., здобувач вищої освіти другого (магістерського) рівня кафедри «Програмне забезпечення систем»;/ Бабенко М.В., к.т.н., доцент, доцент кафедри «Програмне забезпечення систем» / Дніпровський державний технічний університет / 51918, м. Кам'янське Дніпропетровської обл., вул. Дніпробудівська, 2, факультет комп'ютерних технологій та енергетики, кафедра програмного забезпечення систем / E-mail: -, mvbab@ukr.net33

МАСОВО ПАРАЛЕЛЬНА РЕАЛІЗАЦІЯ MORPHONAS / MASSIVELY PARALLEL MORPHONAS IMPLEMENTATION/ Медвідь С.О., магістр комп'ютерних наук, ст. викладач. / Національний університет «Києво-Могилянська Академія» / 04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики / E-mail: s.medvid@ukma.edu.ua ..34

АДАПТИВНА МУЛЬТИМОДАЛЬНА КОГНІТИВНА СИСТЕМА ДЛЯ АВТОНОМНОГО ПРИЙНЯТТЯ РІШЕНЬ У СКЛАДНИХ ЛОГІСТИЧНИХ СЕРЕДОВИЩАХ / Пасека П.І., аспірант/ Теренчук С.А., професор, кандидат фізико-математичних наук, професор кафедри кафедри інформаційних технологій проєктування та прикладної математики / Київський національний університет будівництва і архітектури / 03037, Київ, проспект Повітряних Сил, 31, факультет автоматизації і інформаційних систем, кафедра інформаційних технологій проєктування та прикладної математики / E-mail: paaaa.paul@gmail.com, terenchuksa@ukr.net ..36

ЕКСПЕРИМЕНТАЛЬНИЙ АНАЛІЗ МОДЕЛІ ЕВОЛЮЦІЙНОЇ ГРИ В ІМЕНУВАННЯ / EXPERIMENTAL ANALYSIS OF THE EVOLUTIONARY NAMING GAME MODEL/ Сітков І.П., аспірант 1 року навчання спеціальності «Інженерія програмного забезпечення» факультету інформатики Національного університету «Києво-Могилянська академія»/ Гулаєва Н.М., доцент, кандидат фізико-математичних наук, доцент кафедри інформатики факультету інформатики Національного університету «Києво-Могилянська академія» / Національний університет «Києво-Могилянська Академія» / 04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики / Університет Севільї / Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad de Sevilla, E.T.S. Ingeniería Informática, Avda. Reina Mercedes s/n, 41012 Sevilla, Spain / E-mail: i.sitkov@ukma.edu.ua, gulayeva@ukma.edu.ua38

GPU-ОРИЄНОВАНА БІДІАГОНАЛІЗАЦІЯ В АЛГОРИТМІ СИНГУЛЯРНОГО РОЗКЛАДУ З ШВИДКОЮ ДЕФЛЯЦІЄЮ ТА ПАРАЛЕЛЬНИМ РОЗБИТТЯМ / GPU-NATIVE BIDIAGONAL SVD WITH OPPORTUNISTIC DEFLATION AND PARALLEL SPLITTING / Сухарський С.С., Ph.D студент / Інститут програмних систем НАН України / Київ, 03187, Україна, Проспект Академіка Глушкова, 40, корп. 5, +38 (044) 526-33-19 / E-mail: serhii.sukharskyi@gmail.com40

КАУЗАЛЬНЕ МОДЕЛЮВАННЯ ПОВЕДІНКИ КОРИСТУВАЧА У ЦИФРОВИХ СЕРВІСАХ / Білінський П.О., аспірант кафедри математики НаУКМА. / Національний університет “Києво-Могилянська Академія” / 04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра математики / E-mail: p.bilinskyi@ukma.edu.ua42

Використання Нейромереж для Автоматичного Поновлення Даних у Спеціалізованих Предметних Областях / Application of Neural Networks for Automatic Data Updating in Specialized Subject Domains / Нагнибіда А.А., аспірант PhD програми «Комп’ютерні науки» 4ого року навчання, факультету інформатики Національного університету «Києво-Могилянська академія» / Національний університет “Києво-Могилянська Академія” / 04655, м. Київ, вулиця Григорія Сковороди, 2, НаУКМА, Факультет інформатики / E-mail: a.nahnybida@ukma.edu.ua..44

ПСЕВДООБЕРНЕНА МАТРИЦЯ МУРА-ПЕНРОУЗА / MOORE-PENROSE PSEUDOINVERSE MATRIX / Кравчук О.М., аспірант 3 р.н. докторської школи НаУКМА, факультет інформатики, кафедра математики/ Крюкова Г.В., к.ф.м.н., доцент, доцент кафедри математики факультету інформатики НаУКМА / Національний університет “Києво-Могилянська Академія” / 04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра математики / E-mail: o.kravchuk@ukma.edu.ua, kriukovagv@ukma.edu.ua45

ГРА З НЕПОВНОЮ ІНФОРМАЦІЄЮ НА ПРИКЛАДІ СПОРТИВНОГО БЕТІНГУ / A GAME WITH INCOMPLETE INFORMATION: THE CASE OF SPORTS BETTING / Куцалаба Н.В., студент 2-го курсу магістратури за спеціальністю 113 «Прикладна математика» / Чорней Р.К., завідувач кафедри, доцент, кандидат фізико-математичних наук / Національний університет “Києво-Могилянська Академія” / 04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра математики / This work was partly supported by a grant from the Simons Foundation International [SFI-PD-Ukraine-00014577, O.G.] E-mail: n.kutsalaba@ukma.edu.ua, r.chornei@ukma.edu.ua47

DECISIONET З ПРОПОРЦІЙНИМ РОЗПОДІЛОМ ОБЧИСЛЮВАЛЬНИХ ПОТУЖНОСТЕЙ / DECISIONET WITH PROPORTIONAL ALLOCATION OF COMPUTATIONAL POWER / Мокрий М.В., студент Ph.D програми «Комп’ютерні науки» факультету інформатики Національного університету «Києво-Могилянська академія»/ Швай Н.О., доцент, кандидат фізико-математичних наук, кафедра математики Національного університету «Києво-Могилянська академія» / Національний університет “Києво-Могилянська Академія” / 04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики / E-mail: m.mokryy@ukma.edu.ua, n.shvay@ukma.edu.ua49

МЕТОДИКА ТРЕНУВАННЯ ГЛИБОКОЇ НЕЙРОННОЇ МЕРЕЖІ DEEPFLAV3 ДЛЯ СЕМАНТИЧНОЇ СЕГМЕНТАЦІЇ ОБ’ЄКТІВ НА СУПУТНИКОВИХ ЗОБРАЖЕННЯХ / Колінко П.В., Аспірант / Національний університет “Києво-Могилянська Академія” / 01054, Київ, вул. Григорія Сковороди, каф. мультимедійних систем, тел. (044) 425-77-23 / E-mail: pavlo.kolinko@ukma.edu.ua51

Проектування інтелектуального агента для автономної оптимізації мікросервісів у корпоративних системах / DESIGN OF AN INTELLIGENT AGENT FOR AUTONOMOUS MICROSERVICE OPTIMIZATION IN ENTERPRISE SYSTEMS / Ванін Д.О., Аспірант, без вченого ступеня / Національний університет «Києво-Могилянська академія» / М. Київ, вул. Алматинська 97/1, кв. 30, 02092, +380959063034 / E-mail: danylo.vanin@ukma.edu.ua53

ВИКОРИСТАННЯ MIXTURE OF EXPERTS З LLM-АГРЕГАТОРОМ ДЛЯ ТОРГОВЕЛЬНОГО БОТА / MIXTURE OF EXPERTS WITH LLM AGGREGATION FOR TRADING BOT / Салата К.В., старший викладач / Національний університет “Києво-

Могилянська Академія” / 04655, м. Київ, вулиця Григорія Сковороди, 2, НаУКМА / E-mail: kirill.salata@ukma.edu.ua54

КОНТРОЛЬ ІНТЕНСИВНОСТІ ЕМОЦІЙ В УКРАЇНСЬКОМУ СИНТЕЗІ МОВЛЕННЯ ЧЕРЕЗ КРОС-МОВНЕ ПЕРЕНОСЕННЯ ЗНАНЬ ТА АДАПТЕРИ LORA / EMOTION INTENSITY CONTROL IN UKRAINIAN TTS THROUGH CROSS-LINGUAL TRANSFER KNOWLEDGE AND LORA ADAPTERS / Івашченко Д.С., аспірант / Національний університет “Києво-Могилянська Академія” / 04655, м. Київ, вулиця Григорія Сковороди, 2, НаУКМА / E-mail: d.ivashchenko@ukma.edu.ua56

ВДОСКОНАЛЕННЯ ВІЗУАЛІЗАЦІЇ МОДЕЛЕЙ АРХІТЕКТУР ПРОГРАМ ДЛЯ ЗМЕНШЕННЯ КОГНІТИВНОГО НАВАНТАЖЕННЯ КОРИСТУВАЧА / IMPROVING VISUALIZATION OF SOFTWARE ARCHITECTURE MODELS TO REDUCE USER COGNITIVE LOAD / Франків О.О., аспірант за спеціальністю «Комп’ютерні науки» / Національний університет «Києво-Могилянська академія» / 04070, м. Київ, вул. Г. Сковороди 2, каф. інформатики, тел. (044) 425 60 64 / E-mail: o.frankiv@ukma.edu.ua58

ПРИСКОРЕНЕ МОДЕЛЮВАННЯ ЕПІДЕМІЙ У ВЕЛИКИХ МЕРЕЖАХ: ОПТИМІЗАЦІЯ МЕТОДУ ГІЛЕСПІ ТА ДВОШАРОВИЙ ПІДХІД / ACCELERATED EPIDEMIC SIMULATION IN LARGE-SCALE NETWORKS: OPTIMIZATION OF THE GILLESPIE ALGORITHM AND A TWO-LAYER APPROACH / Куриляк Ю.А., Аспірант, Асистент кафедри ICM Національного університету Львівська політехніка / Міхаель Т.М. Еммеріх., доктор природничих наук, професор університету Юваскуля, Фінляндія / Professor in Multiobjective Optimization, University of Jyväskylä / Національний університет «Львівська політехніка» / Lviv Polytechnic National University / вулиця Степана Бандери, 12, Львів, Львівська область, 79000, Україна / Університет Юваскуля / University of Jyväskylä / Agora, Mattilanniemi 2, 40100 Jyväskylä, Jyväskylä, 40014, Finland / E-mail: yulian.a.kuryliak@lpnu.ua, michael.t.m.emmerich@jyu.fi61

ЗАСТОСУВАННЯ АЛГОРИТМІВ НАВЧАННЯ Q-МЕРЕЖ ДЛЯ ІТЕРАТИВНОЇ ЗАДАЧІ В’ЯЗНЯ / Ткач Н.В., студент 3 курсу Докторської школи за спеціальністю Комп’ютерні науки, старший викладач кафедри інформатики / Національний університет “Києво-Могилянська Академія” / 04655, м. Київ, вулиця Сковороди, 2, НаУКМА, Факультет інформатики. тел.(044) 426 60 64 / E-mail: n.tkach@ukma.edu.ua63

ЕФЕКТИВНЕ НАВЧАННЯ СТРАТЕГІЙ КЕРУВАННЯ ДЛЯ РОБОТИЗОВАНИХ МАНІПУЛЯЦІЙ ЗА ДОПОМОГОЮ ДИСТИЛЯЦІЇ ЗНАНЬ / EFFICIENT POLICY LEARNING VIA KNOWLEDGE DISTILLATION FOR ROBOTIC MANIPULATION / Севергін О.В., студент 2-го курсу магістратури за спеціальністю F1 «Прикладна математика»/ Кузьменко Д.О., PhD студент, старший викладач/ Швай Н.О., кандидат фізико-математичних наук, старший викладач / Національний університет “Києво-Могилянська Академія” / 04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра математики / E-mail: oleksandr.severhin@ukma.edu.ua, kuzmenko@ukma.edu.ua, n.shvay@ukma.edu.ua64

КООРДИНАЦІЯ МУЛЬТИАГЕНТНИХ LLM-СИСТЕМ / COORDINATION OF MULTI-AGENT LLM-BASED SYSTEMS/ Діхтяр І.Ю., аспірант PhD 1-го року навчання, спеціальність «Комп’ютерні науки» / Нагірна А.М., кандидат фізико-математичних наук, завідувачка кафедри інформатики факультету інформатики НАУКМА / Національний університет «Києво-Могилянська академія» / 04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики / E-mail: i.dikhtiar@ukma.edu.ua, a.nahirna@ukma.edu.ua66

РОЛЬ ФУЗИНГУ ТА СУЧАСНИХ ПІДХОДІВ У ВИЯВЛЕННІ ТА ЕКСПЛУАТАЦІЇ ВРАЗЛИВОСТЕЙ «НУЛЬОВОГО ДНЯ» В УМОВАХ ОБМЕЖЕНИХ РЕСУРСІВ / THE ROLE OF FUSING AND MODERN APPROACHES IN DETECTION AND EXPLOITATION OF ZERO-DAY VULNERABILITIES IN A RESOURCE-CONSTRAINED ENVIRONMENT / Ісмагілов А.І., аспірант / Інститут програмних систем Національної академії наук України / 03187, Київ, просп. Академіка Глушкова, 40 корп. 5, тел.: (044) 526-55-07 / E-mail: ismagilov@ukr.net68

СЕКЦІЯ 2. ПРОГРАМНА ІНЖЕНЕРІЯ

РОЗПІЗНАВАННЯ ЕМОЦІЙНИХ СТАНІВ ПІД ЧАС ПЕРЕГЛЯДУ ВІДЕО НА ОСНОВІ ДАНИХ GSR, ПУЛЬСУ ТА ВІЗУАЛЬНОГО ВІДСТЕЖЕННЯ

Глибовець А.М. доктор технічних наук, професор, декан факультету інформатики Національний університет "Києво-Могилянська академія", Хмель С.М. магістр / Hlybovets A.M., Khmel S.M. / Національний університет «Києво-Могилянська академія» / National University of Kyiv-Mohyla Academy / 04655, м. Київ, вулиця Григорія Сковороди, 2, НаУКМА, ФІ (1 корп, кімн. 302), e-mail: a.glybovets@ukma.edu.ua, svitlana.khmel@ukma.edu.ua70

ПРОБЛЕМИ ФРАГМЕНТАЦІЇ КОНТЕКСТУ В RAG СИСТЕМАХ / PROBLEMS OF CONTEXT FRAGMENTATION IN RAG SYSTEMS

Андрощук М.В. аспірант, Androshchuk M.V. / Національний університет «Києво-Могилянська академія» / National University of Kyiv-Mohyla Academy / 04655, м. Київ, вулиця Григорія Сковороди, 2, НаУКМА, ФІ, e-mail: maxym.androshchuk@ukma.edu.ua71

ТЕХНОЛОГІЇ КОМУНІКАЦІЇ ТРАНСПОРТНИХ ЗАСОБІВ VEHICULAR / COMMUNICATION TECHNOLOGIES

Яблонський В. Р. аспірант, Афонін А.О. кандидат фізико-математичних наук, доцент, / Yablonskyi V.R., Afonin A.O. / Національний університет «Києво-Могилянська академія» / National University of Kyiv-Mohyla Academy / 04655, м. Київ, вулиця Григорія Сковороди, 2, НаУКМА, ФІ, e-mail: afonin@ukma.edu.ua, yablonskyivr@ukma.edu.ua73

ДО ПИТАННЯ УЗАГАЛЬНЕНОГО МЕТАПРОГРАМУВАННЯ В C++ / TOWARDS THE ISSUE OF TEMPLATE METAPROGRAMMING IN C++

Бублик В. В. доцент, кандидат фізико-математичних наук, Трохимчук А.А. студент / Bublyk V.V., Trokhymchuk A.A. / Національний університет «Києво-Могилянська академія» / National University of Kyiv-Mohyla Academy / 04655, м. Київ, вулиця Григорія Сковороди, 2, НаУКМА, ФІ, e-mail: boublik@ukma.edu.ua, a.trokhymchuk@ukma.edu.ua75

DATA STREAMING PIPELINE FOR THE QUADCOPTER FLIGHT CONTROL STACK

Завалій Т.І. аспірант, Шаховська Н.Б. д.т.н, професорка Яцишин В.П. аспірант/ T. Zavalii, N. Shakhovska, V. Iatsyshyn / Lviv Polytechnic National University 79000, Lviv, Kn. Romana str., 5, Department of Artificial Intelligence, e-mail: taras.i.zavalii@lpnu.ua, nataliya.b.shakhovska@lpnu.ua, volodymyr.p.yatsyshyn@lpnu.ua77

ГЕНЕРАТИВНИЙ ФРЕЙМВОРК ДЛЯ ПОБУДОВИ ВІЗУАЛЬНО-ТЕКСТОВИХ ДАТАСЕТІВ НА ОСНОВІ ОНТОЛОГІЙ

Чоловський С.О. аспірант, Здырко В.В. аспірант / Cholovskyi S. Zdyrko V. / Національний університет "Києво-Могилянська Академія" / National University of Kyiv-Mohyla Academy 04655, Київ, вул. Григорія Сковороди, 2, ФІ, кафедра інформатики, e-mail: s.cholovskyi@ukma.edu.ua, v.zdyrko@ukma.edu.ua80

ПОРІВНЯЛЬНИЙ АНАЛІЗ DJANGO І EXPRESS.JS: КРИТЕРІЇ ВИБОРУ ТА СФЕРИ ЗАСТОСУВАННЯ У WEB-РОЗРОБЦІ / COMPARATIVE ANALYSIS OF DJANGO AND EXPRESS.JS: CRITERIA FOR SELECTION AND APPLICATION AREAS IN WEB DEVELOPMENT

Печкурова Е.Я., Печкурова О.М./ Piechkurova E., Pyechkurova O. / Національний університет "Києво-Могилянська Академія" / National University of Kyiv-Mohyla Academy 04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра мультимедійних технологій, e-mail: piechkurova@ukma.edu.ua, pyechkurova@ukma.edu.ua82

РОЗРІЗНЕННЯ БЕЗКОШТОВНИХ ЗАСТОСУНКІВ І ПРОБНИХ ВЕРСІЙ ПЛАТНИХ MACOS-ЗАСТОСУНКІВ НА ПУБЛІЧНИХ МАРКЕТПЛЕЙСАХ / DISTINGUISHING

FREE APPS AND TRIAL VERSIONS OF PAID MACOS APPS ON PUBLIC APP MARKETPLACES

Літвінчук З.В., Франків О.О., Петелев Є.Р., Кривоблоцький С.І., Стулова Н.С. / Litvinchuk Z.V., Frankiv O.O., Peteliev Y.R., Krivoblotsky S.I., Stulova N.S. / Національний університет «Києво-Могилянська академія» / National University of Kyiv-Mohyla Academy 04070, м. Київ, вул. Г. Сковороди 2, каф. інформатики, тел. (044) 425 60 64, e-mail: z.litvinchuk@ukma.edu.ua, o.frankiv@ukma.edu.ua, zhenya.peteliev@macpaw.com, krivoblotsky@macpaw.com, nata.stulova@macpaw.com83

МОДЕЛЮВАННЯ АСИНХРОННИХ ПРОЦЕСІВ ТА УПРАВЛІННЯ СТАНОМ ЗА ДОПОМОГОЮ МЕРЕЖ ПЕТРІ

Давиденко А.М./ Davydenko A.M./ Національний університет «Києво-Могилянська академія» / National University of Kyiv-Mohyla Academy 04655, м. Київ, вулиця Григорія Сковороди, 2, НаУКМА, ФІ, e-mail: andrii.davydenko@ukma.edu.ua85

СЕКЦІЯ 3. ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНІ СИСТЕМИ ТА ТЕХНОЛОГІЇ

ІЄРАРХІЧНА МОДЕЛЬ РОЛЕЙ У CRM-СИСТЕМАХ / HIERARCHICAL ROLE MODEL IN CRM-SYSTEMS

Бабенко Ю.М., доктор філософії / Yuri Babenko, PhD / асистент кафедри кібербезпеки та захисту інформації / Assistant at the Department of Cybersecurity and Information Protection / Київський національний університет імені Тараса Шевченка / Taras Shevchenko National University of Kyiv / 04116, Київ, вул. Богдана Гаврилишина, 24 / 04116, Kyiv, Bohdan Havrylyshyn str., 24, e-mail: yurii.babenko@knu.ua; Бабенко М.В., к.т.н. / Mykhailo Babenko, PhD / доцент кафедри «Програмне забезпечення систем» / Associate Professor of the Department of System Software / Дніпровський державний технічний університет / Dnipro State Technical University / 51918, м. Кам'янське Дніпропетровської обл., вул. Дніпробудівська / 51918, Kamianske, Dnipropetrovsk region, Dniprobudivska str., e-mail: mvbab130973@gmail.com / Демидов М.Ю./ Київський національний університет імені Тараса Шевченка / Taras Shevchenko National University of Kyiv / e-mail: maksimdemidov05@gmail.com88

INFORMATION PRACTICE THROUGH DIGITAL ETHICS: WORLDVIEW ASSESSMENT OF AI RELIABILITY

Irma Bagrationi, Doctor of Philosophy / Assistant Professor, Higher Education Accreditation Expert / Batumi Shota Rustaveli State University / Georgia, Batumi 6010, Ninoshvili st. 32/35, e-mail: irma.bagrationi@bsu.edu.ge; Anri Bagratishvili / MA Student in Philosophy / Batumi Shota Rustaveli State University / Georgia, Batumi 6010, Ninoshvili st. 32/35, e-mail: anri.bagratishvili@gmail.com.90

INFORMATION EDUCATION: DIGITAL KNOWLEDGE AND ONLINE LANGUAGE LEARNING

Дідманідзе І.Ш., доктор інформаційних технологій; Дідманідзе Т.І., / Didmanidze I.S., Doctor of Information Technology; Didmanidze T.I., MA Student / Batumi Shota Rustaveli State University / Georgia, Batumi 6010, Ninoshvili st. 32/35 / Contact phone: (+995) 599 27 66 00, e-mail: ibraim.didmanidze@bsu.edu.ge, didmanidzetengiz@gmail.com92

СИСТЕМА FINGERPRINTING ДЛЯ ВИЯВЛЕННЯ БОТІВ І ШКІДЛИВОЇ АКТИВНОСТІ У ВЕБСИСТЕМАХ / FINGERPRINTING SYSTEM FOR DETECTING BOTS AND MALICIOUS ACTIVITY IN WEB SYSTEMS

Думан А.А., студент; Бабенко Ю.М., доктор філософії / Duman A.A., Student; Babenko Yu.M., PhD / Київський національний університет імені Тараса Шевченка / Taras Shevchenko National University of Kyiv / 04116, Київ, вул. Богдана Гаврилишина, 24 / e-mail: alina.duman@knu.ua, yurii.babenko@knu.ua; Бабенко М.В., к.т.н. / Babenko M.V., PhD / Дніпровський державний технічний університет / Dnipro State Technical University / 51918, м. Кам'янське, вул. Дніпробудівська, 2 / e-mail: mvbab130973@gmail.com94

**КАЛІБРУВАННЯ ПАРИ РОБОТИЗОВАНИХ МАНІПУЛЯТОРІВ У СХЕМІ
«ВЕДУЧА–ВЕДЕНА» / CALIBRATION OF A LEADER–FOLLOWER PAIR OF
ROBOTIC MANIPULATORS**

Волинець Євгеній Анатолійович, кандидат фіз.-мат. наук, ст. викладач кафедри мережних технологій НАУКМА 04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра мережних технологій ye.volynets@ukma.edu.ua96

**АВТОМАТИЧНЕ ФОРМУВАННЯ ОНТОЛОГІЇ ТОВАРІВ НА ОСНОВІ АНАЛІЗУ
ДАНИХ ЕЛЕКТРОННОЇ КОМЕРЦІЇ**

Колесніков А.О., аспірант; Жежерун О.П., к.т.н., доцент / Kolesnikov A.O., PhD Student; Zhezherun O.P., Associate Professor / Національний університет «Киево-Могилянська академія» / National University of Kyiv-Mohyla Academy / 04655, Київ, вул. Григорія Сковороди, 2 / e-mail: anton.kolesnikov@ukma.edu.ua, zhezherun@ukma.edu.ua98

СЕКЦІЯ 1. КОМП'ЮТЕРНІ НАУКИ

ІНТЕЛЕКТУАЛЬНІ ПРОГРАМНІ СИСТЕМИ ПІДТРИМКИ ШІ / INTELLIGENT AI SOFTWARE SUPPORT SYSTEMS

Глибовець М.М. / Glybovets M.

Національний університет "Києво-Могилянська Академія" / National University of Kyiv-Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики
E-mail: glib@ukma.edu.ua

In this study presents the current state and prospects for the development of intelligent AI software support systems. The topic is extremely relevant given the rapid development of both AI algorithms and the infrastructure for their deployment. The report outlines the main modern problems facing researchers and developers in this area: Explainability to AI: integration of symbolic and subsymbolic approaches; autonomy and adaptability of intelligent systems; data collection, management and semantic processing; interoperability and modularity; computing resources and optimization.

За останні кілька років виникла потужна екосистема програмних систем підтримки всього життєвого циклу моделей штучного інтелекту. MLOps-платформи, AutoML-інструменти, системи керування знаннями, мультиагентні середовища стали невід'ємною частиною індустріального впровадження ШІ та повсякденного життя людини.

MLOps та оркестрація ML-конвеєрів

Сучасне застосування різних інтелектуальних програмних систем (ПС) супроводу моделей машинного навчання (МН) у виробничих середовищах вимагає використання стандартів для CI/CD моделей, відстеження експериментів, моніторингу дрейфу та управління артефактами. MLOps та оркестрація ML-конвеєрів орієнтовані на забезпечення оптимізації процесів розробки, розгортання та підтримки моделей МН. Вони інтегрують принципи DevOps, Data Engineering та Machine Learning для автоматизації життєвого циклу моделей і забезпечують просту відтворюваність експериментів та керованість моделями, надійне розгортання, моніторинг і зворотний зв'язок.

Архітектурно ML-конвеєр (ML Pipeline) представляє таку послідовність автоматизованих кроків: збір і очищення даних (Data Collection, Cleaning), підготовка ознак (Feature Engineering), навчання моделі (Model Training), оцінка (Evaluation), розгортання (Deployment), моніторинг (Monitoring). Оркестрація координує виконання цих кроків шляхом забезпечення можливостей: побудови складніших послідовностей і завдань; досить просто визначати залежності між етапами; повторного запуску лише тих частин, які змінилися; масштабування виконання на кластерах або у хмарі.

Наразі виділяють такі фреймворки оркестрації: Kubeflow Pipelines (робота у середовищі Kubernetes та інтеграції з TensorFlow, PyTorch), Apache Airflow (для керування робочими процесами (DAGs)), MLflow (для управління експериментами, моделями та їх розгортанням). Прикладами практичних систем є Google Vertex AI (комплексна платформа MLOps у Google Cloud), Amazon SageMaker (підтримує повний життєвий цикл ML, включно з моніторингом).

Перспективними напрямками розвитку бачиться використання нейро-символічних систем MLOps, вдосконалення автоматичного виявлення та виправлення збоїв, повна інтеграція з DataOps і ModelOps [1].

AutoML та демократизація моделювання

Створення ефективних моделей МН потребує глибоких знань у галузях програмування, статистики, оптимізації та аналітики даних. Тому з'явилася концепція автоматизованого машинного навчання AutoML. Вона зорієнтована на спрощення процесу моделювання для користувачів без спеціальної математичної та технічної підготовки [2].

За допомогою AutoML система самостійно обирає найкращі алгоритми і конфігурації, використовуючи методи баєсівської оптимізації, еволюційних стратегій, мета-навчання та нейро еволюційних підходів. Це зменшує технічний бар'єр входження (автоматизуючи складні етапи побудови моделей), підвищує швидкість розробки аналітичних рішень, забезпечує відтворюваність і контроль якості моделей, підтримує інтеграцію з MLOps-платформами для управління життєвим циклом моделі.

З'явилася значна низка відкритих і комерційних AutoML-платформ. Визнаними фреймворками є Auto-sklearn, H2O.ai, Google Cloud AutoML, Amazon SageMaker Autopilot). Google Cloud AutoML представляє набір сервісів, що автоматизують розробку моделей для

зображень, текстів і табличних даних. Бібліотека з відкритим кодом H2O AutoML підтримує стек моделей та автоматичний ансамбловальний підхід. Python-інструменти Auto-sklearn, TPOT, PyCaret орієнтовані на простоту використання, а Microsoft Azure AutoML та Amazon SageMaker Autopilot на хмарні рішення корпоративного рівня. Ці системи інтегруються з платформами Kubeflow, MLflow, TFX для побудови виробничих ML-конвеєрів, що забезпечує єдиний MLOps-підхід до розробки та експлуатації моделей.

Зрозуміло, що використання AutoML має і низку викликів: зниження інтерпретованості моделей, обмежена гнучкість для специфічних задач ручного налаштування, ризику етичного характеру. Тому розвивається напрям Explainable AutoML (XAutoML), який поєднує автоматизацію моделювання з пояснювальністю результатів.

Перспективи AutoML полягають у подальшій інтеграції з нейро-символічними системами ШІ, великомовними моделями (LLM) та self-optimizing pipelines. Очікується створення «інтелектуальних конструкторів» моделей, які зможуть автоматично пояснювати свої рішення та адаптуватися до нових задач без участі людини.

Таким чином, AutoML є ключовим фактором демократизації ШІ-технологій, який змінює парадигму створення аналітичних систем і відкриває можливості для їх масового використання у різних сферах людської діяльності.

Знання (Knowledge Graphs) і семантична підтримка

Для інтеграції розрізнених джерел даних, пояснюваності та семантичного пошуку дедалі частіше застосовують корпоративні knowledge-graphs (Neo4j) та інші підходи [3].

У сучасному цифровому світі завдання ефективного представлення, інтеграції, пошуку і налагодження семантики даних стають критичними. Тому набуває значного розвитку концепція графів знань (Knowledge Graphs, KG) — структурованих репрезентацій сутностей і відносин між ними, які несуть семантичне навантаження. У Semantic Web такі графи задаються у форматі RDF (Resource Description Framework), OWL-онтологій.

Під семантичною підтримкою розуміється набір технологій та методів, які забезпечують: інтерпретацію (розуміння значення даних), інтеграцію (об'єднання гетерогенних джерел), пошук/запити із семантичним змістом, побудова зв'язків (inference, reasoning) на основі знань, підтримку прийняття рішень з урахуванням контексту, семантики, зв'язків. Вона реалізується додаванням до традиційних даних шарів значення, зв'язків, правил.

Окрім класичного RDF/онтологій, сьогодні активно застосовуються методи представлення embeddings (knowledge graph embeddings) — числові вектори для сутностей/відношень, які дозволяють застосовувати ML-алгоритми. Такий підхід поєднує символічні (граф, онтологія) та субсимвольні (вектори, нейромережі) компоненти. Це забезпечує: покращення інтеграцію даних із різномірних джерел, збільшення семантичної виразності, можливість побудови нових знань.

Але з'являються і нові виклики. “Брудні” дані, неповні метадані, неоднорідні джерела значно ускладнюють створення графів. Великі графи (мільйони сутностей, десятки мільйонів триплів) потребують ефективних сховищ. Оцінювання якості графів (точність, повнота, коректність) залишаються досі недостатньо стандартизованими. Важливою є й завдання динамічності та оновлення.

Перспективним напрямком досліджень тут бачиться інтеграція з великими мовними моделями (LLM) а також бачення графів знань як частини нейро-символічних систем: комбінація символічного знання (граф) + статистичного навчання (нейромережі). Поєднання дає можливість: використовувати символічні правила, логіку, онтології (графи); застосовувати статистичні моделі для аналізу, прогнозу, рекомендації; забезпечити пояснюваність; стимулювати «глибоке» семантичне розуміння — не лише шаблонні відповіді, а врахування зв'язків, контексту.

Графи знань і семантична підтримка виступають як фундаментальні інструменти сучасного інтелектуального аналізу даних та систем штучного інтелекту. Вони дозволяють перейти від даних до значення, від ізольованих записів до зв'язків, від статичних репрезентацій до побудови висновків.

Особливо актуальним є їхній синтез із нейро-символічними підходами: в майбутньому великі мовні моделі, графи знань, логіка, правила — все це має інтегруватися у системи, здатні до глибшого розуміння та пояснення.

Поява «AI-копілотів» та мультиагентних середовищ

AI-копілоти (AI copilots) — це інтерактивні інтелектуальні помічники, які працюють паралельно з людиною, допомагаючи виконувати завдання швидше та ефективніше. Вони можуть бути текстовими, візуальними або голосовими. Копілот розуміє контекст дій користувача та пропонує релевантні підказки. Він використовує історію дій, шаблони та дані для передбачення

наступного кроку - прогнозування. Помічник може автоматизувати рутинні задачі, наприклад, автозаповнення коду (GitHub Copilot), створення презентацій, підготовка звітів. AI-копілот може спілкуватися у реальному часі, уточнювати запити та коригувати свої рекомендації [4]. Вже існують реалізовані системи. AI-помічник для програмування GitHub Copilot пропонує код на основі коментарів і коду користувача. Інтегрований в Word, Excel, PowerPoint Microsoft 365 Copilot допомагає в генерації текстів, таблиць, презентацій. Notion AI структурує інформацію та генерує контент.

Мультиагентні системи (MAS, Multi-Agent Systems) в і інтеграції з AI-копілотами можуть аналізувати дані та пропонувати оптимальні стратегії їх обробки та використання. Поєднання цих технологій відкриває нові горизонти у сфері розумних робочих середовищ, автономних транспортних систем, інтелектуальних міст та складних симуляційних моделей. AI-копілоти підвищують продуктивність і точність виконання завдань, а мультиагентні системи дозволяють ефективно моделювати складні процеси та управляти розподіленими ресурсами.

Висновок

Сучасні інтелектуальні програмні системи підтримки ШІ швидко розвиваються у напрямку комплексного супроводу життєвого циклу моделей. Головними викликами залишаються питання надійності, інтерпретованості та відповідності нормативам. Перспективи нових досліджень полягають у поглибленні інтеграції символічних методів з нейронними мережами, ускладненні мультиагентних систем та автоматизації операційного супроводу моделей.

Список джерел:

1. Singh, S., “Unifying DevOps and MLOps Pipelines Via AI-Driven Observability: A Mixed-Methods Study”, Asian Journal of Research in Computer Science, Vol. 18 (6): 334-342 (2025). <https://doi.org/10.9734/ajrcos/2025/v18i6703>.
2. Yao, J., Zhang, L., & Huang, J. (2025). Evaluation of large language model-driven AutoML in data and model management from human-centered perspective. *Frontiers in Artificial Intelligence*, 8. doi:10.3389/frai.2025.1590105. (https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2025.1590105/full?utm_source=chatgpt.com).
3. Vogt, L., Kuhn, T., & Hoehndorf, R. (2024). Semantic units: organizing knowledge graphs into semantically meaningful units of representation. *Journal of Biomedical Semantics*, 15, 7. <https://doi.org/10.1186/s13326-024-00310-5>.
4. Calderón, C., Martí, P., Jordán, J., Palanca, J., & Julian, V. (2025). Cognitive Agents in Urban Mobility: Integrating LLM Reasoning into Multi-Agent Simulations. *Sensors*, 25(18), 5688. <https://doi.org/10.3390/s25185688>

РОЗШИРЕННЯ КРИТЕРІЮ ТЕСТОВОГО ПОКРИТТЯ ДЛЯ ВИЯВЛЕННЯ СЛАБКОЇ ВАЛІДАЦІЇ ВХІДНИХ ДАНИХ / EXTENDING TEST COVERAGE CRITERION TO REVEAL WEAK INPUT VALIDATION

Колчин О.В., Руденко Д.Б. / Kolchyn O.V., Rudenko D.B.

Інститут кібернетики імені В.М. Глушкова НАН України / V.M. Glushkov Institute of cybernetics
NAS of Ukraine

0380 МСП, просп. Академіка Глушкова, 40

E-mail: kolchin_av@yahoo.com, dasharudenko1902@gmail.com

This research addresses the problem of insufficient validation of external inputs in software systems, which can lead to critical vulnerabilities such as out-of-bounds indexing, null pointer dereferences, division by zero, and similar issues. Our approach combines static data-flow and taint analyses to capture value-dependent data propagation. Input-dependent variables are modeled as taint sources, while critical operations are identified as sinks. We introduce an additional data-flow coverage requirement that distinguishes test cases based on the value ranges observed at sink locations. Early experimental results demonstrate the capability of the proposed technique to detect unsafe propagation paths and to reveal subtle instances of weak or missing input validation.

Значна частина вразливостей у програмному забезпеченні виникає через недостатню перевірку або неправильну обробку вхідних даних [1]. Тому аналіз програмного коду з метою виявлення потенційних вразливостей, пов'язаних із некоректною валідацією, є важливим напрямом досліджень у галузі забезпечення надійності програмного забезпечення [2].

Існує широкий спектр підходів до пошуку вразливостей у програмному коді. Частина з них спрямована на підвищення тестового покриття за рахунок автоматизованого генерування вхідних даних (fuzzing), тоді як інші базуються на статичному або динамічному аналізі програм і виявленні типових шаблонів помилок (pattern-based), таких як «нуль у знаменнику» або «доступ поза межами масиву». Однак не всі дефектні стани можна описати через подібні шаблони, і навіть високий рівень структурного або потокового покриття не гарантує виявлення ситуацій, пов'язаних із логічно слабкою валідацією вхідних даних. Розглянемо приклади коду на рис.1,2.

```

1. int bar(int x, bool v) {
2.     int res = x;
3.     if(v) {
4.         if(x < 1)
5.             res = 1;
6.         if(x > 10)
7.             res = 10;
8.     }
9.     if(res == 10)
10.        report("maximum");
11.    process(res);
12.}

```

Рис. 1. Приклад з покриттям потоку даних

```

1. double foo(int x, bool v, int c) {
2.     double res = 1.0;
3.     if(x > 0 || v == 1) {
4.         if(c > 0)
5.             res = div(c, x);
6.         else
7.             res = 0.0;
8.     }
9.     send(res);
10.}

```

Рис.2. Приклад з замаскованою умовою

Для програми, поданої в першому прикладі, покриття за критеріями all-uses та MC/DC може бути досягнуто за допомогою такого тестового набору вхідних даних: $\{(0, T), (11, T), (10, F)\}$. Для змінної *res* перший тест покриває пару визначення–використання 5—11, другий — пару 7—11, а третій тест охоплює пари 2—9 та 2—11. Варто зазначити, що ці тести забезпечують виклик функції *process* лише для двох значень: 1 (тест № 1) та 10 (тести 2, 3). Решта можливих діапазонів залишається неперевіреними, хоча для їх тестування достатньо використати вхідні значення вигляду (x, F) , де $x \neq 10$. У другому прикладі для забезпечення покриття достатньо таких тестів: $\{(0, F, 0), (1, F, 1), (0, T, 0)\}$. Перший тест покриває пару 2—9 для змінної *res*, другий 5—9, третій 7—9; крім того другий тест покриває пари 1—5 для змінних *x* та *c*. У сукупності ці тести задовольняють вимоги критерію MC/DC. Проте жоден із них не покриває значення змінної *x* при використанні у параметрі виклику функції *div* у діапазоні $x \leq 0$. Такий діапазон потенційно може (якщо функція *div* обчислює ділення) призвести до аварійного завершення через ділення на нуль.

Запропонований підхід базується на статичному аналізі абстрактних формальних моделей програм. Статичний підхід дозволяє досліджувати усі можливі шляхи виконання без необхідності повного динамічного запуску програми, що особливо важливо для виявлення потенційних дефектів на ранніх етапах розроблення. Крім того, формальна модель зберігає інформацію про залежності між значеннями змінних та умовами, що робить можливим точніший аналіз покриття на семантичному рівні. Також можливе символічне моделювання станів програми: замість оперування конкретними значеннями змінних використовуються обмеження на їхні можливі значення. Це дає змогу відстежувати діапазони допустимих значень під час виконання різних гілок програми та визначати області, які залишаються непокритими.

Формальна модель розглядається у вигляді спрямованого ациклічного графа (DAG), доповненого додатковими компонентами — змінними та предикатами. Компонент DAG формально визначається як кортеж (V, E, s, I, S) , де: V множина вершин, що представляють умови або оператори програми, $E \subseteq V \times V$ множина спрямованих ребер, які описують переходи керуючого потоку. $s \in V$ початкова вершина, $I \subseteq V$ непорожня множина вершин, які відповідають зовнішнім даним (taint sources), $S \subseteq V$ непорожня множина вершин, що відповідають використанню у критичних операціях (sinks). **Трасою** називається скінченна послідовність вершин (n_1, \dots, n_k) , $k \geq 2$, така, що для кожного $i=1,2,\dots,k-1$ існує ребро $(n_i, n_{i+1}) \in E$.

Визначення (definition-use pair [3]). Нехай x — змінна, а $q \in V$. Позначимо $\text{defs}(q)$ множину всіх змінних, що визначаються (тобто приймають нове значення) у вершині q . Траса $t = (v, q_1, \dots, q_k, v')$, $k \geq 1$, називається вільною від присвоєнь (definition-clear) від вершини v до вершини v' відносно змінної x якщо $x \notin \text{defs}(q_i)$ для всіх $1 \leq i \leq k$. Пара $(d(x, v), u(x, v'))$ називається парою «визначення-використання» (**def-use pair**), якщо існує траса (v, v_1, \dots, v_n, v') , така що її частина (v_1, \dots, v_n) є вільною від присвоєнь змінній x .

Нижче описане ключове визначення розширеного критерію покриття діапазонами значень:

Визначення (range-sensitive def-use). Нехай $DU=(d(x,v), u(x,v'))$ – пара «визначення-використання». Позначимо через $C=\{C_1, C_2, \dots, C_m\}$ скінченне розбиття області значень змінної x на абстрактні діапазони. Тоді набір тестів T_{DU} задовольняє критерій range-sensitive def-use покриття відносно пари DU якщо для кожного діапазону $C_k \in C$ існує траса $t \in T_{DU}$, така що значення змінної x у вершині v' належить діапазону C_k .

Можливі різні сценарії впровадження запропонованого критерію: за наявності попередньо заданих діапазонів [4] можна спрямувати пошук безпосередньо на досягнення їх покриття (наприклад, методом перевірки моделі). Якщо діапазони заздалегідь не задані, можливе їх автоматичне генерування на льоту під час аналізу простору станів моделі. При цьому окремо виділяються критичні значення (наприклад, некоректність індексації, нуль в знаменнику).

На відміну від критерію all-uses, який орієнтується лише на структурне покриття пар «визначення-використання», запропонований метод додатково вимагає покриття семантично різних діапазонів значень, вдосконалюючи таким чином перевірку валідації вхідних даних. Метод дозволяє виявляти випадки, коли певні діапазони можливих значень залишаються неврахованими тестами, попри формальне досягнення покриття за класичними критеріями. Попередні дані застосування критерію свідчать про ~20% зростання кількості тестів у порівнянні з критерієм all-uses, при цьому спостерігається пропорційне зростання мутаційної оцінки.

1. Common Weakness Enumeration (CWE). 2024 CWE Top 25 Most Dangerous Software Weaknesses [Електронний ресурс]. – Режим доступу: https://cwe.mitre.org/top25/archive/2024/2024_key_list.html
2. Herrera, A., Payer, M., Hosking, A. DatAFLow: Toward a Data-Flow-Guided Fuzzer // ACM Transactions on Software Engineering and Methodology. – 2023. – Vol. 32, No. 5. – Article 132. – 31 p. DOI: 10.1145/3587156
3. Rapps, S., Weyuker, E. Data flow analysis techniques for test data selection // Proceedings of the International Conference on Software Engineering. – 1982. – P. 272–277.
4. Kolchin, A., Potiyenko, S. Extending Data Flow Coverage to Test Constraint Refinements // Lecture Notes in Computer Science, vol. 13274. – 2022. DOI: 10.1007/978-3-031-07727-2_17

КОНЦЕПЦІЯ СЕРВІСУ ОФЛАЙН МАП НА ОСНОВІ CESIUMJS

Ткаченко В.О.

Національний Університет “Кієво-Могилянська академія”

04655, м. Київ, вулиця Григорія Сковороди, 2, НаУКМА, Факультет інформатики (1 корп, кімн. 302).

E-mail: vo.tkachenko@ukma.edu.ua; телефон (063) 704 97 37

This paper presents a concept for enabling offline workflows with geospatial data in web applications built on CesiumJS. Since CesiumJS does not natively support offline use, the proposed service introduces a user-controlled process. The workflow begins with selecting a map region and associated datasets, which are preloaded into local persistent storage. During offline operation, rendering relies solely on cached imagery tiles and 3D Tiles, while user edits such as point additions, route creation, and object modifications are recorded in a structured JSON. Upon reconnection, the changes are transmitted to the backend and synchronized with the database, with conflict resolution applied if objects were modified in parallel. This cycle — preparation, offline work, and synchronization — ensures predictable behavior, preserves data integrity, and extends CesiumJS and 3D Tiles to environments where network connectivity is limited or unavailable.

У цій роботі запропоновано концепцію сервісу, що дає змогу працювати з геопросторовими даними в офлайн-режимі у веб-застосунках на основі CesiumJS. Підхід передбачає попереднє завантаження вибраного регіону мапи до локального сховища, збереження змін і об'єктів доданих під час роботи в режимі офлайн та

подальшу синхронізацію з базою даних після відновлення мережі. Розглянуто архітектурні рішення та процес роботи з даними — від підготовки плиток мапи і 3D-моделей до оброблення конфліктів під час зворотної інтеграції.

CesiumJS — це JavaScript бібліотека з відкритим кодом для візуалізації тривимірних геопросторових даних у веб-браузері.[1] Вона підтримує відображення рельєфу, супутникових знімків, векторних шарів та різних типів 3D-наборів, у тому числі у форматі 3D Tiles. Інструментарій бібліотеки включає API для розробки інтерактивних застосунків, а також навчальні приклади й документацію. У матеріалах проєкту підкреслюється потокова організація роботи з даними та можливість інтеграції з онлайн-сервісами зберігання і доставки, зокрема через платформу Cesium ion, що дозволяє працювати з великими наборами просторового контенту у веб-середовищі.

3D Tiles — відкритий стандарт (OGC Community Standard), призначений для ефективного пошарового стрімінгу та візуалізувати великі тривимірні геопросторові дані, такі як фотограмметрії, 3D-будівель, BIM/CAD, точкових хмар. Специфікація описує ієрархію плиток, формати плиток (b3dm, i3dm, pnts, cmpt) і механізми стилізації, що дозволяє клієнту завантажувати лише релевантні рівні деталізації залежно від масштабу та видимості. Завдяки цьому 3D-контент підвантажується поступово й доступний для інтерактивної роботи у веб-браузері.[2]

Важливо зазначити, що у стандартній конфігурації CesiumJS не має можливості працювати в офлайн-режимі. Усі приклади та документація бібліотеки передбачають доступ до онлайн-сервісів, зокрема до об'єктів рельєфу, плиток мапи та 3D Tiles, які зберігаються на серверах. Для багатьох сфер застосування це не є критичною проблемою, однак у випадках, коли користувач працює у польових умовах або в регіонах із відсутнім підключенням до Інтернету, виникає необхідність у локальному збереженні та подальшій синхронізації даних. Саме це мотивує до розробки сервісу який дозволяє працювати офлайн та дає можливість користувачу керувати процесом збереження та синхронізації.

Сесія починається з явної дії користувача — перемикання у режим підготовки даних. Інтерфейс дозволяє виділити територію і визначити набір потрібних шарів: рельєф, растрові зображення (WMTS), набори 3D Tiles тощо. Далі спрацьовує механізм асинхронного завантаження: обчислюються адреси плиток для всіх можливих масштабів, ініціюються паралельні запити з контролем черги, тайм-аутів і повторних спроб. На етапі валідації перевіряється мінімальна повнота покриття: відсутні елементи позначаються для довантаження або повідомляються користувачу.

Після успішного отримання ресурси зберігаються в локальне сховище. Для зображень це структуроване збереження за ключем схеми $\{z\}/\{x\}/\{y\}$; для 3D Tiles — кешування кореневого маніфесту набору (tileset.json) і піддерев плиток, включно з їхніми двійковими вмістами. Сховище організоване так, щоб забезпечити сталий доступ під час рендерингу (перезапуск браузера чи застосунку (PWA) не призводить до втрати даних). Для роботи з локальними файлами використовується File System API. Цей API дозволяє взаємодіяти з файлами на локальному пристрої користувача або на мережевій файлової системі, доступній користувачеві. Основна функціональність цього API включає читання файлів, запис або збереження файлів, а також доступ до структури каталогів.[3]

Після відключення мережі рендеринг працює винятково з локального джерела. Дані отримуються зі сховища за ключами плиток, 3D-набори відкриваються за локальними URL-посиланнями, а система візуалізації застосовує ті самі механізми рівнів деталізації, але без мережевих викликів. У цей час редакторські дії користувача — додавання точок інтересу, ліній/маршрутів, правки атрибутів об'єктів — не відправляються на сервер, а зберігаються у локальному JSON-файлі. Об'єкти мають ідентифікатор (новий або посилання на існуючий) та час події. Це дозволяє оновити об'єкти у безпечному порядку, коли мережа знову з'явиться.

Повернення в онлайн ініціюється користувачем. Об'єкти з JSON-файлу надсилаються через API. Якщо об'єкт змінювався також на сервері під час офлайн-періоду, спрацьовує політика розв'язання конфліктів: пріоритет за часовою міткою або

за явним вибором користувача. Після успішної інтеграції бекенд повертає остаточні ідентифікатори, локальні файли видаляються і клієнт переходить у звичайний мережевий режим.

Такий цикл дає можливість будувати передбачувану поведінку: користувач керує стартом та завершенням офлайн-сесії; регіон і набори даних завжди відомі; зміни не губляться; синхронізація відбувається чітко визначеними кроками. У веб-інтерфейсі відображається поточний статус та індикація завантаження під час отримання даних.

Розроблений підхід розширює можливості CesiumJS у напрямку автономної роботи, відкриваючи перспективи для застосування у сферах де з'єднання з мережею є слабким, або відсутнє повністю. Ключовим елементом підходу є керований користувачем цикл: вибір регіону карти та набору даних для попереднього завантаження, їх збереження у локальному файловому сховищі, накопичення змін під час офлайн-сесії та синхронізація з базою даних після відновлення доступу до мережі. Такий підхід забезпечує безперервність роботи, цілісність внесених змін і механізм їх інтеграції із серверною частиною. Модель дозволяє ефективно поєднати потокову природу CesiumJS із потребою зберігати дані локально.

Література:

1. CesiumJS Documentation — Learn & API. Доступ: cesium.com/learn/ (офіційні матеріали, API-довідник, приклади).
2. OGC 3D Tiles — Специфікація та огляд. Доступ: docs.ogc.org та репозиторій специфікації.
3. File System API - Concepts and Usage: Доступ: https://developer.mozilla.org/en-US/docs/Web/API/File_System_API (офіційні матеріали, API-довідник, приклади).

SYSTEM AND APPROACH FOR AUTOMATED PHOTO-LIKE FACIAL IMAGE GENERATION FOR USE IN TESTING, TRAINING OR EVALUATION

Bohdan Artiushenko, PhD (Candidate of Technical Sciences),

National University of Kyiv-Mohyla Academy

04655, Kyiv, Hryhorii Skovoroda St 2., Faculty of Informatics, Department of Informatics

E-mail: b.artiushenko@ukma.edu.ua, <https://orcid.org/0009-0007-6107-4205>

This research focuses on systems, apparatuses, and methods for automatically generating photo-like facial image items that may be used on an exam or test. The test may be used for evaluating a test-taker's proficiency in familiar faces recognition, training in facial recognition, prosopagnosia or other similar purpose or goal. A combination of generative artificial intelligence, image and face recognition and manual check is proposed and evaluated.

Prosopagnosia, or face blindness, is widely considered as hardly improbable cognitive disorder, thus training on face recognition is often suggested [2]. It's also widely believed that it's difficult for normal humans to distinguish faces of another nationality.

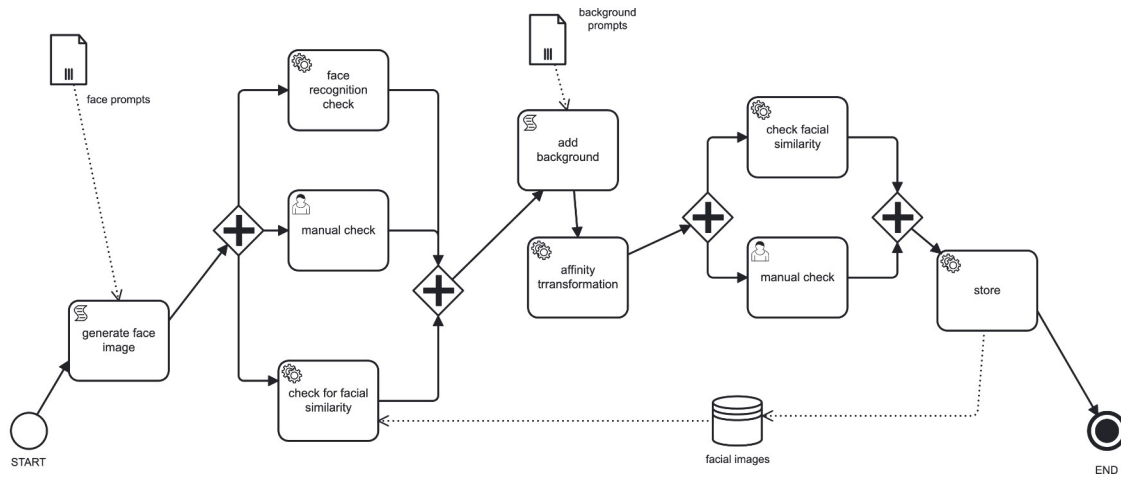


Fig. 1. Facial photo-like image generation BPM diagram

One of possible training methods is to show sequentially human faces of various repeatable people with various backgrounds, colors, light with opportunity for the trainee to mark found repeated faces. A limiting issue in facial recognition by humans is a need for an appropriate facial photo-like library with faces which are not known to the patient. In this paper we suggest an approach and method in general to generate needed photos for the above training faces. Automatically generated tests are widely known, and for example, used by famous Duolingo products[1], and could be adjusted with Generative Artificial Intelligence (for example, Stable Diffusion [3]) for generating image-based test, especially for face recognition skill. When generating photos with AGA new challenges arise, for example there are a lot of generated images with ugly or questionable human faces, in addition we need to guarantee then newly generated phases are not repeated of existing one. Involving operator, even so costly, but could be used to filter out non-human-face, thus filtering repeated faces in a database of thousands of faces by a human is questionable. Conducted testing proves that nearly 95% of generated images are to be filtered out by operator as of bad quality. Due to those limitations, we have added automatic check of generated images and utilize facial recognition systems (OpenCV [4], Dlib etc.). Prompts to generate images both photos and backgrounds are to be systematically reviewed, that leads to need of batch processing. The suggested approach to generate human faces images for training patients is presented in Fig. 1 and consists of steps described on Table 1. As an input for generating images textual prompts, and AGA models are used as an output image with photo on solid background and on various background to be obtained. Also, statistics of good vs bad image is to be collected for prompts adjustment.

Table 1 Step and stages descriptions

No	Step	Description	Technology used	Efficiency
1	Generate face images	Generate images based on stored prompts	Generative AI, like Stability-AI or other	More than 99% prompts generate an image
2	Face recognition check	Test if generated image has a big, uncovered face	OpenCV, Dlib	Rejects 60% of generated images
3,7	Check facial similarity	Test if generated image contains faces same to previously generated	Dlib	Less than 1% of generated faces are found to be same

4,8	Manual check	Manually checking	N/A	Rejects 40% of generated images
5	Add background	Generate new image by changing background of existed one	Generative AI, like Stability-AI or other	More than 99% prompts generate an image, 30% of those have nearly no difference from original
6	Affinity change	Affinity transformation (mirroring, rotations, color shift)	System and numerical tools	Generated images are nearly always good
9	Store	Storing the result to the DataBase	Any Database or file format can be used	N/A

Conclusions: using generative artificial intelligence in combination with facial recognition could be used as a basis for automated photo-like facial image generation for use in testing, training or evaluation. The suggested apparatus was used to generate hundreds of success images, which were picked from up to tens of thousands of generated images total) and could be checked out on web resource: faciemo.web.app.

REFERENCES

1. Attali Y., Runge A., Yancey K., LaFlair G. T. Systems and methods for automated generation of passage-based items for use in testing or evaluation: pat. US 20230080674A1. – Appl. US17/942,626 filed 12 Sep 2022; publ. 16 Mar 2023. – Applicant and assignee: Duolingo, Inc. (US). – 12 p.
2. DeGutis J. M., Chiu C., Grosso M. E., Cohan S. Face processing improvements in prosopagnosia: successes and failures over the last 50 years // *Frontiers in Human Neuroscience*. – 2014. – Vol. 8. – Art. 561. – DOI: 10.3389/fnhum.2014.00561.
3. Generative Models [Software package] / Stability AI. GitHub. Available at: <https://github.com/Stability-AI/generative-models> (accessed 26 Sep 2025).
4. Pulli, K., Baksheev, A., Korniyakov, K., Eruhimov, V. Realtime Computer Vision with OpenCV. *ACM Queue*. 2012. Vol. 10, No 4. P. 40:40–40:56. DOI: 10.1145/2181796.2206309.

ПРО ЗАСТОСУВАННЯ ЦИФРОВИХ ДВІЙНИКІВ В ІНТЕЛЕКТУАЛІЗОВАНИХ МЕРЕЖАХ ПУБЛІЧНИХ СЕРВІСІВ / ON THE APPLICATION OF DIGITAL TWINS IN INTELLECTUALIZED NETWORKS OF PUBLIC SERVICES

Олецький О.В., Моголівський В.О. / Oletsky O.V., Moholivskyi V.O.

Національний університет “Києво-Могилянська Академія” / National University of Kyiv-Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра мультимедійних систем

E-mail: oletsky@ukma.edu.ua, v.moholivskyi@ukma.edu.ua

Issues concerning possible use of digital twins and LLMs within a potential architecture of intellectual agents’ networks for e-governance, e-government, and e-participation systems are discussed. Some approaches for simulating decision-making and estimating their consequences are suggested.

Останнім часом відбувся помітний прогрес в розвитку засобів штучного інтелекту, який пов'язується з появою та розвитком великих мовних моделей (LLM). Вони непогано справляються з задачами, пов'язаними з обчисленнями, аналізом даних, логічним виведенням (хоч і не без проблем). LLM можуть пояснювати хід розв'язку задачі, що наближає їх до систем підтримки прийняття рішень, рекомендаційних систем, експертних систем. Тому вони знаходять все більш широке застосування в різних сферах людської діяльності.

З огляду на це в усьому світі робляться зусилля, спрямовані на інтелектуалізацію публічних сервісів, засобів e-governance, e-government, e-participation. Показовим прикладом у цьому контексті є проект Bürokratt, який реалізується в Естонії [3]. В своїй основі цей проект може бути охарактеризований як такий, що спрямований на побудову мережі чатботів та віртуальних помічників, які надають допомогу кінцевим користувачем у спілкуванні з державними установами. Подібні спроби робляться й у інших країнах; розвиваються методики, спрямовані на оцінювання ступені зрілості впровадження штучного інтелекту в державному управлінні.

При застосуванні LLM суттєво скорочується традиційний технологічний розрив між концепціями та постановками задач, які спочатку існують в словесному вигляді (причому не завжди достатньо добре сформульованому) та їх реалізаціями у вигляді програмних комплексів, графічних інтерфейсів і т.п., оскільки багато пов'язаних з цим питань LLM можуть розв'язувати "з коробки". Якщо казати про мережеву організацію системи інтелектуальних агентів як сервісів, які взаємодіють з урядовими установами, кінцевими користувачами а також між собою, то слід виокремити принаймні три рівні вузлів у таких сервісах: вузли, безпосередньо пов'язані з кінцевими користувачами; вузли-хаби, які акумулюють апробовані експертні знання з окремих питань; вузли, пов'язані з державними органами, що приймають рішення та/або забезпечують публічні сервіси.

У цьому контексті видається перспективним будувати моделі публічних сервісів на основі концепції цифрових двійників [1]. Концепція цифрового двійника – це по суті розвиток концепції імітаційної моделі, але підсилює її на основі більш повноцінного доступу до реальних даних, а також повноцінних засобів вимірювання, контролю та моніторингу. Про цифрові двійники частіше згадується в контексті Internet of Things, Industry 4.0. Але ця концепція може бути застосована і для моделювання систем державного управління та публічних сервісів.

На цій основі можна говорити про побудову формалізованих підходів до моделювання, і в рамках цих моделей розробляти методики вирішення типових задач. Як приклад наведемо можливий підхід до аналізу того, як суспільство оцінює ефективність або неефективність державного адміністрування. В ідеалі управлінські рішення мають прийматися так, щоб здійснювався перехід до ситуацій (станів) з кращими значеннями тих чи інших критеріїв. Математично задача може бути сформульована наступним чином: від поточного стану $x^{(1)}$ зі значеннями критеріїв $(f_1(x^{(1)}), \dots, f_K(x^{(1)}))$ перейти до стану $x^{(2)}$ зі значеннями критеріїв $(f_1(x^{(2)}), \dots, f_K(x^{(2)}))$ так, щоб $f_k(x^{(2)}) \geq f_k(x^{(1)}) \forall i = \overline{1, K}$ (тут K – кількість критеріїв), і хоча б для одного з критеріїв вказана нерівність була строгою. Після ухвалення та реалізації рішень здійснюється оцінка (моніторинг) нової ситуації. У випадку державного управління такий моніторинг може здійснюватися як силами самих управлінських органів, так і з боку суспільства.

Але це складна задача, і її складність зумовлена, зокрема, тим, що критерії оцінювання не обов'язково є кількісними, багато з них носить оціночний та/або розмитий характер; часто ситуація є Парето-оптимальною в тому сенсі, що неможливо покращити значення одних критеріїв, не погіршивши при цьому значення інших критеріїв; часто буває так, що покращення може настати лише в деякій перспективі.

Деякі підходи, пов'язані з організацією імітаційного експерименту щодо прийняття управлінських рішень та оцінювання їх наслідків на основі поєднання моделі «стан-ймовірність дії» з дворівневою схемою методу аналізу ієрархій, описуються в загальних рисах в [5]. При цьому зміна суспільних настроїв на основі оцінки результатів прийнятих рішень моделюється в термінах зміни перехідних ймовірностей між станами (які характеризують міру позитивного або негативного ставлення агента) в рамках марковського ланцюга – можливо, неоднорідного або ж умовно однорідного, в якому перехідні ймовірності розглядаються як випадкові величини. При цьому перспективним видається розвиток підходів на основі порівняння фактично ухваленого рішення з тим, яке було б ухвалено більшістю голосів на загальному референдумі. Проведення референдуму з кожного окремого питання є абсолютно неможливим у реальному житті, але може бути здійснено при моделюванні на основі цифрового двійника.

Перспективним видається поєднання можливостей такого цифрового двійника з можливостями LLM; останні можна використовувати, зокрема, для інтелектуального аналізу звернень громадян; для збору та аналізу постів та коментарів у соціальних мережах тощо.

Видається безумовно доцільною організація взаємодії інтелектуальних агентів та/або їх компонент, принаймні в рамках окремого вузла можливої мережі, на основі мікросервісної архітектури. При цьому слід звернути увагу на підходи до координації мікросервісів на основі машин станів, описані в [4]. Декларативний опис на основі машин станів дозволяє застосовувати математичні методи до аналізу конфігурації мережі та мір важливостей окремих вузлів – зокрема на основі PageRank-подібних методик [4, 2].

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

5. A comprehensive review of Digital Twin from the perspective of total process: data, models, networks and applications [Electronic resource] / Honghai Wu [et al.] // *Sensors*. – 2023. – Vol. 23, no. 19. – P. 8306. – Mode of access: <https://doi.org/10.3390/s23198306> (date of access: 20.10.2025). – Title from screen.

6. Author-level Eigenfactor metrics: evaluating the influence of authors, institutions, and countries within the social science research network community [Electronic resource] / Jevin D. West [et al.] // *Journal of the american society for information science and technology*. – 2013. – Vol. 64, no. 4. – P. 787–801. – Mode of access: <https://doi.org/10.1002/asi.22790> (date of access: 20.10.2025). – Title from screen.

7. Dreyling III R. Technology push in AI-enabled services: how to master technology integration in case of Bürokratt [Electronic resource] / Richard Dreyling III, Tanel Tammet, Ingrid Pappel // *SN Computer science*. – 2024. – Vol. 5, no. 6. – Mode of access: <https://doi.org/10.1007/s42979-024-03064-0> (date of access: 20.10.2025). – Title from screen.

8. Oletsky O. On supervising and coordinating microservices within web applications on the basis of state machines / O. Oletsky, V. Moholivskiy // *CEUR workshop proceedings*. – 2024. – Vol. 3909. – P. 442–454.

9. Oletsky O. Modeling changes of opinions as transition probabilities within one- and two-level model "State-Probability of Action" / O. Oletsky, D. Peleshko, V. Moholivskiy // *CEUR workshop proceedings*. – 2025. – Vol. 4004. – P. 52–61. Richard Dreyling III, Tanel Tammet, Ingrid Pappel. Technology Push in AI-Enabled Services: How to Master Technology Integration in Case of Bürokratt // *SN Computer Science*. – 2024.

NOWCASTING ЯК СУЧАСНИЙ ПІДХІД ДО ОЦІНЮВАННЯ ВВП УКРАЇНИ: ПОРІВНЯННЯ З ТРАДИЦІЙНИМИ МОДЕЛЯМИ ПРОГНОЗУВАННЯ / NOWCASTING AS A MODERN APPROACH TO ESTIMATING UKRAINE'S GDP: COMPARISON WITH TRADITIONAL FORECASTING MODELS

Болотов Є. С., Дрін С. С. / Bolotov Y., Drin S.

Національний університет “Кієво-Могилянська Академія” / National University of Kyiv-Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра математики
E-mail: yehor.bolotov@ukma.edu.ua, svitlana.drin@ukma.edu.ua

This study evaluates advanced forecasting models (MF-VAR and MF-FAVAR) for estimating Ukraine's GDP, which address issues associated with traditional low-frequency data. The MF-FAVAR model stands out for its effectiveness in integrating large, diverse data sets, including regional “short data” and high-frequency indicators such as nighttime lighting, Google Trends, and Twitter activity, into common factors. This approach is vital for short-term forecasting, especially during wartime. In future studies, this model will be applied to forecast the gross regional product (GRP) of the Lviv region.

У сучасному макроекономічному аналізі існує гостра потреба у швидкій оцінці ключових показників, таких як ВВП (Nowcast), до їх офіційного оприлюднення. Ця проблема ускладнюється питанням даних з різною періодичністю. Наприклад, ВВП зазвичай оприлюднюється щоквартально, тоді як інші важливі показники можуть бути доступними щомісяця або навіть щодня. Для України, особливо в умовах війни, використання класичних методів є менш ефективними. Традиційні економічні дані можуть бути недоступними або

ненадійними. Вирішення цієї проблеми вимагає використання спеціалізованих моделей, які можуть ефективно обробляти великі масиви даних з різною частотою та з різних джерел.

Метою даного дослідження є порівняльний аналіз точності прогнозів ВВП України. До таких сучасних підходів до прогнозування на основі моделей змішаної частоти належать векторної авторегресії (MF-VAR) [1] та їх факторних розширень (MF-FAVAR) з традиційними моделями прогнозування. Особливий акцент робиться на їх здатності інтегрувати багаточастотні та нетрадиційні джерела інформації. Це дослідження є особливо актуальним, оскільки, хоча в українській науці вже існують дослідження змішаних частотних моделей на національному рівні (наприклад, U-MIDAS [2]), сучасні економічні підходи, такі як MF-FAVAR, ще не застосовувалися для регіональних прогнозів.

Для досягнення цієї мети було запропоновано використовувати модель MF-FAVAR. Вона підходить для роботи з великими панелями даних та «short data» зі значно коротшими часовими інтервалами. Ця модель інтегрує інформаційний зміст багатьох регіональних предикторів, узагальнених у невелику кількість спільних факторів. MF-FAVAR ефективно вирішує проблему незбалансованих панелей даних («ragged edge»), що виникають через асинхронну публікацію статистичних даних, використовуючи фільтр Калмана. Модель можна описати системою рівнянь, що розділяє змінні за частотою. Динаміка r латентних факторів F_t описується рівнянням стану:

$$F_t = A_1 F_{t-1} + A_2 F_{t-2} + \dots + A_p F_{t-p} + u_t.$$

Рівняння вимірювання пов'язують ці фактори зі спостережуваними факторами:

$$\begin{aligned} X_t^{(H)} &= \Lambda^{(H)} F_t + \eta_t^{(H)}, \\ Y_t^{(L)} &= \Lambda^{(L)} F_t + \eta_t^{(L)}, \end{aligned}$$

де: $X_t^{(H)}$ - об'єднана панель високочастотних офіційних та альтернативних даних; $Y_t^{(L)}$ - низькочастотні показники (ВВП); F_t - вектор латентних факторів (споживчі настрої, стан ринку праці); $\Lambda^{(H)}$, $\Lambda^{(L)}$ - матриця факторних навантажень для високо- та низькочастотних даних; A_i - матриці коефіцієнтів VAR-частини. Кожна дана матриця показує, як минулі значення факторів впливають на поточні значення; $u_t, \eta_t^{(H)}, \eta_t^{(L)}$ - вектор залишків відповідних рівнянь.

Для оцінки точності прогнозування поза вибіркою будуть використовуватися стандартні показники, включаючи середньоквадратичну похибку прогнозу (RMSFE) та симетричну середню абсолютну процентну похибку (SMAPE):

$$\begin{aligned} RMSFE &= \sqrt{\frac{1}{H} \sum_{h=1}^H (\hat{Y}_{T+h|T} - Y_{T+h})^2}, \\ SMAPE &= \frac{100\%}{n} \sum_{t=1}^n \frac{|\hat{y}_t - y_t|}{(|\hat{y}_t| + |y_t|) / 2}. \end{aligned}$$

Ключовим елементом дослідження є унікальний набір даних за 1991-2025 роки, що поєднує дані різної періодичності. Цільовою змінною є ВВП Львівської області, отриманий від Головного управління статистики. Він доповнюється офіційними даними середньої частоти, включаючи економічні показники (промисловість, торгівля, сільське господарство) з порталу статистики Львівської області та дані про ринок праці від Львівської обласної служби зайнятості. Модель також інтегруватиме високочастотні альтернативні дані: ділова активність (реєстрація/ліквідація приватних підприємств та юридичних осіб) від Opendatobot, ціни на ринку нерухомості, а також пошукова активність та настрої від Google Trends та Twitter. Окремим викликом для дослідження є інтерпретація даних про нічне освітлення (NLI) [3], оскільки цей показник, ймовірно, відображав не економічну активність, а фактори обмежень безпеки (відключення електроенергії) та дефіциту енергії (відключення електроенергії), що вимагає спеціальної обробки в моделі.

Попередні результати показали, що використання факторних моделей на основі великих наборів даних має переваги над простими авторегресійними (AR) моделями. Щодо нетрадиційних даних, поєднання декількох корелятивів значно покращує точність відстеження економічної активності. Моделі, що використовують лише один показник, працюють гірше. В емпіричному дослідженні регіональної економічної активності в Україні комбінована модель, що використовує ці нетрадиційні дані, показала надзвичайно високу пояснювальну силу (R^2 у вибірці 99,60%) [3]. Водночас такий високий R^2 вимагає обережної інтерпретації через потенційний ризик перенавчання. Слід зазначити, що хоча динамічні факторні моделі мають

теоретичні переваги, емпіричні порівняння часто показують, що вони не завжди систематично перевершують надійні статичні факторні моделі (Stock & Watson) [4], які зазвичай реалізуються за допомогою аналізу головних компонент (PCA), з точки зору точності прогнозування.

Отримані результати вказують на можливість істотного поліпшення якості короткострокових економічних прогнозів (Nowcasting) для ВВП України та її регіонів. Моделі MF-VAR/MF-FAVAR мають фундаментальні переваги над традиційними моделями, оскільки вони здатні обробляти змішані частоти даних і великі незбалансовані набори даних за допомогою ефективного використання спільних факторів. Це особливо важливо для регіонального прогнозування, де для отримання своєчасних оцінок необхідна інтеграція неоднорідних і високочастотних показників.

Список джерел:

- [1] Koop G., McIntyre S., Mitchell J., Poon A., Wu P. Incorporating Short Data into Large Mixed-Frequency VARs for Regional Nowcasting // Federal Reserve Bank of Cleveland Working Paper Series. – 2023. – No. 23-09
- [2] Зомчак Л. М., Ракова А. С. U-MIDAS-модель прогнозування ВВП України на даних змішаної частоти // Східна Європа: Економіка, Бізнес та Управління. – 2020. – Вип. 2 (25). – С. 436–441.
- [3] Constantinescu M., Karpner K., Szumilo N. The Warcast Index: Estimating Economic Activity without Official Data during the Ukraine War in 2022 // NBU Working Papers. – 2024. – № 03/2024.
- [4] Stock J. H., Watson M. W. Forecasting Using Principal Components From a Large Number of Predictors // Journal of the American Statistical Association. – 2002. – Vol. 97, № 460. – P. 1167–1179.

ФРАКТАЛЬНО-ДИFUЗІЙНІ ГЕНЕРАТИВНІ МОДЕЛІ: ІЄРАРХІЧНИЙ ПІДХІД ДО СИНТЕЗУ ЗОБРАЖЕНЬ / FRACTAL-DIFFUSION GENERATIVE MODELS: A HIERARCHICAL APPROACH TO IMAGE SYNTHESIS

Шалімов А. В., Авраменко О. В. / Shalimov A., Avramenko O.

Національний університет “Києво-Могилянська Академія” / National University of Kyiv-Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра математики
E-mail: a.shalimov@ukma.edu.ua, o.avramenko@ukma.edu.ua

This study extends the Fractal Generative Model by replacing its autoregressive generator with a diffusion-based generator at the top level of the hierarchy. FractalGen builds images recursively: each level invokes a similar generative module on subregions, enforcing self-similar structure across scales. In the original model, this module is autoregressive, generating pixels sequentially and enabling exact likelihoods. We introduce a new generator type that uses a conditional diffusion U-Net to produce a coarse global canvas in parallel via denoising. Lower levels of the hierarchy remain unchanged and further refine local detail. This hybrid fractal-diffusion architecture preserves the recursive, self-similar assembly of FractalGen while replacing pixel-by-pixel sampling with parallel diffusion at the latent level. We evaluate a baseline DDPM against our approach on ImageNet using visual-fidelity and distributional metrics.

Сучасні методи генерації зображень ґрунтуються на масштабних авторегресійних і дифузійних моделях, здатних створювати реалістичні та семантично узгоджені зображення. Дифузійні підходи [1] формують зображення шляхом ітеративного знешумлення латентного вектора, забезпечуючи високу якість, але потребуючи значних обчислень. Авторегресійні моделі факторизують розподіл за регіонами, що забезпечує узгодженість між підобластями, проте залишаються повільними через послідовну природу генерації. Для подолання цих обмежень запропоновано фрактальні генеративні моделі (Fractal Generative Models, FGM) [2], які рекурсивно застосовують один і той самий генератор на різних масштабах, формуючи самоподібну ієрархію, що концептуально узгоджується з фрактальною природою [3]. У початковій реалізації FGM використовували авторегресійні модулі, які забезпечували узгодженість між підобластями, але зберігали низьку швидкість через послідовну генерацію.

У цьому дослідженні пропонується модифікація FGM із використанням дифузійної нейронної мережі на архітектурі U-Net як альтернативного генератора. Такий підхід підвищує паралельність процесу, скорочує час синтезу й інтегрує ієрархічний самоподібний принцип у стандартну дифузійну модель. Конфігурація працює в латентному просторі [4], що зменшує ймовірність артефактів, спричинених відсутністю явного авторегресійного обумовлення. Верхній рівень ініціалізує гаусівський шум, який знешумлюють дочірні вузли, що діють на

маскованих підобластях із контекстним урахуванням глобального рівня [5]. Узгодженість між підобластями забезпечується спільною ініціалізацією шуму та контекстним навчанням.

Порівняння роботи моделі проводимо з використанням датасетів MNIST та ImageNet у роздільності 64×64 . Для оцінки якості генерації застосовуємо метрику FID (Fréchet Inception Distance), яка відповідає за візуальну якість зображень. Також виконано експерименти з варіюванням глибини фрактала, що відповідає за збільшення рівнів ієрархії та кількості дочірніх вузлів.



Рис. 1 Згенеровані зображення при $N = 50$.



Рис. 2 Згенеровані зображення при $N=20$.

Під час експериментів метрику FID оцінювали для різних глибин фрактала $N \in \{20, 30, 50, 80, 100, 150\}$. Результати показали немонотонну залежність якості від глибини: FID зменшується від $81.2 (N=20)$ до $72.4 (N=50)$, після чого знову зростає до ≈ 78 при $N \geq 100$. Така поведінка свідчить про наявність оптимальної кількості рекурсивних рівнів, за якої досягається баланс між глобальною структурою та локальною деталізацією. Як видно з рис. 1–2, глибші ієрархії ($N=50$) формують узгоджену глобальну композицію та плавні переходи між підобластями, тоді як менша глибина ($N=20$) спричиняє локальні артефакти й втрату семантичної цілісності сцени.

Отримані результати підтверджують, що рекурсивний фреймворк, описаний у роботі [2], може ефективно поєднуватися з дифузійними модулями без послідовної авторегресії. Збережено ієрархічну композицію сцени та покращено можливість паралельної генерації підрегіонів. Наступні кроки включають теоретичний аналіз ініціалізації первинного шуму в латентному просторі та дослідження взаємозв'язку між спектральними й візуальними характеристиками згенерованих зображень залежно від глибини фрактала.

У перспективі такий фрактально-дифузійний підхід може бути поширений на інші типи даних, зокрема тривимірні сцени, відеопослідовності та мультимодальні моделі типу “текст \rightarrow зображення”. Завдяки ієрархічній структурі, подібні архітектури потенційно здатні поєднувати глобальну семантичну цілісність з локальною адаптивністю, що відкриває можливості для побудови більш універсальних генеративних систем нового покоління.

5. Ho, J., Jain, A., Abbeel, P. Denoising Diffusion Probabilistic Models. *Advances in Neural Information Processing Systems*. 2020. Vol. 33. P. 6840–6851.
6. Li, T., Sun, Q., Fan, L., He, K. Fractal Generative Models. *arXiv preprint arXiv:2502.17437*. 2025.
7. Mandelbrot, B. B. *The Fractal Geometry of Nature*. Revised and enlarged edition. New York: W. H. Freeman, 1983.
8. Rombach, R., Blattmann, A., Lorenz, D., Esser, P. High-Resolution Image Synthesis with Latent Diffusion Models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022. P. 10684–10695.
9. Lugmayr, A., Danelljan, M., Romero, A., Yu, F., Timofte, R., Van Gool, L. RePaint: Inpainting Using Denoising Diffusion Probabilistic Models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022. P. 11461–11471.

**ІНТЕГРАЦІЯ МЕТОДУ АНАЛІЗУ ІЄРАХІЙ ІЗ НЕЧІТКОЮ ЛОГІКОЮ ДЛЯ
ВИРІШЕННЯ БАГАТОКРИТЕРІАЛЬНИХ ЗАДАЧ / INTEGRATION OF THE
ANALYTICAL HIERARCHY METHOD WITH FUZZY LOGIC FOR SOLVING MULTI-
CRITERION PROBLEMS**

Тригуб О.С. / Tryhub O.S.

Національний університет «Києво-Могилянська Академія» / National University of Kyiv-Mohyla
Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики

E-mail: oleksandr.tryhub@ukma.edu.ua

The decision-making process is often accompanied by uncertainty and subjectivity of expert assessments. The Analytical Hierarchy Process (AHP) method is one of the most widely used methods of multi-criteria choice. However, the classical AHP is based on fairly clear numerical estimates according to the proposed scale [1], which does not always adequately reflect the real judgments of experts. There are known situations when, when conducting the procedure of pairwise comparisons of the importance of the elements of the problem, the decision-maker (DM, expert) has difficulties with choosing such a clear numerical estimate. In such situations, the use of the AHP method with elements of fuzzy logic (Fuzzy Logic AHP, FAHP) is quite promising, which allows better taking into account such fuzziness in pairwise comparisons [2,3]. The results of solving the model problem demonstrate that the use of elements of fuzzy logic can lead to a change in the leader among the alternatives.

Розглянемо модельну задачу вибору найкращої альтернативи програмного забезпечення (ПЗ) для аналітики даних серед трьох альтернатив: А1 (Програма X), А2 (Програма Y), А3 (Програма Z). Оцінювання здійснюється за трьома критеріями: К1 — функціональність, К2 — зручність використання, К3 — вартість програмного продукту.

Розв'яжемо задачу спочатку класичним методом АНР. Обчислення будемо здійснювати за допомогою програмної системи дослідження слабоструктурованих задач багатокритеріальної оптимізації [4]. Оцінки експерта щодо парних порівнянь критеріїв відносно їх впливу на вибір ПЗ наведені у матриці:

$$\begin{pmatrix} \text{ПЗ} & \text{К1} & \text{К2} & \text{К3} \\ \text{К1} & 1 & 2 & 3 \\ \text{К2} & 0.5 & 1 & 3 \\ \text{К3} & 0.333 & 0.333 & 1 \end{pmatrix}$$

Результати парних порівнянь експертом альтернатив за кожним з трьох критеріїв К1, К2, К3 представлені у наступних трьох матрицях:

$$\begin{pmatrix} \text{К1} & \text{X} & \text{Y} & \text{Z} \\ \text{X} & 1 & 3 & 5 \\ \text{Y} & 0.333 & 1 & 2 \\ \text{Z} & 0.200 & 0.500 & 1 \end{pmatrix} \quad \begin{pmatrix} \text{К2} & \text{X} & \text{Y} & \text{Z} \\ \text{X} & 1 & 0.500 & 3 \\ \text{Y} & 2 & 1 & 4 \\ \text{Z} & 0.333 & 0.250 & 1 \end{pmatrix} \quad \begin{pmatrix} \text{К3} & \text{X} & \text{Y} & \text{Z} \\ \text{X} & 1 & 2 & 0.500 \\ \text{Y} & 0.500 & 1 & 0.333 \\ \text{Z} & 2 & 3 & 1 \end{pmatrix}$$

Обчислені величини узгодженостей для наведених матриць ПЗ, К1, К2 та К3 становлять 4,62%, 0,32%, 1,6% та 0,79% відповідно, що свідчить про коректність процедури парних порівнянь.

Нормалізовані вектори локальних пріоритетів, обчислені для кожної матриці становлять:

$$V_{\text{пз}} = (0.528; 0.333; 0.139), V_{\text{К1}} = (0.648; 0.230; 0.122), \\ V_{\text{К2}} = (0.320; 0.558; 0.122), V_{\text{К3}} = (0.297; 0.163; 0.140).$$

Вектор глобальних пріоритетів, обчислений методом лінійної згортки призвів до результату: $V_{\text{глоб}} = (0.490; 0.330; 0.180)$, тобто лідером серед альтернатив є А1 (Програма X).

Розв'яжемо тепер цю задачу методом FАHP, з використанням елементів нечіткої логіки. Для коректності порівняння двох підходів залишимо матрицю вагомостей критеріїв відносно їх впливу на придбання ПЗ без змін, а при проведенні парних порівнянь альтернатив по кожному критерію будемо використовувати нечіткі трикутні числа вигляду (L, M, U) з подальшою їх

дефазифікацією у вигляді $(L+M+U)/3$. Тут M – чітке значення нечіткого трикутного числа, L та U – ліва та права межі, які визначають ступінь «розмитості» числа [3]. Наведемо судження експерта щодо парних порівнянь альтернатив за кожним критерієм:

Пари альтернатив	Критерій K1	Критерій K2	Критерій K3
(A_1, A_2)	$(1.0; 1.1; 1.2)$ дефазифікація = 1.1	$(0.5; 0.6; 0.7)$ дефазифікація = 0.6	$(1.0; 1.1; 1.2)$ дефазифікація = 1.1
(A_1, A_3)	$(1.2; 1.3; 1.4)$ дефазифікація = 1.3	$(1.0; 1.1; 1.2)$ дефазифікація = 1.1	$(0.5; 0.7; 0.9)$ дефазифікація = 0.7
(A_2, A_3)	$(1.0; 1.2; 1.4)$ дефазифікація = 1.2	$(1.1; 1.3; 1.5)$ дефазифікація = 1.3	$(0.5; 0.65; 0.8)$ дефазифікація = 0.65

Такі дані зберігають всі відношення між альтернативами в такому ж вигляді, як і при розв'язанні задачі класичним методом аналізу ієрархій АНР, результати якого наведені вище. Таким чином, для методу ФАНР матриці парних порівнянь альтернатив за кожним критерієм матимуть вигляд:

$$\begin{pmatrix} K1 & X & Y & Z \\ X & 1 & 1.1 & 1.3 \\ Y & 0.909 & 1 & 1.2 \\ Z & 0.769 & 0.833 & 1 \end{pmatrix} \quad \begin{pmatrix} K2 & X & Y & Z \\ X & 1 & 0.6 & 1.1 \\ Y & 1.667 & 1 & 1.3 \\ Z & 0.909 & 0.769 & 1 \end{pmatrix}$$

$$\begin{pmatrix} K3 & X & Y & Z \\ X & 1 & 1.1 & 0.7 \\ Y & 0.909 & 1 & 0.650 \\ Z & 1.429 & 1.538 & 1 \end{pmatrix}$$

Величини узгодженостей матриць K1, K2 та K3 становлять 0%, 1,1% та 0% відповідно. Обчислені нормалізовані вектори локальних пріоритетів набувають вигляду:

$$V_{K1} = (0.373; 0.341; 0.286), V_{K2} = (0.285; 0.424; 0.291), V_{K3} = (0.300; 0.275; 0.425).$$

За цими даними був обчислений методом лінійної згортки вектор глобальних пріоритетів:

$V_{глоб} = (0.334; 0.359; 0.307)$, тобто лідером серед альтернатив тепер вже є A2 (Програма Y). Результати розрахунків глобальних пріоритетів за обома методами запишемо у вигляді порівняльної таблиці глобальних пріоритетів альтернатив:

Альтернатива	Класичний АНР	Нечіткий АНР
A1	0.490	0.334
A2	0.330	0.359
A3	0.180	0.307

Таким чином, використання методу з елементами нечіткої логіки ФАНР може призводити до іншого розв'язку задачі багатокритеріальної оптимізації, ніж при використанні класичного методу аналізу ієрархій АНР. В певних задачах багатокритеріального вибору метод ФАНР дозволяє краще враховувати невизначеність у судженнях експертів, ніж метод АНР, що може суттєво впливати на прийняття остаточного рішення.

Список джерел

- Saaty T.L. The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation. New York: McGraw-Hill, 1980. 287 p.
- Желдак Т. А., Коряшкіна Л. С., Ус С. А. Нечіткі множини в системах управління та прийняття рішень: навч. посібник. Дніпро: НТУ «ДП», 2020. 386 с.
- Коротка Л. І., Зеленцов Д. Г., Науменко Н. Ю., Ляшенко О. А., Солодка Н. О. Обчислювальний інтелект: теорія нечітких множин: навч. посібник. Дніпро: ДВНЗ УДХТУ, 2020. 161 с.

13. Trygub R., Trygub O., Gorborukov V. Researching semistructured problems of multicriteria optimization using the software system // Scientific Notes of National University of Kyiv-Mohyla Academy. Computer Science. Vol. 151, 2013, P. 79-88.

ТОПОЛОГІЧНА СКЛАДНІСТЬ ВУЛИЧНИХ МЕРЕЖ УКРАЇНСЬКИХ МІСТ: ЕНТРОПІЙНИЙ АНАЛІЗ

Смиш О.Р.

Національний університет «Києво-Могилянська академія»
вул. Сковороди 2, м. Київ, 04070, Україна, o.smysh@ukma.edu.ua

This study presents a comprehensive entropy-based data analysis of street network configurations across all 463 Ukrainian cities, using graph-theoretic methods and information entropy measures. The research quantifies the topological complexity of urban network systems through the computation of two entropy measures: H_o and φ . Analysis of OpenStreetMap data reveals significant regional differentiation in entropy characteristics across Ukrainian cities, with eastern regions demonstrating higher φ values, while western regions exhibit the inverse pattern. The results indicate a dependency between city size and structural organization in Ukrainian urban systems, confirming a correlation between total network length and the φ parameter. The findings reveal that maintaining regular grid-like structures becomes increasingly complex for Ukrainian cities, particularly when considering historical development patterns and topographical constraints that have shaped urban morphology.

Ентропійні характеристики урбаністичних мережевих структур є формальними показниками організаційної складності та функціональної узгодженості міського середовища в термінах аналізування графів. Використання методів оброблення даних та алгоритмів мережевого аналізування для дослідження топологічних параметрів дорожньої інфраструктури уможливило вимірювання ступеня впорядкованості міських систем і виявлення стабільних закономірностей їхнього просторового розвитку. Стаття Geoff Boeing [1] демонструє ефективність запропонованих ентропійних мір для кількісного оцінювання просторового порядку в міських мережах, ілюструючи значні відмінності між містами різних регіонів світу.

МЕТОДОЛОГІЯ

Створено програму для автоматичного аналізування всього списку офіційних міст України (463 міста станом на 2025 рік), розподілених на 27 адміністративних регіонах, охоплюючи Автономну Республіку Крим та міста Київ і Севастополь зі спеціальним статусом. Завдяки наявним відкритим даним із ресурсу OpenStreetMap [2] для кожного міста обчислено його мережеві параметри.

У дослідженні використано H_o як основну ентропійну міру топологічної складності за формулою Шеннона:

$$H_o = - \sum_{i=1}^n P(o_i) \log_e P(o_i),$$

де n — загальна кількість бінів, i — індексує біни, а $P(o_i)$ — пропорція орієнтацій, що потрапляють у i -й бін.

Додатковою ентропійною характеристикою мережевої структури є значення φ — індикатор орієнтаційного порядку, що кількісно визначає, наскільки мережа міста відповідає геометричній логіці впорядкування єдиної сітки [1]:

$$\varphi = 1 - \left(\frac{H_o - H_g}{H_{max} - H_g} \right)^2,$$

де $H_{max} = \ln(36) = 3,584$ нат — максимальна ентропія для 36 бінів орієнтацій при рівномірному розподілі, $H_g = \ln(4) = 1,386$ нат — ідеалізована ентропія ідеальної чотириосьової сітки з рівними пропорціями вулиць. Значення $\varphi = 0$ відповідає повністю неупорядкованій структурі, а $\varphi = 1$ — ідеальній сітці.

РЕЗУЛЬТАТИ

Аналіз ентропійних характеристик показав, що міра H_o українських міст варіюється в діапазоні від 3,571 до 1,855 з середнім значенням 3,289, що свідчить про значну варіабельність топологічної складності між різними містами України. Ентропійна міра φ демонструє діапазон від 0,011 (м. Харків) до 0,954 (м. Українськ, Донецька обл.) із середнім значенням 0,231, що вказує на низьку загальну впорядкованість українських міст.

Регіональний аналіз, що показано на рисунку 1, виявив, що найвищі значення φ має Донецька область. Серед областей найнижче середнє значення показника φ міст демонструє Івано-Франківська область.

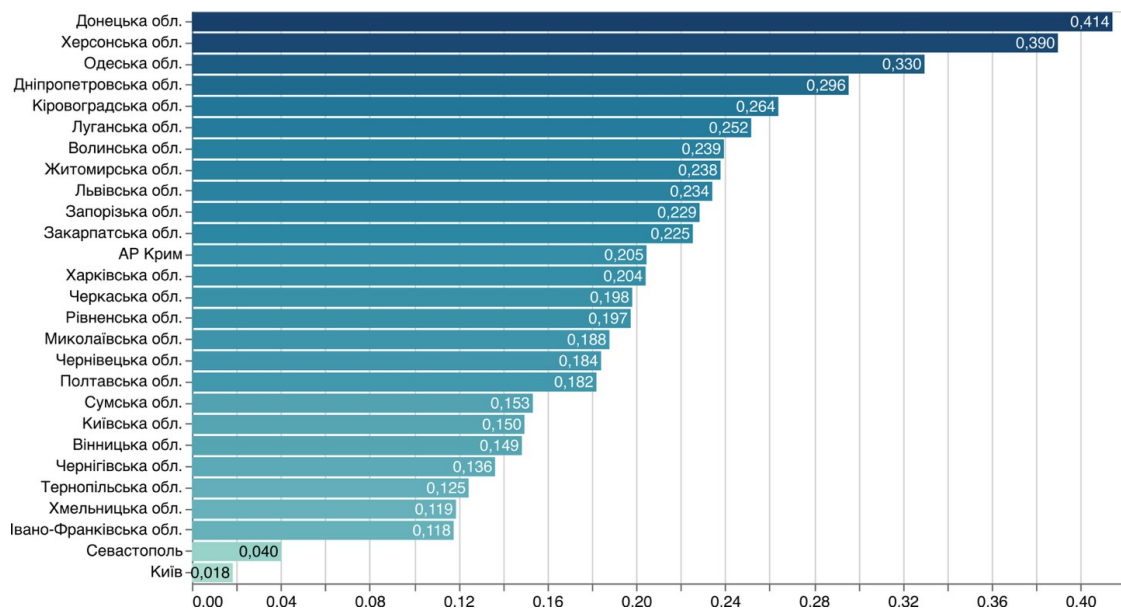


Рисунок 1. Значення φ для регіонів України

ВИСНОВКИ

Регіональні особливості фіксують чітку географічну диференціацію ентропійних характеристик. Східні регіони, зокрема Донецька область, характеризуються вищими значеннями φ , що може відобразити історично сформовану промислову специфіку містобудування з переважанням лінійних планувальних схем. Західні регіони, особливо Івано-Франківська, Хмельницька та Тернопільська області, мають низькі показники φ , що свідчить про менш структуровану мережеву організацію.

Застосування ентропійних мір H_o та φ демонструє дієвість для кількісного оцінювання топологічної складності урбаністичних мережевих систем, підтверджуючи методологічну валідність підходів для проведення аналізу структурних властивостей українських міських агломерацій. Дослідження виявило статистично значущі відмінності в ентропійних характеристиках між регіонами України, що відображає історично сформовані особливості урбаністичного розвитку та регіональну специфіку планувальних рішень.

Отримані результати про всі українські міста можливо застосовувати для оцінювання міського планування, оптимізації транспортних мереж та прогнозування

розвитку урбаністичних систем, що надає практичний інструментарій для містобудівної політики та стратегічного територіального розвитку, особливо в повоєнний період відбудови міст.

Дані про дорожню інфраструктуру для цього дослідження отримано з ресурсу OpenStreetMap [2], що є відкритим картографічним проєктом, який підтримується спільнотою добровільних маперів, котрі самостійно вносять і оновлюють інформацію про дороги. Оскільки редагування даних є відкритим для всіх користувачів, якість і повнота картографічного покриття можуть істотно відрізнятися від реальної ситуації. Через це результати аналізу варто тлумачити з належною обережністю, передусім для міст, що перебувають під окупацією або зазнали значних руйнувань унаслідок воєнних дій, де актуальність та правдивість даних порушено.

ЛІТЕРАТУРА

1. Boeing G. Urban spatial order: street network orientation, configuration, and entropy. *Applied Network Science*. 2019. Vol. 4, no. 1. URL: <https://doi.org/10.1007/s41109-019-0189-1> (date of access: 27.08.2025).
2. A Review of OpenStreetMap Data. *Mapping and the Citizen Sensor*. 2017. P. 37–59. URL: <https://doi.org/10.5334/bbf.c> (date of access: 27.08.2025).

КЛЮЧОВІ СЛОВА: аналіз даних, геоінформаційні технології, ентропія, теорія графів, вуличні мережі, мережевий аналіз, топологічний аналіз, топологія, обчислювальна урбаністика, урбаністичні системи.

PINN MODELING OF INTERFACIAL GRAVITY–CAPILLARY WAVES / МОДЕЛЮВАННЯ ВНУТРІШНІХ ГРАВІТАЦІЙНО-КАПІЛЯРНИХ ХВИЛЬ ЗА ДОПОМОГОЮ PINN

Avramenko O., Sotnikov M. / Авраменко О., Сотніков М.

National University of Kyiv-Mohyla Academy / Національний університет “Києво-Могилянська Академія”

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра математики

E-mail: o.avramenko@ukma.edu.ua, m.sotnikov@ukma.edu.ua

This paper presents an automated computational framework for modeling hydrodynamic processes using physics-informed neural networks (PINNs). The modular system integrates all stages of numerical experimentation — from data generation and model training to validation and accuracy evaluation — ensuring reproducibility, flexibility, and scalability. The framework was verified on the classical problem of interfacial gravity–capillary waves between two incompressible fluids, using the analytical solution as a benchmark for numerical assessment. Computational experiments showed that increasing the number of training points from 400 to 1000 improved accuracy and convergence, with the *Extended* configuration achieving 98.86% accuracy and a MAPE of 1.14%, while *Adaptive_LR* remained stable. The results confirm the reliability and efficiency of the proposed PINN-based framework for solving complex hydrodynamic problems governed by nonlinear partial differential equations.

Physics-informed neural networks (PINNs) are a powerful tool for solving forward and inverse problems governed by nonlinear partial differential equations. The foundational framework introduced by Raissi, Perdikaris, and Karniadakis [1] demonstrated how deep neural networks can incorporate physical laws through equation residuals. Later improvements by Lin and Chen [2] enhanced accuracy for localized wave phenomena, while Ren et al. [3] extended the method to large-scale geophysical simulations. These developments established the basis for applying PINNs to complex hydrodynamic systems, including interfacial gravity–capillary waves used here as a verification benchmark. The developed software complex offers an integrated, modular environment for constructing, training, and analyzing PINNs, automating all stages of computation — from data generation and model training to accuracy evaluation and visualization — ensuring reproducibility and scalability of numerical experiments.

The core of the software complex is a configuration management system that defines standardized neural network architectures and training parameters, including the number of layers, neurons, epochs, and learning rates. This design enables quick reconfiguration of experiments and systematic exploration of the parameter space, allowing researchers to efficiently test various PINN setups. The training workflow integrates all stages of numerical experimentation — automated generation of analytical data, model training, and validation against reference solutions — as well as batch execution of multiple experiments with detailed logging of timing, errors, and performance metrics.

Data management and monitoring components ensure deterministic organization of results, automatically collect metadata from completed runs, and generate summary tables with quantitative indicators such as MAE, RMSE, MAPE, and R^2 . Visualization and diagnostic tools provide interactive comparison of trained models, producing plots that display the temporal evolution of wave profiles and deviations from analytical benchmarks. The analytical datasets used for both training and validation can be regenerated with increased spatial resolution to achieve higher accuracy in detailed simulations. All components are integrated through an interactive command-line interface that unifies the workflow and allows the user to execute all operations — from configuration and training to analysis and visualization — without manual scripting.

The developed software complex was verified on the classical problem of interfacial gravity–capillary waves between two incompressible, inviscid fluids of different densities. This benchmark allows testing the accuracy and stability of numerical methods against an analytical solution. The system describes potential fluid motion in two semi-infinite layers separated by a deformable interface, whose evolution is governed by kinematic and dynamic conditions accounting for gravity and surface tension. For verification, the linearized formulation was used, which admits an exact analytical solution expressed as a superposition of harmonic components with a known dispersion relation linking frequency and wavenumber. The analytical solution presented in [4] served as a reference for validating the accuracy and convergence of the PINN-based model.

A comprehensive computational analysis was conducted to assess the impact of the number of training points on the performance of physics-informed neural networks (PINNs) within the developed framework. The comparison involved two datasets consisting of 400 and 1000 training points, respectively, both derived from analytical wave profiles. The objective was to determine how an increase in data resolution affects model accuracy, stability, and generalization. The results demonstrated that increasing the number of training points generally improved model performance, particularly for architectures trained over longer epochs. The *Extended* configuration (7 layers, 25 neurons, 20,000 epochs) achieved the highest accuracy, improving from 97.19% to 98.86%, with a corresponding reduction in the mean absolute percentage error (MAPE) from 2.81% to 1.14%. This improvement indicates that the model effectively utilized additional data to refine its approximation of the analytical solution. The *Adaptive_LR* model (8 layers, 32 neurons, 8,000 epochs) exhibited exceptional stability, maintaining identical accuracy (98.22%) and MAPE (1.78%) across both datasets. This result confirms the robustness of adaptive learning rate strategies with respect to training data density. The *Deep* configuration (15 layers, 50 neurons, 5,000 epochs) showed a marginal yet consistent improvement (+0.05%), reflecting stable convergence in deeper architectures. Conversely, models with wider or larger architectures (*Wide* and *Large*) experienced significant degradation when trained with more data, likely due to overfitting or suboptimal hyperparameter selection. Their performance drop underscores the importance of model regularization and parameter tuning when increasing dataset size.

Statistical evaluation revealed that two models improved, one remained stable, and two deteriorated in performance, while two new configurations were introduced. The highest coefficient of determination ($R^2 > 0.9999$) was consistently achieved by the best-performing networks, confirming their high fidelity to the analytical benchmark. Overall, the analysis established that increasing the number of training points from 400 to 1000 yields measurable benefits for well-optimized architectures, particularly those employing extended training durations and adaptive optimization schemes. The *Extended* model was identified as the most accurate configuration, suitable for scientific research requiring maximal precision, while *Adaptive_LR* was recommended as the most reliable option for practical applications due to its numerical stability.

The study confirmed the accuracy, stability, and adaptability of the developed PINN framework for modeling internal gravity–capillary waves. The system successfully reproduces analytical results and provides a unified environment for systematic testing and optimization of neural network architectures. Future research will focus on extending the framework to nonlinear and multidimensional regimes, incorporating adaptive sampling and hybrid data-driven approaches, and

applying it to inverse problems and experimental data analysis. These developments are expected to enhance the efficiency, generalization capability, and applicability of PINNs to a wide range of complex hydrodynamic and geophysical systems.

1. Raissi M., Perdikaris P., Karniadakis G. E. *Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations*. Journal of Computational Physics. 2019. Vol. 378. P. 686–707. DOI: 10.1016/j.jcp.2018.10.045
2. Lin S., Chen Y. *A two-stage physics-informed neural network method based on conserved quantities and applications in localized wave solutions*. Journal of Computational Physics. 2022. Vol. 457. P. 111053. DOI: 10.1016/j.cpc.2022.111053
3. Ren P., Rao C., Chen S., Wang J.-X., Sun H., Liu Y. *SeismicNet: physics-informed neural networks for seismic wave modeling in semi-infinite domain*. Computer Physics Communications. 2024. Vol. 295. P. 109010. DOI: 10.1016/j.cpc.2023.109010
4. Avramenko O. *Deviation of two liquid half-spaces interface with surface tension: multiscale approach*. Mohyla Mathematical Journal. 2024. Vol. 7. P. 51–56. DOI: 10.18523/2617-70807202451-56

ML-ПІДХІД ДО ІДЕНТИФІКАЦІЇ ЕМОЦІЙ У ЗВУКОВИХ ЗАПИСАХ НА БАЗІ MFCC / MACHINE LEARNING APPROACH TO EMOTION IDENTIFICATION BASED ON MFCC

Тимошевський Д.С. / Tymoshevskiy D.S.

Національний університет “Києво-Могилянська Академія” / National University of Kyiv-Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра математики

E-mail: d.tymoshevskiy@ukma.edu.ua

This research presents the development of a custom Ukrainian dataset of human voice recordings labeled by seven key emotional states: angry, disgust, fear, happy, neutral, sad, and surprise. Each audio sample was processed to extract Mel-Frequency Cepstral Coefficients (MFCCs), which serve as primary acoustic features representing emotional cues. A Random Forest classifier was trained on these features to identify emotional categories from speech. The model achieved reliable accuracy and demonstrated robustness against minor variations in tone and recording conditions. The study highlights the efficiency of classical ensemble methods for emotion recognition and provides a reproducible dataset for further research in affective computing and audio-based AI systems.

У сучасному світі, де взаємодія людини з комп’ютерними системами набуває все більш природного характеру, розпізнавання емоцій за голосом стає одним із ключових напрямів розвитку штучного інтелекту. Голос є одним із найвиразніших каналів передавання емоційного стану, тому аналіз аудіосигналів відкриває можливості для створення інтелектуальних систем, здатних розуміти не лише зміст висловлювань, а й їх емоційне забарвлення. Такі технології мають широкий спектр практичного застосування — від покращення роботи віртуальних асистентів, систем підтримки клієнтів і адаптивних освітніх платформ до психологічного моніторингу стану користувачів. Окремим напрямом, що набуває актуальності, є використання аналізу емоційного тону голосу для оцінювання ефективності ведення гри в настільних рольових іграх, де емоційна залученість та динаміка голосу відображають якість ігрового процесу та комунікації між учасниками. Створення відкритого датасету аудіозаписів, розподілених за сімома базовими емоціями, а також навчання моделі машинного навчання для їх розпізнавання є важливим кроком до розвитку систем афективних обчислень, що дозволяють глибше розуміти поведінку людини та підвищувати якість інтерактивних ігрових і навчальних середовищ.

Для розпізнавання емоцій за голосом надзвичайно важливо коректно описати акустичні особливості мовлення, що відображають інтонацію, тембр, силу голосу та інші параметри, пов’язані з емоційним станом людини. Саме тому в дослідженні були обрані мел-частотні кепстральні коефіцієнти (MFCC) — один із найефективніших [1] та найпоширеніших способів представлення аудіосигналів у задачах аналізу мовлення. Їхня перевага полягає у здатності моделювати сприйняття звуків людиною: шкала мелів відображає нелінійну чутливість слуху до частот, наближену до психоакустичних властивостей слухового апарату. Таким чином, MFCC дають змогу отримати компактне, але інформативне відображення спектральних характеристик звуку, що добре підходить для подальшої обробки методами машинного навчання. На відміну

від сирих аудіосигналів або простих спектрограм, ці коефіцієнти зменшують вплив шумів, варіацій гучності та характеристик мікрофонів, забезпечуючи стабільність моделі. Крім того, використання MFCC дозволяє суттєво скоротити розмірність вхідних даних, зберігаючи при цьому ключову інформацію про тембр та артикуляцію. Завдяки цим властивостям MFCC стали стандартом де-факто у задачах розпізнавання мовлення, емоційного аналізу та біометричної ідентифікації голосу.

Одним із ключових викликів у сфері автоматичного розпізнавання емоцій за голосом є відсутність якісних україномовних відкритих наборів даних. Більшість наявних датасетів, що використовуються у світових дослідженнях, містять аудіозаписи англійською, німецькою чи китайською мовами, що робить їх малопридатними для моделювання емоційного забарвлення української мови. Це пов'язано з фонетичними, інтонаційними та просодичними відмінностями, які безпосередньо впливають на акустичні характеристики емоційного мовлення. Відтак створення українського датасету аудіозаписів, класифікованих за базовими емоціями, є необхідним кроком для розвитку національних досліджень у сфері афективних обчислень і голосової аналітики.

Розроблений датасет представлений у двох формах:

1. У вигляді набору коефіцієнтів MFCC із відповідними мітками емоцій (angry, disgust, fear, happy, neutral, sad, surprise), що дозволяє безпосередньо використовувати дані для навчання та тестування моделей машинного навчання.
2. У вигляді сирих аудіофайлів, розсортованих по папках, що відповідають кожній емоції. Така структура забезпечує гнучкість використання — як для подальшої обробки, так і для візуалізації спектрограм або порівняння з іншими наборами даних.

Для задачі класифікації емоцій за голосом у цьому дослідженні було обрано модель Random Forest, яка поєднує простоту реалізації, високу точність і стійкість до шумів у даних [2]. Суть цього методу полягає у створенні ансамблю рішень — великої кількості незалежних дерев рішень, кожне з яких навчається на випадковій підмножині даних і ознак. Остаточний результат визначається за принципом голосування між деревами, що дозволяє зменшити ризик перенавчання, притаманний іншим моделям. Такий підхід забезпечує стабільність результатів навіть за умов обмеженого обсягу даних чи варіативності записів.

Перевагою Random Forest у порівнянні з нейронними мережами чи іншими складними моделями є менша потреба в попередньому налаштуванні параметрів та вища інтерпретованість результатів — можна оцінити важливість кожного з коефіцієнтів MFCC для розпізнавання певної емоції. Крім того, метод добре працює з даними різного масштабу, не потребує нормалізації та демонструє високу узагальнювальну здатність на невеликих вибірках, що особливо важливо на етапі створення нового датасету. Завдяки цим властивостям Random Forest став оптимальним вибором для побудови базової, але надійної моделі розпізнавання емоцій на основі акустичних ознак голосу.

Під час експериментів було проведено порівняльний аналіз точності моделі Random Forest, натренованої на різних мовних наборах даних. Модель, навчена на англійськомому датасеті, продемонструвала лише 19% правильних класифікацій емоцій у випадку розпізнавання українських аудіозаписів, що підтвердило значну залежність емоційних характеристик голосу від мовних та фонетичних особливостей. Натомість модель, навчена безпосередньо на створеному українському датасеті, досягла 81% точності, що свідчить про її ефективність і релевантність до особливостей українського мовлення. Отримані результати підтверджують доцільність розробки локалізованих наборів даних для задач розпізнавання емоцій за голосом.

У результаті проведеного дослідження було створено базовий україномовний датасет аудіозаписів, класифікованих за сімома емоційними станами: angry, disgust, fear, happy, neutral, sad, surprise. На основі отриманих мел-частотних кепстральних коефіцієнтів (MFCC) побудовано модель класифікації емоцій за допомогою алгоритму Random Forest, який показав стабільну точність та високу узагальнювальну здатність на обмеженій кількості даних. Обґрунтовано вибір MFCC як оптимального набору ознак для опису емоційних характеристик мовлення, а також доведено ефективність Random Forest як інтерпретованої та надійної моделі для цієї задачі. Створений датасет має практичну цінність як відкрита основа для подальших досліджень у сфері афективних обчислень, розпізнавання мовлення та емоційного аналізу.

Список джерел:

1. Valiveti H., Budati A. K. Feature extraction algorithms to improve the speech emotion recognition rate. *Int J Speech Technol* **23**, 2020, 45–55. URL: <https://doi.org/10.1007/s10772-020-09672-4>

2. Kilimci Z. H., Bayraktar Ülkü, Kucukmanisa A. Evaluating raw waveforms with deep learning frameworks for speech emotion recognition. *Multimed Tools Appl*, 2025. URL: <https://doi.org/10.1007/s11042-025-20930-y>

РОЗРОБЛЕННЯ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ МОДЕЛЮВАННЯ ПРОЦЕСУ КРИСТАЛІЗАЦІЇ ЗЛИВКУ З ВИКОРИСТАННЯМ НЕЧІТКОЇ ЛОГІКИ / DEVELOPMENT OF AN INTELLIGENT SYSTEM FOR MODELING THE CRYSTALLIZATION PROCESS OF A MELT USING FUZZY LOGIC

Павлюченко В., Бабенко М. / Pavliuchenko V., Babenko M.

Дніпровський державний технічний університет / Dniprovsky State Technical University
51918, м. Кам'янське Дніпропетровської обл., вул. Дніпробудівська, 2, факультет комп'ютерних технологій та енергетики, кафедра програмного забезпечення систем
E-mail: mvbab@ukr.net

This research explores the use of fuzzy logic methods to achieve intelligent control during the cooling stage of ingot formation in mold casting processes. The main challenges we ran into when trying to automate these metallurgical processes were: dealing with uncertainty in sensor readings, transforming human operator experience into computational algorithms, and ensuring rapid real-time control responses. So we built a Mamdani fuzzy inference system to tackle these problems to turn what the operators were telling us into actual control commands. The system processes three main input parameters: molten metal temperature, casting speed, and cooling rate, and applies a rule base of 27 fuzzy if-then statements to determine adaptive cooling corrections. Experimental tests confirmed the model's high reliability, achieving a correlation of 0.91 with expert evaluations and maintaining reaction times below one second. The results look promising for actually using this in a real plant. We coded everything in Python, mainly using, particularly scikit fuzzy, NumPy, and SciPy.

Проблема керування охолодженням зливу під час його кристалізації ускладнюється значною параметричною невизначеністю, нелінійною природою теплових процесів і необхідністю включення експертного досвіду операторів у формалізовані алгоритми прийняття рішень. Традиційні детерміністичні підходи потребують детальних математичних моделей і повної інформації про граничні умови, що не завжди можливо в умовах реального виробництва [1]. Це викликає попит на методологію, яка здатна працювати з якісними характеристиками процесу та обмеженою кількістю даних.

Основною ідеєю є використання нечіткої логіки для узагальнення та відтворення досвіду операторів під час процесу кристалізації зливу. Такий підхід дозволяє перетворити практичні знання фахівців у систему алгоритмічних правил, що оперують лінгвістичними змінними замість точних числових значень. Наприклад, замість жорсткого рівняння вигляду «якщо температура розплаву 1560 °С, то витрати охолоджуючої рідини 350 л/хв.», застосовуються більш гнучкі правила: "якщо температура ВИСОКА, швидкість витягування СЕРЕДНЯ і поточне охолодження НИЗЬКЕ, тоді СУТТЄВО ЗБІЛЬШИТИ інтенсивність охолодження". Такий підхід повніше відповідає способу мислення операторів, які зазвичай використовують якісні оцінки для аналізу стану процесу.

Технологічний алгоритм складається з чотирьох етапів обробки інформації. На першому етапі – фазифікації – конкретні виміри від датчиків (наприклад, температура 1560 °С, швидкість 1,2 м/хв) перетворюються у ступені належності до нечітких термів за допомогою відповідних функцій належності, як трикутної чи трапецієподібної. Так, температура 1560 °С може одночасно мати ступінь належності 0,7 до терму «висока» і 0,3 до «нормальна». На наступному етапі – агрегації антецедентів – визначається ступінь активації кожного правила, застосовуючи операції мінімуму («I») або максимуму («АБО»). Третій етап – композиція висновків – дозволяє об'єднати результати спрацьованих правил у нечітку множину вихідної змінної. Завершальний, четвертий етап – дефазифікація – перетворює отриману нечітку множину у конкретне числове значення корекції інтенсивності охолодження, використовуючи, наприклад, метод центру ваги [2].

Основною перевагою цього методу є здатність ефективно працювати з конфліктними або частковими даними, виключаючи помилки під час обчислень. Наприклад, якщо датчик температури фіксує показник, що знаходиться на межі двох категорій, система автоматично враховує обидва значення з відповідними ваговими коефіцієнтами. Це суттєво відрізняється від класичних підходів, які вимагають суворого порогового визначення і можуть викликати різкі зміни керуючого сигналу. Завдяки активації множини правил забезпечується плавне регулювання: підвищення температури з 1555 °С до 1565 °С призводить до поступової зміни рекомендацій, а не до раптових стрибків.

Програмна реалізація виконана на Python із залученням бібліотек scikit-fuzzy для роботи з нечіткою логікою, NumPy і SciPy для математичних розрахунків, Matplotlib для візуалізації та Tkinter для створення графічного інтерфейсу. Математична модель теплопровідності базується на одновимірному нестационарному рівнянні Фур'є з граничними умовами третього роду, яке розв'язується чисельно методом Рунге-Кутти четвертого порядку з інтервалом інтегрування 0,1 секунди.

Підхід до формування бази знань передбачає розбиття експертної інформації операторів на прості правила прийняття рішень. Для трьох вхідних змінних із 3 – 4 термами кожна теоретично можливе формування бази з 27 – 48 правил. Проте на практиці використовується лише підмножина найважливіших комбінацій. Наприклад, ситуація типу «дуже висока температура + висока швидкість + низьке охолодження» класифікується як критична і потребує максимального керуючого впливу (+ 50%). Водночас конфігурація «низька температура + низька швидкість + високе охолодження» вимагає значного зниження впливу (– 50%), щоб уникнути термічного шоку металу.

Проведені випробування на реальних виробничих даних засвідчили високу результативність запропонованого підходу. Аналіз показав, що рекомендації, сформовані нечіткою логічною системою, майже повністю збігаються з рішеннями, які зазвичай приймають досвідчені оператори. Значення коефіцієнта кореляції 0,91 підтверджує тісний взаємозв'язок між результатами системи та експертними оцінками, що свідчить про адекватне відтворення логіки фахівців у моделі. Один повний цикл обчислень виконується за приблизно 0,15 секунди на типовому персональному комп'ютері, що значно швидше порівняно з класичним методом скінченних елементів. Важливо, що система демонструє надійну роботу навіть за змін умов чи параметрів процесу, забезпечуючи стабільні та коректні рекомендації без збоїв.

Метод демонструє високу адаптивність і зручність масштабування. Щоб додати нові параметри, наприклад марку сталі, товщину виробу або якість мастильного матеріалу, достатньо розширити базу правил, не змінюючи принципи роботи системи. Вона здатна підлаштовуватися під реальні умови експлуатації, автоматично коригуючи параметри функцій належності на основі даних про роботу обладнання.

Практичні результати впровадження цього підходу підтверджують доцільність використання нечіткої логіки для розв'язання завдань інтелектуального керування у металургійній сфері. Поєднання досвіду фахівців із потужністю обчислювальних алгоритмів створює ефективну альтернативу класичним математичним моделям і підвищує якість прийняття технологічних рішень. Також підхід характеризується стійкістю до збоїв окремих сенсорів або впливу шумів у вимірювальних каналах. Завдяки використанню нечітких функцій належності система може працювати навіть у разі отримання неповних або неточних даних, аналізуючи ситуацію за ступенем приналежності до встановлених термів. У промислових умовах, де вплив навколишніх факторів (температура, вібрація, пил) ускладнює точність вимірювань, такі властивості надають значні переваги порівняно з класичними алгоритмами керування.

Подальший розвиток методу може враховувати інтеграцію модулів прогнозування динаміки температурного поля за допомогою методів машинного навчання. Такий підхід дасть змогу реалізувати не тільки реактивні, а й проактивні управлінські дії: наприклад, попереднє регулювання інтенсивності охолодження у відповідь на прогнозовані зміни швидкості розливання чи температури металу. Комбінація таких рішень на основі нечітко-логічного підходу відкриває шлях до створення повноцінних систем «розумного лиття», здатних до автоматичної адаптації в умовах реального часу.

Список джерел

1. Mendel J.M. Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions. 2nd Edition. Springer International Publishing, 2018. 684 p. URL: <https://doi.org/10.1007/978-3-319-51370-6> (дата звернення: 20.10.2025)
2. Karaboga D., Kaya E. Adaptive network based fuzzy inference system (ANFIS) training approaches: a comprehensive survey. *Artificial Intelligence Review*. 2018. Vol. 52, no. 4. P. 2263–2293. URL: <https://doi.org/10.1007/s10462-017-9610-2> (дата звернення: 30.10.2025).

Масово паралельна реалізація MorphoNAS / Massively Parallel MorphoNAS Implementation
Медвідь С.О. / Medvid S.

Our GPU-accelerated MorphoNAS (Morphogenetic Neural Architecture Search) enhances the biologically motivated developmental model of the original MorphoNAS, from sequential developmental simulations to large scale parallel simulation. Like the original CPU-based MorphoNAS system, the GPU enhanced version models morphogenesis (neural development) through reaction-diffusion, progenitor differentiation and axon-guided wiring that grow recurrent controllers. Unlike the original CPU-based version, the GPU-based version has the ability to formulate growth and control in parallel execution of thousands of genomes. Through use of sparse, device-resident representations of developmental dynamics and recurrent rollouts, we produce behaviorally equivalent recurrent controllers with up to three orders of magnitude speedup. As a result, we have enabled the possibility of performing evolutionary search over populations of developmental programs, transforming MorphoNAS from a proof-of-concept model of artificial morphogenesis and adaptive architecture discovery to a scalable framework for those objectives.

Background

MorphoNAS [1] is a biologically inspired approach to neural architecture search that views the development of recurrent neural networks as a process of morphogenesis [2].

In its original form, MorphoNAS executes **development** sequentially by first simulating morphogen secretion, diffusion, inhibition, and cell fate transitions from a single progenitor cell to a recurrent neural network. Because of this sequential nature of the process, it limits the size of the search space that can be explored. It exhibits both control-flow irregularity and memory access irregularity [3].

MorphoNAS also defines how the resulting networks will behave as **controllers**. The original MorphoNAS implementation performs the control loop sequentially by evaluating the behavior of each developed network individually and executing them in parallel using multiple CPU cores. However, this approach scales poorly. The cost of the rollouts increases quadratically with the number of neurons, and must be performed for each of the controller-environment pairs.

MorphoNAS GPU Implementation

To construct the **developmental** component of MorphoNAS, we preserve the morphogenetic rules used in the original MorphoNAS implementation and reformulate them to enable batched, massively parallel execution: diffusion and secretion are implemented as convolutional CUDA kernels; genome parameters are broadcast across the developmental grid of each network being grown; and cell division, differentiation, and synaptogenesis are performed for thousands of networks at a time.

For the **controller** component of MorphoNAS, we preserve the exact same control laws as the original MorphoNAS implementation: here, we represent them in a sparse CSR-like format representation (i.e., row pointers, edge indices, weights); then concatenate multiple such sparse representations of controllers, along with additional metadata describing their properties (neuron range, edge range, input neuron indices, output neuron indices, and internal settling depth) into large device buffers. We then execute the rollout portion of the MorphoNAS control loop using a single CUDA kernel that evaluates all active controllers in parallel. Controllers' observations are injected into the appropriate input neurons; the kernel then iteratively updates neurons of each controller using on-GPU ping-pong buffers for state; and finally, the kernel collects activity of output neurons of each controller as an action. Temporal memory of controllers is preserved throughout the entire loop, without incurring the overhead associated with moving the hidden state between host and device [4].

Functionally, "Evaluate one dense recurrent controller sequentially" becomes "Evaluate a population of sparse recurrent controllers simultaneously," allowing us to execute the rollout portion of the MorphoNAS control loop in parallel over the entire population of controllers. Therefore, we can evaluate hundreds or thousands of different controllers simultaneously, and apply selection based on their performance on the task at hand.

Performance and Scaling

Developmental growth is now population-scale. GPU-driven growth reaches 9,710 networks/s on RTX 5090 and 5,379 networks/s on RTX 3090, versus 15-20 networks/s on high-core-count CPUs (12-core Apple M2 Max: 15.2 networks/s; 56-core Xeon: 14.4 networks/s; 192-core AMD EPYC: 19.9 networks/s). Poor CPU speedup is attributed to limited multithreading scalability in the BLAS/MKL implementations. Apple's Accelerate demonstrated the highest parallel efficiency which confirms

findings in [5]. GPU implementation results in ~354x-639x speedup over CPU-based growth, even against 192-core servers.

Recurrent controller rollout accelerates 1-2 orders of magnitude depending on the networks diversity. For the real evolutionary runs with highly diverse networks, GPU recurrent rollout reaches ~0.5-1.1 million propagations per second depending on GPU and network class. This represents roughly 6-10x speedup over a parallel 12-core Apple M2 Max CPU baseline (113,837 propagations/s). Importantly, the GPU implementation preserves the exact recurrent control loop used on CPU.

The bottleneck moves from development to behavior. With GPU acceleration, developmental growth becomes effectively free compared to other parts. Growth on RTX 5090 proceeds at 9,710 networks/s, whereas even the best CPU setup reaches only 19.9 networks/s for the same genomes.

Implications. GPU implementation is able to simulate growth of thousands of networks in parallel and apply parallel evaluation of thousands controllers at a time. It has eliminated morphogenesis as a computational bottleneck in MorphoNAS and significantly widened target search space. The combination of biological realism with large-scale parallelism has potential to transform MorphoNAS into a scalable platform for discovering adaptive, self-organizing neural architectures.

Sources

- [1] M. Glybovets and S. Medvid, "MorphoNAS: Embryogenic Neural Architecture Search Through Morphogen-Guided Development," *arXiv preprint arXiv:2507.13785*, 2025, [Online]. Available: <https://arxiv.org/abs/2507.13785>
- [2] A. M. Turing, "The Chemical Basis of Morphogenesis," *Bulletin of Mathematical Biology*, Jan. 1952, doi: 10.1016/S0092-8240(05)80008-4.
- [3] M. Burtscher, R. Nasre, and K. Pingali, "A quantitative study of irregular programs on GPUs," in *2012 IEEE International Symposium on Workload Characterization (IISWC)*, Nov. 2012, pp. 141–151. doi: 10.1109/IISWC.2012.6402918.
- [4] G. Diamos *et al.*, "Persistent RNNs: Stashing Recurrent Weights On-Chip," in *Proceedings of The 33rd International Conference on Machine Learning*, PMLR, June 2016, pp. 2024–2033. Accessed: Oct. 31, 2025. [Online]. Available: <https://proceedings.mlr.press/v48/diamos16.html>
- [5] P. Hübner, A. Hu, I. Peng, and S. Markidis, "Apple vs. Oranges: Evaluating the Apple Silicon M-Series SoCs for HPC Performance and Efficiency," in *2025 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, June 2025, pp. 45–54. doi: 10.1109/IPDPSW66978.2025.00013.

АДАПТИВНА МУЛЬТИМОДАЛЬНА КОГНІТИВНА СИСТЕМА ДЛЯ АВТОНОМНОГО ПРИЙНЯТТЯ РІШЕНЬ У СКЛАДНИХ ЛОГІСТИЧНИХ СЕРЕДОВИЩАХ

Пасєка П.І., Теренчук С. А.

Київський національний університет будівництва і архітектури
03037, Київ, проспект Повітряних Сил, 31, факультет автоматизації і інформаційних систем,
кафедра інформаційних технологій проектування та прикладної математики
E-mail: paaaa.paul@gmail.com, terenchuksa@ukr.net

This research focuses on a conceptual framework for an adaptive multimodal cognitive system designed to optimize logistics processes in dynamic and uncertain environments. Proposed architecture for a continually updated world-model (knowledge graph) from fused sensory inputs that employs interacting with cognitive agents to decompose logistics objectives, plan actions, and re-plan in response to incoming observations. Unlike traditional static optimization approaches, the proposed system implements perception–action feedback loops enabling real-time adaptation and robustness to incomplete or adversarial data. Expected contributions include a formalization of the perception-to-planning pipeline, a hybrid agentic optimization scheme, and evaluation scenarios demonstrating improved resilience and decision latency in non-standard logistics tasks. The research contributes to developing intelligent logistics systems capable of operating effectively in complex, non-standard, or military scenarios.

Сучасні логістичні системи функціонують у складних умовах, що характеризуються високою динамічністю, невизначеністю та нестачею достовірних даних. Це особливо актуально у контексті гуманітарних або військових операцій, де рішення щодо постачання, транспортування

чи розподілу ресурсів мають прийматися в реальному часі та часто за відсутності повної інформації про середовище.

Класичні підходи до логістичної оптимізації – від математичних алгоритмів до евристичних і нейронних методів – ефективні лише за умов статичного або добре визначеного середовища. Водночас технології комп’ютерного зору, аудіоаналізу та сенсорного сприйняття відкривають можливість створення інтелектуальних систем, здатних динамічно оцінювати ситуацію й адаптувати свої рішення відповідно до змін середовища.

У цій роботі пропонується концепція адаптивної мультимодальної когнітивної системи, яка інтегрує сприйняття навколишнього середовища (зображення, відео, звукові записи, телеметричні дані [1]) із процесом оптимізації логістичних дій. Система поєднує модулі розпізнавання, аналізу, міркування та планування, утворюючи замкнений цикл сприйняття–дії, подібний до когнітивного циклу людини.

Сучасні дослідження в області транспортної логістики зосереджуються переважно на задачах оптимізації маршрутів, розподілу ресурсів та прогнозуванні попиту з використанням класичних нейронних мереж або методів підкріплювального навчання. Відомі рішення на основі XGBoost, DQN [2, 3] демонструють ефективність у стабільних середовищах, проте погано масштабуються до умов змінної структури даних або непередбачуваних подій.

Водночас активно розвиваються напрямки мультимодального розпізнавання — інтеграції різноманітної інформації для побудови узагальненої “картини світу”. Проте більшість таких систем зосереджена на сприйнятті, не інтегруючи його із рівнем прийняття рішень. Таким чином, залишається відкритою проблема створення єдиної когнітивної моделі, здатної зв’язати сенсорне сприйняття з оптимізаційним міркуванням.

Для вирішення цієї проблеми запропоновано стан середовища в певний момент часу t визначати на основі множини мультимодальних спостережень, а оптимальний план дій, який мінімізує сукупну вартість або ризик логістичних операцій знайти наступним чином:

$$P^{\square}(t) = \arg \min_P E_{E(t)} [C(P, E(t))],$$

де:

$E(t)$ – це стан середовища в момент часу t (погодні умови, наявність транспорту, загрози, доступність маршрутів тощо);

$C(P, E(t))$ – функція вартості (ризик) для плану дій P у конкретному стані середовища $E(t)$.

Оскільки $E(t)$ не завжди можна точно виміряти або передбачити, його вважають випадковою величиною. Тоді математичне сподівання означає очікувану середню вартість плану P , обчислену за всіма можливими станами середовища з урахуванням їх імовірностей.

Особливістю запропонованого рішення є те, що стан середовища не є фіксованим, а постійно оновлюється на основі нових сенсорних даних і результатів дій системи. Це створює замкнутий когнітивний цикл, де сприйняття безпосередньо впливає на стратегію оптимізації.

Система, що пропонується для автономного прийняття рішень у складних логістичних середовищах складається з чотирьох взаємопов’язаних шарів:

4. Шар сприйняття, що використовує моделі комп’ютерного зору, аудіоаналізу та обробки сенсорних сигналів для виділення релевантних об’єктів, подій і контекстів.
5. Шар злиття даних, у якому отримані спостереження перетворюються у граф знань де вершини описують об’єкти, а ребра – просторово-часові чи логічні зв’язки між ними. Цей граф виступає «моделлю світу», що постійно оновлюється. [4]
6. Шар міркування та оптимізації: на основі графа працює набір агентів – когнітивних модулів, кожен із яких відповідає за підзадачу (маршрутизацію, прогнозування, пріоритизацію). Агенти взаємодіють, використовуючи генеративні моделі для формування нових стратегій дій. [5]
7. Шар адаптації та зворотного зв’язку: система оцінює результати дій, порівнює їх із очікуваннями та коригує як параметри моделей сприйняття, так і стратегії планування. Таким чином утворюється когнітивний контур де стан середовища у наступний момент часу формується на основі впливу дій агентів на середовище та результатах спостережень після їх виконання.

Завдяки такій архітектурі система здатна самонавчатися, накопичувати досвід і поступово покращувати якість своїх рішень у динамічному середовищі.

Наукова новизна запропонованого підходу полягає у створенні єдиної мультимодальної когнітивної архітектури, що поєднує:

1. розпізнавання та аналіз навколишнього середовища з різних джерел даних;

2. побудову уніфікованої моделі світу у вигляді графа знань;
3. агентне прийняття рішень із генеративними властивостями;
4. адаптивний контур зворотного зв'язку між сприйняттям і плануванням.

Очікуваними результатами є:

1. підвищення стійкості логістичних систем до невизначеності та неповноти даних;
2. скорочення часу ухвалення рішень у динамічному середовищі;
3. покращення узгодженості між сенсорними спостереженнями та діями системи.

Таким чином ця робота закладає основи для створення інтелектуальних автономних логістичних систем, здатних працювати у складних або ворожих умовах – від кризових гуманітарних місій до військових операцій.

Список джерел:

1. Bajcsy R. Active Perception. *Proceedings of the IEEE*. 1988. P. 996–1005.
2. Baltrušaitis T., Ahuja C., Morency L.-P. Multimodal Machine Learning: A Survey and Taxonomy. *ArXiv:1705.09406*. 2018.
3. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System. *ArXiv:1603.02754*. 2016.
4. A Comprehensive Survey on Graph Neural Networks. / Z. Wu et al. *ArXiv:1901.00596*. 2019.
5. ReAct: Synergizing Reasoning and Acting in Language Models. / S. Yao et al. *ArXiv:2210.03629*. 2022.

ЕКСПЕРИМЕНТАЛЬНИЙ АНАЛІЗ МОДЕЛІ ЕВОЛЮЦІЙНОЇ ГРИ В ІМЕНУВАННЯ / EXPERIMENTAL ANALYSIS OF THE EVOLUTIONARY NAMING GAME MODEL

Сітков І.П.¹, Гулаєва Н.М.^{1,2} / Sitkov I.¹, Gulayeva N.^{1,2}

¹Національний університет “Києво-Могилянська Академія” / ¹National University of Kyiv-Mohyla Academy,

²Університет Севільї / ²Universidad de Sevilla

¹04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики

²Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad de Sevilla, E.T.S.

Ingeniería Informática, Avda. Reina Mercedes s/n, 41012 Sevilla, Spain

E-mail: i.sitkov@ukma.edu.ua, gulayeva@ukma.edu.ua

Agent-based modelling is the principal method for studying natural language emergence and evolution due to the extreme complexity of the process. Lipowskis' Evolutionary Naming Game model lacks statistically-based studies on its reproducibility and sensitivity to parameter variations. Our research aims to analyse the effect of the model parameters on its dynamics and to verify the reproducibility of known results. We implemented the Evolutionary Naming Game model and conducted a series of experiments described by Lipowskis. Our experimental analysis confirmed the reproducibility of main behavioural patterns and demonstrated the model general robustness to parameter variations, although some mismatches in numerical values were observed. The statement suggesting a correlation between agents' languages and their learning abilities was refuted.

Причини виникнення та механізми розвитку природної мови залишаються предметом досліджень у лінгвістиці, біології та комп'ютерній науці протягом багатьох років. Вивчення еволюції мови в реальних умовах є складним через тривалість мовотворчих процесів, динамічність соціального середовища та велику кількість зовнішніх чинників. У зв'язку зі сказаним, основним методом дослідження мовотворчих процесів стає обчислювальне моделювання.

У роботі розглядається агентна модель виникнення та розвитку мови – Еволюційна Гра в Іменування (ЕГІ) [1-3]. Модель представлена популяцією агентів, розмішених у двовимірній квадратній сітці з довжиною сторони L . Кожен агент характеризується навчальною здібністю, віком, словником та рівнем знань (сума ваг слів словника). Симуляція відбувається протягом фіксованої кількості ітерацій, де на кожній ітерації виконуються $L \times L$ операцій над агентами. Модель передбачає два типи операцій: комунікацію агентів та оновлення популяції. Під час комунікації випадковий агент виконує роль мовця та пропонує випадкове слово зі свого словника агенту-слухачу, обраному серед сусідів мовця. Якщо словник агента-слухача містить пропонуване слово, то комунікація є успішною та обидва агенти збільшують вагу слова у своїх

словниках на значення навчальної здібності. В іншому разі агент-мовець зменшує вагу слова на значення навчальної здібності, а агент-слухач додає невідоме слово до свого словника. Під час оновлення популяції агент або утворює нащадка, або гине (визначається ймовірністю виживання). Утворений нащадок або отримує випадковий словник та навчальну здібність, або успадковує ці дані від батька (визначається ймовірністю мутації). Модель ЕГІ характеризується низкою параметрів, такими як кількість ітерацій, розмір агентної сітки, розмір словника агентів, ймовірність комунікації (p_{comm}), ймовірність мутації, міра впливу віку та рівня знань агентів на ймовірність їх виживання тощо.

У роботах [1-3] описано такі властивості моделі ЕГІ. По-перше, популяція агентів за певної ймовірності комунікації здійснює біолінгвістичний перехід (БЛП) – різкий перехід від лінгвістично неоднорідного стану до монолінгвальності, від низьких значень навчальних здібностей та успішності комунікації до високих. Відзначимо, що в [1, 2] наявність БЛП розглядається як демонстрація ефекту Болдвіна. По-друге, у роботах [1-3] заявлена стійкість моделі до зміни значень її параметрів, при цьому в [3] наголошується на чутливості алгоритму до певних модифікацій. По-третє, у [3] акцентується увага на кореляції між кластерами агентів, що володіють спільною мовою та мають однакові навчальні здібності.

Зазначимо, що в роботах [1-3] відсутні детальні описи умов проведення експериментів та не наведені статистичні дані, на основі яких були зроблені висновки. Тож складається враження, що частина висновків зроблена за результатами лише одного прогону алгоритму, а висновок про кореляцію між мовами та навчальними здібностями агентів спирається лише на візуальне порівняння кластерів. Зі сказаного випливає необхідність перевірки відтворюваності результатів та проведення статистично обґрунтованого аналізу властивостей моделі.

Нами було виконано програмну реалізацію моделі ЕГІ відповідно до опису алгоритму в [1, 2] та скоригованого псевдокоду з [3] (виправлено ймовірність мутації мови й навчальної здібності, усунуто зайве оновлення популяції). Було відтворено такі експерименти з [1-3]:

1 – виконання перших 8000 ітерацій із $p_{comm}=0.02$ та використання $p_{comm}=0.98$ для наступних 42000 ітерацій;

2 – виконання 100000 ітерацій з лінійним збільшенням p_{comm} від 0.1 до 0.5;

3 – збільшення (зменшення) p_{comm} з кроком 0.01 та зі сталою кількістю ітерацій на одне значення p_{comm} .

Аналіз стійкості моделі був проведений для таких значень параметрів:

10. розмір словника агентів (1, 5, 1000);
11. довжина сторони агентної сітки L (20, 30, 40, 60, 70);
12. початковий вік агентів (0, 1, 5, 10);
13. вплив віку на ймовірність виживання (0, 0.005, 0.05, 0.1, 0.5, 4);
14. визначення множини сусідів (окіл Мура, окіл фон Неймана, 4 сусіди на діагоналях).

Порівняльний аналіз отриманих результатів із заявленими в [1-3] засвідчив часткову відтворюваність поведінкових патернів моделі. Експериментально були підтверджені загальні тенденції у розвитку популяцій агентів, зокрема, БЛП, зворотний БЛП, кластеризація агентів за мовами та навчальними здібностями, уніфікація мов та навчальних здібностей у ході еволюції, а також сигмоїдна форма графіків у експерименті 1. Утім, не було досягнуто точного збігу в значеннях ймовірності комунікації БЛП та формі графіків навчальних здібностей і успішності комунікації. Наприклад, у експерименті 2 ймовірність комунікації БЛП у [2] склала 0.4, у [3] – 0.12, у наших експериментах – 0.16875 ± 0.01031 . Розбіжності в результатах можна пояснити відмінностями в умовах проведення експериментів та відсутністю достатньої кількості зібраних статистичних даних.

Результати експериментів 1-3 підтвердили стійкість моделі до зміни значень досліджуваних параметрів: були зафіксовані кількісні зміни в показниках ЕГІ (ймовірність комунікації БЛП, чисельність популяції, середній вік та рівень знань агентів) при збереженні загальних тенденцій.

Кореляцію між мовами та навчальними здібностями агентів було перевірено з використанням метрик кластерного аналізу. Результати експериментів продемонстрували повну відповідність між множинами кластерів за відсутності мутацій та комунікації. Застосування мутації порушило повну узгодженість кластерів. Долучення операції комунікації призвело до зниження кореляції до рівня випадкової, що, у підсумку, дозволило спростувати твердження.

Список використаних джерел

1. Lipowski A., Lipowska D. Bio-linguistic transition and Baldwin effect in an evolutionary naming-game model. *International Journal of Modern Physics C*. 2008. Т. 19, вип. 3. С. 399–407. URL: <https://doi.org/10.1142/S0129183108012248>.
2. Lipowska D. Naming Game and Computational Modelling of Language Evolution. *Computational Methods in Science and Technology*. 2011. Т. 17, вип. 1-2. С. 41–51. URL: <https://doi.org/10.12921/CMST.2011.17.01.41-51>.
3. Moloney P. Evolutionary Naming Game: an Agent-based Model for the Emergence of Language : M.Sc. Computer Science - Augmented and Virtual Reality / Trinity College Dublin. Dublin, 2021. URL: <https://publications.scss.tcd.ie/theses/diss/2021/TCD-SCSS-DISSERTATION-2021-042.pdf> (дата звернення: 31.10.2025).

GPU-ОРІЄНОВАНА БІДІАГОНАЛІЗАЦІЯ В АЛГОРИТМІ СИНГУЛЯРНОГО РОЗКЛАДУ З ШВИДКОЮ ДЕФЛЯЦІЄЮ ТА ПАРАЛЕЛЬНИМ РОЗБИТТЯМ / GPU-NATIVE BIDIAGONAL SVD WITH OPPORTUNISTIC DEFLATION AND PARALLEL SPLITTING

Сухарський С. С. / Sukharskyi S. S.

Інститут програмних систем НАН України

Київ, 03187, Україна, Проспект Академіка Глушкова, 40, корп. 5, +38 (044) 526-33-19

E-mail: serhii.sukharskyi@gmail.com

This research focuses on a GPU-native bidiagonal SVD that performs scale-aware, opportunistic deflation with immediate splitting into independent subproblems. We replace full upper/lower sweeps with an implicit-shift (Wilkinson) paired left/right Givens bulge-chasing, sharply reducing iterations. Householder reduction uses operations optimization techniques, and rotations on U and V are applied in tiled, GEMM-like passes. Avoiding per-rotation kernel launches and improving inter-block coordination yield 2–3× speedups over the initial version [5]. Early splitting exposes parallelism across subblocks, improving SM occupancy and time-to-solution for large dense matrices. The method preserves numerical stability and near-Jacobi orthogonality via scale-aware tests. We also propose a “step-0” heuristic targeting the largest-magnitude bidiagonal entries to hasten early deflation. Ongoing evaluation benchmarks against Jacobi, QR/DC, and polar-decomposition SVDs across different matrix sizes and conditions.

У роботі розглянуто GPU-орієнтований алгоритм SVD, що намагається усунути проблеми традиційних підходів завдяки масштабованій, опортуністичній дефляції на бідіагональній формі з негайним розбиттям на незалежні підзадачі. Таке раннє розбиття перетворює один великий, послідовний конвеєр складності $O(n^3)$ на кілька менших, паралельних ядер, що покращує використання локального кешу, підвищує завантаження потокових мультипроцесорів (SM) та забезпечує ефективні за пропускну здатністю, GEMM-подібні оновлення накопичених ортогональних матриць U та V . У підсумку істотно зменшується час обчислень для великих щільних матриць, що є критично важливим для задач на кшталт аналізу основних компонент (PCA), спектральної фільтрації та аналітики даних.

Наявні бібліотеки зазвичай надають перевагу надійним методам: однобічному Якобі (відмінна ортогональність; сильний у пакетній обробці малих/середніх задач) [4, 5], класичним QR/Розділяй-і-Пануй (DC) (висока надійність, але часто повільніші на GPU без ретельного блокування) [1], а також новішим SVD на основі полярного розкладу (висока швидкість на дуже великих щільних задачах) [2]. Існують багатоетапні GPU-розклади — перетворення щільних матриць у бідіагональну форму (інколи через проміжну стрічкову), — однак подальша бідіагоналізація часто лишається «вузким місцем» [3] через притаманну послідовність переслідування «горба» (bulge chasing) та відсутність якнайранішої дефляції з динамічним плануванням підзадач. Це окреслює потребу для розробки алгоритму бідіагоналізації як частини алгоритму сингулярного розкладу з використанням усіх можливостей графічних процесорів, що поєднує неявні зсуви, жадібну дефляцію й розбиття, а також ефективне паралельне застосування обертань до накопичених множників U і V .

Завдання цього дослідження спроектувати та оцінити повний алгоритм SVD, оптимізований під сучасні графічні процесори, який охоплює: початкову редуцію матриці до бідіагональної форми з використанням ефективних перетворень Хауссхолдера; механізм діагоналізації з неявним зсувом і негайною дефляцією та розбиттям; застосування накопичених обертань до матриць U і V у паралельних проходженнях; ефективне планування підзадач за допомогою окремих ядер і динамічних черг на сучасних архітектурах GPU. Кінцева мета —

досягти суттєвого виграшу в продуктивності (ціль $\geq 1.5\times$ порівняно з сильними базовими QR/DC), зберігаючи рівень ортогональності Якобі та чисельну стабільність.

Постановка задачі. Для матриці

$$A \in R^{m \times n},$$

необхідно ефективно обчислити її сингулярний розклад (SVD)

$$A = U \Sigma V^T,$$

на графічному процесорі мінімізувавши час виконання за умови суворих чисельних допусків на залишок

$$\|A - U \Sigma V^T\| \|A\|$$

та ортогональність векторних матриць

$$\|U^T U - I\|, \|V^T V - I\|.$$

Рішення має чітко визначити й реалізувати: перевірку дефляції та відповідну політику розбиття для бідіагоналі; схему оновлень з неявними зсувами під час діагоналізації; плиткову стратегію застосування обертань для ефективного оновлення накопичених U та V . Дослідження завершується еталонною реалізацією та суворою експериментальною оцінкою відносно сучасних базових реалізацій SVD Якобі, QR/DC і полярних методів на широкому діапазоні розмірів матриць і забезпечення виконання додаткових умов.

Після реалізації двоетапного SVD [5] ми перейшли до оптимізації його виконання на GPU. Перший підхід — запуск окремого ядра для кожного обертання Гівенса, що застосовується до U та V , — виявився повільнішим за монолітне ядро через значні накладні витрати від тисяч ітераційних запусків. Далі ми зосередилися на взаємодії між блоками CUDA та загальному завантаженні SM, що дало прискорення у 2–3 рази відносно початкової версії й дозволило зрівнятися з наявними базовими підходами. Наступним кроком ми досліджували розбиття великої задачі на менші: початкова стратегія ітеративного «пропускання» рядків/стовпців під час повного «обнулення» спочатку верхньої, а потім нижньої діагоналі, занадто повільно створювала нулі між потенційними підблоками. Це мотивувало перехід до неявно-зміщеного режиму: на кожному кроці застосовувати 2×2 обертання Гівенса спершу до верхньої, а тоді відразу й до нижньої діагоналі, використовуючи зсув Вілкінсона для пришвидшення збіжності. За попередніми тестами такий підхід різко зменшує кількість ітерацій і створює підґрунтя для реалізації розширеної логіки ранньої дефляції з метою подальшого розбиття матриці.

Висновки. Ми показали, що вибір гранулярності ядер критично впливає на продуктивність бідіагоналізації SVD на GPU: злиття багатьох обертань у більші ядра перевершує запуск «по одному обертанню» завдяки мінімізації накладних витрат. Поліпшення координації між блоками й підвищення завантаження SM забезпечують відчутне прискорення (у 2–3 рази) ще до зміни власне алгоритму. Заміна повних проходів «угору-вниз» на парне застосування обертання зліва/справа зі зсувом Вілкінсона істотно скорочує число ітерацій і раніше відкриває можливості для дефляції. Такий перехід не лише узгоджується з усталеними реалізаціями, а й дає практичний шлях розкладати великі задачі на незалежні підзадачі з подальшою паралельною обробкою. Запропонований дизайн відповідає сильним сторонам GPU (небагато великих ядер, плиткові GEMM-подібні оновлення U, V) при збереженні чисельної надійності. У подальшій роботі ми завершимо логіку якнайскорішої дефляції та оцінимо прискорення на великих матрицях відносно реалізацій Якобі, QR/DC і полярних методів SVD. Ми також плануємо запровадити новий «крок 0», який виконуватиметься перед основною логікою діагоналізації та застосовуватиме обертання Гівенса до найбільших за модулем елементів бідіагональної матриці, щоб отримати якнаймасивнішу діагональ.

Список джерел

1. Dongarra, J., Gates, M., Haidar, A., Kurzak, J., Łuszczek, P., Tomov, S., & Yamazaki, I. (2018). The singular value decomposition: Anatomy of optimizing an algorithm for extreme scale. *SIAM Review*, 60(4), 808–865. <https://doi.org/10.1137/17M1117732>
2. Nakatsukasa, Y., & Higham, N. J. (2013). Stable and efficient spectral divide and conquer algorithms for the symmetric eigenvalue decomposition and the SVD. *SIAM Journal on Scientific Computing*, 35(3), A1325–A1349. <https://doi.org/10.1137/120876605>
3. Tomov, S., Dongarra, J., & Baboulin, M. (2010). Towards dense linear algebra for hybrid GPU accelerated manycore systems. *Parallel Computing*, 36(5–6), 232–240.
4. Lahabar, S., & Narayanan, P. J. (2009). Singular value decomposition on GPU using CUDA. In *Proceedings of the 2009 IEEE International Symposium on Parallel & Distributed Processing* (pp. 1–10). IEEE. <https://doi.org/10.1109/IPDPS.2009.5161058>
5. Sukharskyi, S. (2026). Two-Step Singular Value Decomposition on a Single GPU with CUDA. In: Potapov, I., et al. *Digitalisation and Digital Transformation. RTC-Digital 2023*.

КАУЗАЛЬНЕ МОДЕЛЮВАННЯ ПОВЕДІНКИ КОРИСТУВАЧА У ЦИФРОВИХ СЕРВІСАХ

Білінський П.О. / Bilinskyi P.

Національний університет “Києво-Могилянська Академія” / National University of Kyiv-Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра математики
E-mail: p.bilinskyi@ukma.edu.ua

This research focuses on developing approaches to model the behavior of a customer in digital services. A conceptual framework is proposed based on Structural Causal Models along with a basic example of causal model for customer retention prediction. The problem of causal identification is outlined, and existing approaches are briefly discussed. This work defines the direction of further research and presents basic methodology for the next steps. A representation learning approach is proposed as a potential tool to learn causal relations from the temporal customer behavior data.

Каузальний аналіз є потужним інструментом для моделювання залежностей між випадковими величинами. На відміну від класичних методів статистики і машинного навчання, каузальний аналіз моделює причинно-наслідкові зв'язки між спостережуваними величинами, що дозволяє більш точно оцінювати ефект від втручань [1], а також - робить моделі більш стійкими до змін у середовищі [3]. Одним із важливих та недосліджених можливих застосувань є пояснення та передбачення поведінки користувача у цифрових сервісах. Актуальність проблеми зумовлена, з одного боку, поширенням таких сервісів (мобільні застосунки, державні портали та ін), з іншого боку - наявністю детальних даних про активність користувача та відсутності методів, що б перетворювали складні та неоднорідні дані послідовностей подій у моделі, що пояснюють поведінку користувача та допомагають оцінювати ефект від втручань та змін у сервісі.

В цій роботі окреслюється напрямок дослідження - розробка методів побудови каузальних представлень часових даних із застосуванням до аналізу поведінки користувача в цифровому продукті.

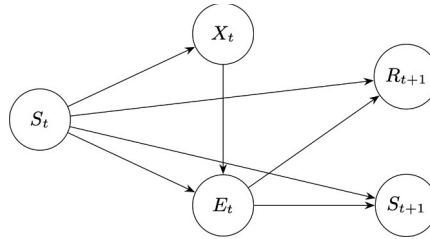
У структурній каузальній моделі [2], спостережувані випадкові величини X_1, X_2, \dots, X_N асоціюються із вершинами орієнтованого ациклічного графу (каузального графу), що описує залежності між ними у функціональній формі:

$$X_i := f_i(P A_i, U_i)$$

де $P A_i$ - вершини-батьки X_i , а U_i - випадкові шуми, незалежні між собою.

Якщо граф коректно визначений - тоді відомі відношення незалежності, що дозволяє отримати розклад сумісного розподілу $P(X_1, X_2, \dots, X_N)$ на незалежні компоненти. Маючи сумісний розподіл, можна оцінювати значення параметрів на основі спостережуваних даних.

Ми пропонуємо базову модель поведінки користувача як процесу із дискретним часом:



Де S_t - стан користувача, X_t - втручання / зміни в сервісі, E_t - рівень активності користувача, R_t - ймовірність повернення користувача в наступний момент часу.

На практиці кількість спостережуваних величин часто є значно більшою. Враховуючи це, основний виклик полягає у відшуванні, або ідентифікації каузального графу із спостережуваних даних. Важливим результатом є принципова неідентифіковність каузальної структури без додаткових припущень [2].

Враховуючи цей результат, основні підходи до відшукування каузальної структури використовують такі техніки:

2. Встановлення обмежень класу функцій f та використання статистичної незалежності для відшування каузального графу. До цих методів відноситься алгоритм Пітера-Кларка.
3. Додаткові дані - спостереження процесу при втручанні або при інших умовах. Принцип розрідженого зсуву механізмів забезпечує стабільність каузальних механізмів за різних умов [3].

Виклики та невирішені проблеми:

15. Темпоральність. Існує добре розвинена теорія каузального аналізу для стаціонарних моделей. Проте на практиці в багатьох є динамічна складова - система еволюціонує із часом, сила і тип залежностей між величинами може змінюватись.

До прикладу, ймовірність утримання користувача R_{t+1} монотонно зростає в залежності від загального часу, проведеного на сервісі (евристичне правило).

16. Побудова представлень. На практиці часто є багато спостережуваних величин, при цьому кожна окрема величина рідко має суттєву інтерпретацію, проте група величин може описувати компоненту каузального графу. До прикладу, у потоці подій окрема подія не має суттєвої значущості, але послідовність певних подій у певному інтервалі може відображати суттєву ознаку.

У подальшій роботі планується розробити метод виявлення каузальної структури процесу взаємодії користувача із цифровим сервісом та випробувати його на реальних даних. Один із можливих підходів - використання методів навчання представлень, які дозволяють відокремити латентні фактори, що відповідають причинним компонентам процесу. Зокрема, метод CausalVAE демонструє, що використання архітектури варіаційного автокодувальника із обмеженнями на незалежність і інваріантність механізмів може сприяти виявленню структурних залежностей у досліджуваному процесі [4].

5. Pearl J. *Causality: Models, Reasoning, and Inference*. – 2nd ed. – Cambridge: Cambridge University Press, 2009. – 484 p.
6. Peters J., Janzing D., Schölkopf B. *Elements of Causal Inference: Foundations and Learning Algorithms*. – Cambridge, MA: The MIT Press, 2017. – 288 p.
7. Schölkopf B., Locatello F. [та ін.] *Towards Causal Representation Learning* [Електронний ресурс] // *Proceedings of the IEEE*. – 2021. – Vol. 109, No. 5. – P. 612–634. – Режим доступу: <https://doi.org/10.1109/JPROC.2021.3058954>.

8. Yang C., Zhang Y. [та ін.] *CausalVAE: Disentangled Representation Learning via Neural Structural Causal Models* [Електронний ресурс]. – 2021. – Режим доступу: <https://doi.org/10.48550/arXiv.2106.08679>

Використання Нейромереж для Автоматичного Поновлення Даних у Спеціалізованих Предметних Областях / Application of Neural Networks for Automatic Data Updating in Specialized Subject Domains

Нагнибіда А. / Nahnybida A.

Національний університет “Києво-Могилянська Академія” / National University of Kyiv-Mohyla Academy


04655, м. Київ, вулиця Григорія Сковороди, 2, НаУКМА, Факультет інформатики
E-mail: a.nahnybida@ukma.edu.ua


The work “Application of Neural Networks for Automatic Data Updating in Specialized Subject Domains” presents a neural-network-based framework for automatic knowledge updating in specialized domains using only authoritative sources. The pipeline integrates an intake channel from official registers and APIs, deterministic version diffs and drift monitors for change detection, neural information extraction and retrieval-augmented generation for fact capture, and human-in-the-loop validation for high-risk or low-confidence cases. Authors pair continual (incremental) training with active expert reviews to reduce manual effort while preserving primary sources and traceability. Automated factuality checks re-verify model outputs against sources before committing versioned updates to a knowledge base. The approach targets lower staleness and compliance risk while maintaining explainability and auditable end-to-end operations.

Робота “Використання Нейромереж для Автоматичного Поновлення Даних у Спеціалізованих Предметних Областях” присвячена дослідженню нейромережових підходів до безперервного оновлення знань у спеціалізованих областях на основі офіційних джерел із застосуванням каналу надходження даних, виявлення змін, RAG-витягу фактів і валідації людиною (human-in-the-loop) для подолання інформаційного перевантаження.

У спеціалізованих областях інформація швидко застаріває через швидке зростання кількості джерел і публікацій; через обмеженість людських ресурсів вичерпне охоплення релевантної інформації та її своєчасна актуалізація в базах знань є практично недосяжними. Пропонується підхід до автоматичного поновлення даних на основі нейромереж, які навчаються на академічних і офіційних матеріалах, з подальшою валідацією людиною (human-in-the-loop).

Спеціалізовані предметні області швидко змінюються: виходять нові клінічні протоколи, оновлюються статті законів і підзаконні акти; у фінансовому секторі переглядаються ліміти та кредитні ставки. Ручне відстеження змін є дорогим і схильним до помилок, що призводить до хибних рішень і невідповідності нормам. Мета цієї роботи - описати практичну нейромережову архітектуру для безперервного поновлення даних із гарантованою простежуваністю першоджерела та валідацією, де машина автоматизує рутину, а експерт підтверджує критичні зміни.

Для навчання та оновлень використовуються лише авторитетні джерела: офіційні реєстри, керівні протоколи й стандарти, рецензовані наукові публікації, новинні стрічки регуляторів/інституцій, а також офіційні API. За форматом це структуровані (таблиці, реєстри), напівструктуровані (PDF, DOCX), і неструктуровані тексти (статті, постанови). Таке обмеження джерел мінімізує ризики LLM галюцинацій і полегшує аудит походження даних. 

На рівні обробки даних система використовує нейромережові компоненти на кожному ключовому етапі: збирання з офіційних джерел через сканування вебресурсів, API та вебхуки супроводжується нормалізацією, очищенням і дедуплікацією, після чого зміни виявляються як детермінованим порівнянням версій (diff), так і детекторами дрейфу, які поєднують статистичні та нейромережові підходи. Для витягу фактів застосовуються модулі інформаційного вилучення на базі глибоких моделей (розпізнавання сутностей і зв'язків) та великі мовні моделі у зв'язці з ретривером (RAG): LLM генерує твердження з посиланнями на джерела й оцінкою впевненості; ті твердження, що не проходять поріг, автоматично передаються на активне рецензування експертом. Це зменшує навантаження на людей і прискорює поновлення даних без втрати контрольованості процесу. 

Валідація даних поєднує формальні доменні правила (цілісність, діапазони, міжнормові залежності) з автоматизованою перевіркою фактичності на основі пошуку й повторної оцінки узгодженості LLM-висновків із першоджерелами; критичні або неоднозначні зміни ескалюються експерту, а вся історія фіксується в журналі аудиту. Такий підхід мінімізує ризики галюцинацій мовних моделей і підвищує простежуваність оновлень.

Для нейромереж застосовується поетапне безперервне навчання з методами регуляризації, що стримують «забування»; передбачено періодичне повторне тренування на основному наборі та цільове донавчання під нові типи документів. Щоб зменшити обсяг ручної праці, використовується навчання з активним відбором прикладів: модель самостійно пропонує найінформативніші випадки для перевірки фахівцем. Достовірність тверджень контролюється перевіркою з опором на пошук джерел: кожен факт повторно оцінюється на узгодженість із першоджерелами, а виявлені невідповідності позначаються як ризикові.

Якість та офіційність джерел критично впливають на точність; LLM схильні до галюцинацій, тому в системі обов'язково включені механізми посилення на першоджерела, оцінки невизначеності та ручного затвердження для високоризикових змін. Для захисту приватності запропоновані знеособлення (деідентифікація) персональних даних, чіткі правила зберігання інформації та контроль доступу. Достовірність тверджень забезпечується простежуваністю від кожного факту до першоджерела й збереженням проміжних матеріалів (порівнянь версій, витягів). Для безпеки процесу включені протоколювання та ведення журналів подій, можливість повернення до попередніх версій, а також визначений порядок передавання ризикованих оновлень на розгляд експерту. Узагальнені моделі можуть гірше переноситися між предметними областями; це компенсується налаштуванням на спеціалізовану предметну область і перевірками на виявлення ризикованих оновлень у даних.

Автором запропоновано практичну схему автоматичного поновлення знань у вузьких предметних областях: збір і оброблення даних - виявлення змін - витяг і перевірка фактів - оновлення у версійованому шарі знань - ручна експертна перевірка. Вже автоматизовано пошук, нормалізацію, порівняння версій та початкове вилучення фактів; ручними залишаються підтвердження критичних змін і розв'язання суперечностей. Подальша робота - розширення кола офіційних джерел і програмних інтерфейсів доступу, поліпшення виявлення ризикованих змін у даних, посилення перевірок узгодженості (самоузгодженість і звірка за кількома джерелами), а також стандартизація показників актуальності й повноти оновлень.

Список джерел:

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
2. Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., & Wermter, S. (2019). Continual lifelong learning with neural networks: A review. *Neural Networks*, 113, 54–71.
3. Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM Computing Surveys*, 46(4), 44.
4. Settles, B. (2009). Active Learning Literature Survey. University of Wisconsin–Madison, CS Tech Report 1648.
5. Sculley, D., Holt, G., Golovin, D., et al. (2015). Hidden technical debt in machine learning systems. *NeurIPS*.

ПСЕВДООБЕРНЕНА МАТРИЦЯ МУРА-ПЕНРОУЗА / MOORE-PENROSE PSEUDOINVERSE MATRIX

Кравчук О.М., Крюкова Г.В. / Kravchuk O., Kriukova G.

Національний університет «Києво-Могилянська Академія» / National University of Kyiv-Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра математики

E-mail: o.kravchuk@ukma.edu.ua, kriukovagv@ukma.edu.ua

The Moore-Penrose pseudo-inverse is a foundational concept in modern numerical linear algebra, offering a principled approach to solving ill-posed and inconsistent systems arising in machine learning and other fields. This paper explores the pseudo-inverse from five distinct perspectives — axiomatic, variational, regularization, spectral, and algebraic graph theory — highlighting its theoretical depth and practical relevance across disciplines such as machine learning, signal processing, and network analysis.

1. The Axiomatic Perspective

The Moore-Penrose pseudo-inverse of a matrix A , denoted A^{+} , is uniquely defined by four conditions:

- $AA^{+}A = A$
- $A^{+}AA^{+} = A^{+}$
- $(AA^{+})^{\top} = AA^{+}$
- $(A^{+}A)^{\top} = A^{+}A$

These conditions ensure that A^{+} behaves as a generalized inverse that respects both the structure of A and the symmetry of the solution space. The axiomatic definition is particularly powerful because it applies to any matrix, regardless of rank or shape, and guarantees uniqueness.

2. The Variational Perspective

From an optimization standpoint, the pseudo-inverse provides the minimum-norm solution to the least-squares problem. For an inconsistent system $Ax = b$, the solution $x = A^{+}b$ minimizes the Euclidean norm $\|x\|$ among all solutions that minimize $\|Ax - b\|$. This is especially relevant in machine learning, where overdetermined systems arise in regression and classification tasks, and regularization is often used to prevent overfitting.

3. The Regularization Perspective

The pseudo-inverse can be interpreted as the limit of regularized solutions to inverse problems. In particular, Tikhonov regularization solves:

$$x_{\lambda} = \arg \min_x \{ \|Ax - b\|^2 + \lambda \|x\|^2 \}.$$

As $\lambda \rightarrow 0$, the solution $x_{\lambda} \rightarrow A^{+}b$. This perspective connects the pseudo-inverse to functional analysis and reproducing kernel Hilbert spaces (RKHS), where regularization plays a central role in ensuring stability and generalization. The RKHS framework also allows for kernelized versions of the pseudo-inverse, enabling nonlinear generalizations.

4. The Spectral Perspective

The Singular Value Decomposition (SVD) offers a computationally stable and insightful method for constructing the pseudo-inverse. Given $A = U\Sigma V^{\top}$, where diagonal matrix Σ contains the singular values, the pseudo-inverse is: $A^{+} = V\Sigma^{+}U^{\top}$.

Here, Σ^{+} is formed by inverting the non-zero singular values. This decomposition is crucial for numerical stability, especially when dealing with ill-conditioned matrices. Moreover, the SVD reveals the rank and null space structure of A , which are essential in dimensionality reduction techniques like Principal Component Analysis (PCA).

5. The Algebraic Graph Theory Perspective

In graph theory, the Moore-Penrose inverse of the Laplacian matrix L of a connected graph has a rich combinatorial interpretation. The entry $(L^{+})_{ij}$ corresponds to the effective resistance between vertices i and j , a concept rooted in electrical network theory. This connection enables applications in network analysis, such as measuring centrality, robustness, and flow efficiency. The pseudo-inverse thus bridges linear algebra with discrete structures, offering tools for analyzing complex systems.

6. Applications

The Moore-Penrose pseudo-inverse is not only a theoretical construct but also a practical tool with widespread applications across scientific and engineering disciplines. Its versatility stems from its ability to handle non-square, rank-deficient, and ill-conditioned matrices, which frequently arise in real-world problems.

6.1 Machine Learning and Data Science

In supervised learning, particularly linear regression, the pseudo-inverse provides a closed-form solution for the optimal weights for design matrix and target vector. This is especially useful when the design matrix is not full rank or when the number of features exceeds the number of samples (underdetermined systems). Although gradient-based methods dominate large-scale learning, the pseudo-inverse remains essential for understanding model behavior and for initializing iterative algorithms. In applications like collaborative filtering (e.g., recommendation systems), the pseudo-inverse aids in estimating missing entries by leveraging low-rank structure via SVD.

6.2 Signal Processing and Control Theory

In signal reconstruction and system identification, the pseudo-inverse helps recover signals from incomplete or noisy measurements. For example, in compressed sensing, it can be used to approximate solutions when sparsity constraints are relaxed. In control theory, the pseudo-inverse is used to compute the least-squares solution for actuator commands in redundant robotic systems, ensuring minimal energy usage or joint movement.

6.3 Quantum Computing and Physics

In quantum mechanics, the pseudo-inverse appears in the context of generalized measurements and state reconstruction. It is used to define the inverse of non-unitary operators and to solve systems arising from quantum tomography. The spectral perspective via SVD aligns well with the decomposition of quantum states and operators, providing a bridge between linear algebra and quantum theory.

6.4 Graph-Based Learning and Network Analysis

The algebraic graph theory perspective opens doors to applications in network science. Effective resistance, derived from the pseudo-inverse of the Laplacian, serves as a metric for connectivity and robustness. In semi-supervised learning on graphs, the pseudo-inverse helps propagate labels across nodes by solving Laplacian-based optimization problems.

Conclusion

The Moore-Penrose pseudo-inverse is more than a computational tool — it is a conceptual bridge linking diverse mathematical domains. Whether viewed through the lens of axioms, optimization, regularization, spectral theory, or graph combinatorics, its versatility and depth make it indispensable in both theory and practice. As data-driven disciplines continue to evolve, the pseudo-inverse remains a cornerstone for solving systems that defy classical assumptions.

ГРА З НЕПОВНОЮ ІНФОРМАЦІЄЮ НА ПРИКЛАДІ СПОРТИВНОГО БЕТІНГУ / A GAME WITH INCOMPLETE INFORMATION: THE CASE OF SPORTS BETTING

Кущалаба Н.В., Чорней П.К. / Kutsalaba N., Chornei R. / This work was partly supported by a grant from the Simons Foundation International [SFI-PD-Ukraine-00014577, O.G.]

Національний університет “Києво-Могилянська Академія” / National University of Kyiv-Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра математики
E-mail: n.kutsalaba@ukma.edu.ua, r.chornei@ukma.edu.ua

This study presents a predictive framework for football match outcomes that integrates the Elo rating system, multiple linear regression, and a Bayesian game-theoretic interpretation of sports betting under incomplete information. Match statistics and bookmaker odds were collected via web scraping from Sofascore. The regression model achieved an adjusted $R^2 = 0.3493$, with significant predictors including shots on target, total chances, and total shots. Bookmaker accuracy was 22.37% for over/under goals and 28.33% for win/loss outcomes, revealing market inefficiency. Modeling the interaction between the bettor and the bookmaker as a Bayesian game with asymmetric information allows identifying value bets and optimizing decision-making strategies under uncertainty.

У сучасному спортивному аналітичному середовищі прогнозування результатів матчів є однією з ключових задач, що поєднує методи статистики, машинного навчання та теорії ігор.

Спортивний беттинг, зокрема футбольний, може розглядатися як взаємодія раціональних агентів — команд, аналітиків та букмекерів — які приймають рішення в умовах неповної та асиметричної інформації. У даному дослідженні запропоновано поєднати класичні статистичні методи прогнозування з теоретико-ігровою інтерпретацією через концепцію Баєсівських ігор (Bayesian games).

Через обмеженість наявних відкритих датасетів дані були зібрані самостійно з платформи *Sofascore* за допомогою технологій веб-скрапінгу [1]. Отримані масиви містять статистику матчів, коефіцієнти букмекерів і показники гравців, що дозволяє проводити комплексний кількісний аналіз.

Для первинного прогнозування використовувався рейтинг Elo, розроблений Арпадом Ело для оцінки сили гравців у шахах [2]. Цей рейтинг забезпечує можливість не лише ранжувати команди, але й оцінювати ймовірності результатів матчів. Його базова форма задається як

$$E_A = \frac{1}{1 + 10^{(R_B - R_A)/400}},$$

де E_A — очікуване число очок для команди A , R_A , R_B — попередні рейтинги команд. Новий рейтинг визначається за формулою

$$R'_A = R_A + K(S_A - E_A),$$

де K — коефіцієнт регуляції, S_A — набрані очки.

Для аналізу факторів, що впливають на кількість забитих голів, застосовано множинну лінійну регресію [3]:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon.$$

У нашій моделі прогнозованою змінною (Y) була *total_goals*, а найбільш інформативними предикторами виявилися *total_on_target*, *total_chances* та *total_shots*. Отримана модель:

$$total_{goals} = 1.40 + 0.336 \cdot total_{\square} + 0.172 \cdot total_{chances} - 0.059 \cdot total_{shots},$$

із скоригованим коефіцієнтом детермінації $R_{adj}^2 = 0.3493$, що пояснює близько 35% варіації результату. Незалежність змінних підтверджується низькими значеннями VIF (<2), що свідчить про відсутність мультиколінеарності.

Було також перевірено точність букмекерських прогнозів. Для найпоширеніших тоталів (>2.5, >3.5) середня точність склала 22.37%, а для передбачення результатів *win/loss* — 28.33%. Це свідчить про наявність значного рівня непередбачуваності у футбольних матчах.

У моделі Баєсівської гри букмекер володіє приватною інформацією $\theta_B \in \Theta$ про точність власної моделі, а гравець спостерігає лише публічний сигнал — коефіцієнти $k = (k_W, k_D, k_L)$. Ймовірність, імплікована коефіцієнтами, обчислюється як $p_{implied} = 1/k_i$, а value bet визначається умовою $p_{model} > p_{implied}$, що сигналізує про потенційну неефективність ринку [4].

Подальший аналіз включатиме перевірку гіпотези про ринкову ефективність $H_0: E[Value] = 0$, а також застосування критерію Келлі для оптимального вибору розміру ставки

$$f^{\square} = \frac{p \cdot k - 1}{k - 1},$$

і подальше тестування на історичних даних.

Отримані результати підтверджують можливість формалізації спортивного беттингу як Баєсівської гри та відкривають перспективи для побудови ефективних стратегій прогнозування і керування ризиками на основі ймовірнісних моделей.

Список літератури

1. Як використовувати web scraper для збору даних з інтернету з Python. URL: <https://dou.ua/forums/topic/43070/> (дата звернення: 23.09.2025).
2. Winston W. L., Pelechrisis K., Nestler S. *Mathletics: How Gamblers, Managers, and Fans Use Mathematics in Sports*. Princeton University Press, 2022. 432 p.
3. *Simple Linear Regression. Linear Models in Statistics*. Hoboken, NJ, USA, pp. 127–136.
4. Zou Q., Song K., Shi J. *A Bayesian In-Play Prediction Model for Association Football Outcomes*. Applied Sciences, 2020, Vol. 10, No. 8, p. 2904.

**DECISIONET З ПРОПОРЦІЙНИМ РОЗПОДІЛОМ ОБЧИСЛЮВАЛЬНИХ
ПОТУЖНОСТЕЙ / DECISIONET WITH PROPORTIONAL ALLOCATION OF
COMPUTATIONAL POWER**

Мокрий М. В., Швай Н. О. / Mokryi M., Shvai N.

Національний університет “Києво-Могилянська Академія” / National University of Kyiv-
Mohyla
Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики
E-mail: m.mokryy@ukma.edu.ua, n.shvay@ukma.edu.ua

This research focuses on DecisioNet (DN) — a Neural Decision Tree model that combines the generalization capabilities of deep networks with the computational efficiency of decision trees via conditional computation. The tree structure of the standard DN model allocates uniform computational power regardless of the number of classes per node. This study introduces Unbalanced DecisioNet (Unbalanced DN), which allocates computational power proportionally to the number of classes assigned to a node using hierarchical clustering. The results are compared with the original DN and baseline Network in Network models on CIFAR-10 and FashionMNIST datasets. Validation performed on CIFAR-10 demonstrated that Unbalanced DN improves classification accuracy compared to DN, which confirms the effectiveness of allocating increased computational resources to complex clusters while reducing resources for simple clusters.

У цьому дослідженні розглянуто нейронну мережу з бінарною деревоподібною структурою DecisioNet (DN) [1], яка належить до категорії нейронних дерев рішень [2] та представлено нову версію моделі з пропорційним розподілом обчислювальних ресурсів.

Традиційні нейронні мережі обробляють кожний вхідний зразок через усі параметри моделі, тоді як модель DN використовує умовні обчислення. Кожний вхідний зразок проходить по одному шляху від кореневого вузла до одного з листових вузлів дерева, активуючи лише частину від загальної кількості параметрів під час виведення. Класи призначаються вузлам на основі ієрархічної кластеризації за матрицею плутанини попередньо навченої базової моделі. Модель DecisioNet використовує архітектуру нейронної мережі Network in Network (NiN) [3].

Під час навчання використовується комбінована функція втрат, яка складається з функції втрат класифікації L_{cls} та маршрутизації L_{σ} , в якій параметр β контролює ступінь впливу над маршрутизацією вхідного зразка по дереву:

$$L = L_{cls} + \beta \cdot L_{\sigma}$$

У процесі навчання застосовується метод покращеного семантичного хешування [4], який поєднує диференційовану та дискретну маршрутизацію у рівних пропорціях. Це забезпечує зворотне поширення під час навчання, зберігаючи при цьому ефективність виведення завдяки вибору одного маршруту від кореня дерева до його листового вузла.

Структура збалансованого дерева моделі DN рівномірно розподіляє обчислювальні ресурси незалежно від розподілу класів між вузлами дерева. Тому з метою підвищення ефективності використання обчислювальних ресурсів було запропоновано підхід, за якого вони розподіляються пропорційно до обсягу даних, що обробляються на кожному вузлі.

Оновлена версія DN — UnbalancedDecisioNet (Unbalanced DN), реалізує нерівномірний розподіл обчислювальних ресурсів, дозволяючи виділяти більше ресурсів для складніших підзадач класифікації. Розподіл обчислювальних ресурсів для вузла i визначається за допомогою формули:

$$C_i = C_l \cdot 2^l \cdot \frac{n_i}{N}$$

Де N — загальна кількість класів, n_i — кількість класів призначених вузлу i та C_l — базова кількість параметрів для рівня дерева l у конфігурації DN.

Валідація створеної моделі була проведена на наборах даних FashionMNIST та CIFAR-10. Було натреновано базову модель NiN, а також моделі DN2 та Unbalanced DN2, де індекс «2» позначає глибину дерева.

Застосований метод ієрархічної кластеризації демонструє логічну структуру розподілу даних. Для набору даних CIFAR-10 початковий розподіл на першому рівні дерева ділить дані на два кластери: тварин та транспортні засоби. Для FashionMNIST перший рівень розподілу відокремлює взуття від одягу. Структуру дерева моделі Unbalanced DN2 для набору даних CIFAR-10, а також розподіл класів і обчислювальних ресурсів відповідно до ієрархічної кластеризації, наведено на рисунку 1.

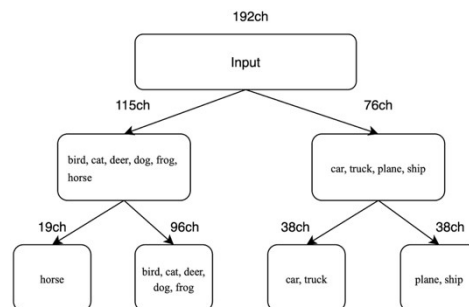


Рисунок 1. Структура моделі Unbalanced DN2 для CIFAR-10

Загальні результати експериментів наведено в таблиці 1. Вони демонструють, що моделі DecisioNet досягають подібної до базової моделі NiN результативності, використовуючи меншу кількість параметрів під час навчання, і ще меншу — під час виведення. Модель Unbalanced DN2 продемонструвала покращення точності класифікації на наборі даних CIFAR-10.

Таблиця 1. Загальні результати класифікації

Модель	Набір даних	Точність класифікації, %	К-сть параметрів
Baseline NiN	CIFAR-10	88.39	966,986
DN2	CIFAR-10	85.97	736,907
Unbalanced DN2	CIFAR-10	87.18	739,519
Baseline NiN	FashionMNIST	92.40	957,386
DN2	FashionMNIST	92.80	727,307
Unbalanced DN2	FashionMNIST	92.81	820,891

Емпірична валідація результатів моделі Unbalanced DN була проведена на наборах даних CIFAR-10 та FashionMNIST. Вона продемонструвала, що залежно від характеристик набору даних, складні вузли показують покращення точності класифікації за умов збільшення обчислювальних ресурсів, тоді як прості вузли зберігають високу результативність навіть при їх зменшенні.

Робота демонструє, що архітектура нейронного дерева рішень може включати нерівномірний розподіл обчислювальних ресурсів, відповідно до складності задачі класифікації, виходячи за межі рівномірного розподілу ресурсів по дереву рішень. У подальших дослідженнях планується перевірити ефективність моделі Unbalanced DN на більш складних наборах даних, зокрема CIFAR-100, а також дослідити вплив збільшення глибини архітектури DN на її ефективність.

Список джерел:

- Li H., Song J., Xue M., Zhang H., Ye J., Cheng L., Song M. A survey of neural trees // arXiv preprint arXiv:2209.03415. — 2022.
- Gottlieb N., Werman M. DecisioNet: A binary-tree structured neural network // Proceedings of the Asian conference on computer vision. — 2022. — P. 1675-1690.

3. Lin M., Chen Q., Yan S. Network in network // arXiv preprint arXiv:1312.4400. — 2013.
4. Kaiser Ł., Bengio S. Discrete autoencoders for sequence models // arXiv preprint arXiv:1801.09797. — 2018.

**МЕТОДИКА ТРЕНУВАННЯ ГЛИБОКОЇ НЕЙРОННОЇ МЕРЕЖІ DEEPLABV3
ДЛЯ СЕМАНТИЧНОЇ СЕГМЕНТАЦІЇ ОБ'ЄКТІВ НА СУПУТНИКОВИХ
ЗОБРАЖЕННЯХ / METHODOLOGY FOR TRAINING DEEPLABV3 DEEP
NEURAL NETWORK FOR SEMANTIC SEGMENTATION OF OBJECTS IN
SATELLITE IMAGES**

Колінько П.В./ Kolinko P.V.

Національний університет “Києво-Могилянська Академія”/ National university Kyiv-
Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, каф. мультимедійних систем

E-mail: pavlo.kolinko@ukma.edu.ua

Semantic segmentation of satellite imagery remains one of the most challenging tasks in computer vision, requiring accurate pixel-level classification of diverse objects such as vegetation, buildings, and terrain features. DeepLab v3 is a state-of-the-art deep learning architecture with about 60-70M parameters to capture multi-scale contextual information. In this research, we proposed and tested a comprehensive training methodology for DeepLab integrated with PointRend head, utilizing ResNet-101 as backbone with layer-based learning rates and a gradual layer unfreezing strategy for the whole backbone. A composite loss function that is composed by weighted class-dependent loss metrics. Proposed approach allows to achieve the optimal results, prevent model overfitting and accelerate model convergence.

Семантична сегментація є фундаментальною та найобширнішою задачею комп'ютерного зору, що полягає у класифікації кожного пікселя зображення відповідно до його класу, на виході формуючи так звану “маску”. Застосування семантичної сегментації до супутникових знімків супроводжується специфічними викликами: надвисока роздільність зображень, велика варіативність масштабів об'єктів (для дерев - як поодинокі дерева, так і лісові масиви, для водойм - невеликі затоки або широкі озера), атмосферні перешкоди, сезонні зміни забарвлення рослинності, та значний дисбаланс класів. Проте, семантична сегментація супутникових зображень має дуже широке застосування. Вона охоплює моніторинг сільськогосподарських культур та оцінку врожайності, лісових ресурсів, міське планування (розпізнавання забудови) та аналіз змін землекористування, екологічний моніторинг водойм та прибережних зон, тощо. Успішне вирішення цієї задачі вимагає не лише потужних мереж глибокого навчання, але й спеціалізованих методів тренування, що враховують унікальні характеристики супутникових даних та специфіку цільових об'єктів сегментації.

Для вирішення такої задачі семантичної сегментації використовують великі згорткові нейронні мережі, розмірність яких нерідко доходить до 90+M параметрів, в залежності від архітектури. Прикладом такої глибокої нейронної мережі є Deeplabv3 [1]. Вона складається з таких ключових компонентів: CNN backbone, такий як ResNet з atrous convolutions, та модуля Atrous Spatial Pyramid Pooling (ASPP). Основною перевагою DeepLab над традиційними encoder-decoder архітектурами (як FCN або SegNet) є збереження високої роздільності feature maps завдяки атрозним згорткам, що дозволяє уникнути втрати просторової інформації. Порівняно з архітектурами на базі трансформерів, таких як наприклад Segment anything model (SAM), DeepLab [2] демонструє меншу обчислювальну складність та вимоги до пам'яті завдяки локальній природі згорткових операцій.

Використання backbone мережі ResNet-101 з попереднім навчанням на масштабному датасеті ImageNet (1.2 мільйона зображень, 1000 класів) забезпечує

модель надійними даними для низькорівневого виділення загальних ознак. Критичною модифікацією архітектури є налаштування є заміною stride-based downsampling в layer3 та layer4 на atrous convolutions, що зменшує крок нейронної мережі і дозволяє зберігати більшу роздільну здатність features, що критично важливо для точної локалізації об'єктів на супутникових знімках з високою деталізацією [3]. До шару layer3 базової мережі доданий допоміжний класифікатор FCNHead (auxiliary classifier), забезпечуючи пряме навчання з учителем для ознак середнього рівня та полегшуючи поширення градієнтів назад через глибоку мережу. Також, архітектура мережі доповнена PointRend [4], що дозволяє відбирати точки, поєднуючи низькорівневі з високорівневими передбаченнями, що мають найбільшу впевненість.

Основний запропонований підхід є фрагментарне тренування. І хоча загальний план тренування передбачає викривання CosineAnnealingLR із попереднім розігрівом за допомогою LinearLR, кожен із зазначених шарів нейронної мережі, що описаний вище, має власний початковий learning rate для корекції швидкості навчання. До прикладу, класифікатор нейронної мережі DeeplabHead має більш швидкий та агресивніший план навчання (~0.0001), порівняно з backbone мережі (~0.00005). Також, ResNet-101 backbone поступово розморожується відносно передбаченого плану, після проходження тренування певної кількості епох. Таким чином, відбувається поступова, “м’яка” розморозка шарів, починаючи від верхнього шару layer4, розморозка якого відбувається після проходження ~30% епох, до найнижчого шару, дотренування якого починається після ~85% епох. Останнім важливим введенням є використання комбінованої функції втрат на основі зваженої покласово кросс-ентропії. Перед початком тренування обраховується частота входжень класу на тренувальному датасеті, що дозволяє сильніше штрафувати модель за помилки класу, що зустрічається найменше. До зваженої Cross-Entropy по передбаченням класифікатора додається значення зваженої кросс-ентропії по auxiliary classifier, PointRend [4] та середньозважене mIoU з різними ваговими коефіцієнтами.

Таким чином, запропонований підхід покращує якість результатів після тренування на невеликих за обсягом датасетах, але має ряд недоліків, такий як збільшення часу тренування однієї епохи, ризик потрапляння на “плато” на нерізноманітних датасетах.

Список використаних джерел

1. Thomas, A. J., & Boerkamp, C. (2025). Enhanced DeepLab based nerve segmentation with optimized tuning. arXiv. <https://doi.org/10.48550/arXiv.2507.13394>
2. Matthew, B., Genc, B., Huang, Y., Belter, D., & Tang, Z. (2024, June 1). Comparative analysis of popular semantic segmentation architectures: U-Net, DeepLab, and FCN. ResearchGate. <https://www.researchgate.net/publication/394470887>
3. Gao, X., Chen, C., Gan, Y., and Liu, Y.: Comparison of Extraction accuracy of Sugarcane from different resolution satellite images using Deep lab V3+ Mode, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-1-2024, 179–184, <https://doi.org/10.5194/isprs-archives-XLVIII-1-2024-179-2024>, 2024.
4. Max Reinhardt, Gregor Geigle, Radu Timofte, and Goran Glavaš. 2024. Improving Vision-Language Cross-Lingual Transfer with Scheduled Unfreezing. In *Proceedings*

**ПРОЄКТУВАННЯ ІНТЕЛЕКТУАЛЬНОГО АГЕНТА ДЛЯ АВТОНОМНОЇ
ОПТИМІЗАЦІЇ МІКРОСЕРВІСІВ У КОРПОРАТИВНИХ СИСТЕМАХ / DESIGN OF AN
INTELLIGENT AGENT FOR AUTONOMOUS MICROSERVICE OPTIMIZATION IN
ENTERPRISE SYSTEMS**

Ванін Д.О. / Vanin D.

Національний університет “Києво-Могилянська Академія” / National University of Kyiv-Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики
E-mail: danylo.vanin@ukma.edu.ua

This study introduces a conceptual framework for an intelligent agent designed to autonomously optimize microservice-based enterprise systems. Microservice architecture [3] serves as the basis for scalable and modular enterprise solutions, supporting agile DevOps practices and enabling independent deployment, updating, and management of services. While these features accelerate release cycles and improve maintainability, they also create persistent operational and architectural challenges [4, p. 633-634] that often necessitate manual intervention. These issues are especially significant in corporate settings, where requirements are dynamic and influenced by organizational, legal, and economic factors. In response, the software engineering field is increasingly exploring AI-driven approaches [1, p. 12]. Although machine learning has been utilized for anomaly detection [2], auto-scaling, and traffic routing, there remains a gap in the development of fully autonomous agents capable of intelligent architectural analysis and modification. The proposed agent addresses these challenges through a modular architecture that includes monitoring, decision-making, execution, and safety verification modules. The decision component employs reinforcement learning and formal quality models to balance objectives such as performance, cohesion, and latency, while maintaining adherence to system constraints. This design seeks to reduce architectural drift, improve service modularity, and support adaptive system behavior in real time. The agent is intended as a foundational element for intelligent DevOps and self-optimizing enterprise software systems.

Architectural decisions within microservice systems exert long-term effects on performance, reliability, and maintainability, yet are frequently made manually or rely on limited metrics. Static rule-based approaches are insufficient in environments with fluctuating workloads, evolving services, and variable infrastructure constraints. Current solutions often focus on optimizing isolated objectives, such as CPU utilization or response time [1, p. 13-14], while overlooking critical architectural quality attributes, including cohesion, coupling, and transactional autonomy. There is a demonstrated need for an autonomous agent capable of holistic analysis of service systems and optimization of architecture-level quality within operational boundaries. Such an agent should learn from historical actions, adapt to evolving patterns, and ensure safe operation within enterprise constraints.

The methodology initiates with formal modeling of the system as a dynamic service dependency graph, augmented by runtime telemetry. The agent derives a state vector of quality attributes from this graph, including latency (measured as average or 95th-percentile response time across call paths), service cohesion (functional similarity of operations within a service), coupling (frequency of inter-service calls), and deployment complexity. A multi-objective reward function is constructed based on these attributes. The microservice optimization problem is formalized as a sequential decision-making process under uncertainty. Each system state is described by a high-dimensional vector of features, including latency distributions, cohesion/coupling indicators, service error rates, and deployment metadata. The decision module selects an action $a_t \in A$ based on the current state $s_t \in S$, and receives feedback via a reward function R_t , computed as

$R_t = w_1 \cdot Q_{\text{perf}}(s_t, a_t) + w_2 \cdot Q_{\text{cohesion}}(s_t) - w_3 \cdot Q_{\text{coupling}}(s_t) - w_4 \cdot Q_{\text{complexity}}(s_t)$, where Q_{perf} evaluates latency/throughput, Q_{cohesion} and Q_{coupling} quantify architectural quality, and $Q_{\text{complexity}}$ penalizes excessive service granularity or communication overhead. The w_i encode system-specific priorities. This formalization enables reinforcement learning algorithms to optimize not only for performance but also for maintainability and scalability.

The agent's core logic is implemented using a reinforcement learning framework, specifically a model-free, on-policy algorithm such as Proximal Policy Optimization (PPO). This approach enables the agent to learn which architectural actions, including splitting, merging, or relocating services. The reward model integrates domain-specific constraints and preferences, such as latency thresholds, availability requirements, cost limitations, and adherence to data locality policies.

The system architecture comprises four loosely coupled modules. The observation module collects real-time metrics and dependency data. The decision module processes the current system state and selects actions based on its learned policy. The execution module implements changes to the infrastructure through DevOps APIs. The safety module conducts verification of proposed changes to ensure compliance with requirements. The agent interfaces with its environment via programmable endpoints, including Kubernetes APIs. This design supports both assistive and autonomous operation. Figure 1 illustrates the overall architecture of the intelligent agent.

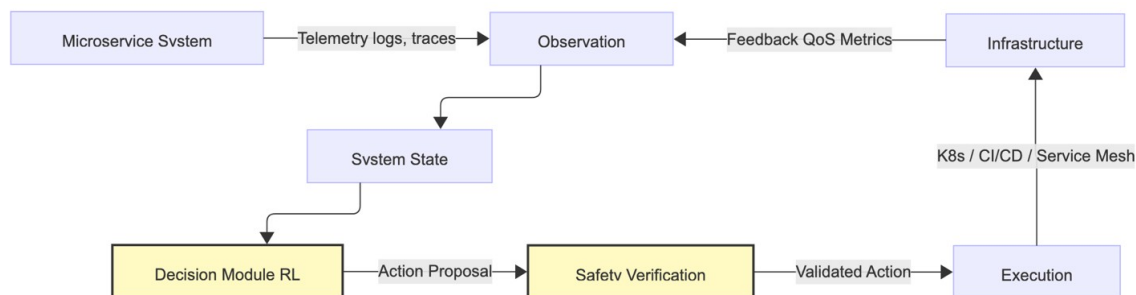


Figure 1. Architecture of the intelligent agent. Components: (1) Observation layer; (2) RL-based Decision module; (3) Execution interface (e.g. Kubernetes API); (4) Safety Validator; (5) Feedback from environment.

The proposed system is based on ongoing dissertation research dedicated to developing self-optimizing intelligent agents for microservice-based software architectures. This architecture is expected to enable the agent to mitigate architectural drift and optimize quality attributes over time. The framework will be assessed in a controlled simulation environment using microservice benchmarks such as Online Boutique [5], comparing agent-driven adaptations to static baselines. Anticipated results include demonstrating that an intelligent agent can serve as an effective decision-support tool or autonomous controller in complex, evolving enterprise systems.

Список джерел:

1. AI techniques in the microservices life-cycle: a systematic mapping study [Electronic resource] / Sergio Moreschini [et al.] // Computing. – 2025. – Vol. 107, no. 4. – P. 100. – Mode of access: <https://doi.org/10.1007/s00607-025-01432-z>. – Title from screen.
2. GAL-MAD: towards explainable anomaly detection in microservice applications using graph attention networks [Електронний ресурс] / Lahiru Akmeemana [та ін.]. – [Б. м. : б. в.], 2025. – (Препринт). – Режим доступу: <https://arxiv.org/abs/2504.00058>. – Назва з екрана.
3. Lewis J. Microservices [Електронний ресурс] / James Lewis, Martin Fowler // martinfowler.com. – Режим доступу: <https://martinfowler.com/articles/microservices.html> (дата звернення: 31.10.2025). – Назва з екрана.
4. Microservices: architecture, container, and challenges [Електронний ресурс] / Guozhi Liu [та ін.] // 2020 IEEE 20th international conference on software quality, reliability and security companion (QRS-C). – [Б. м.], 2020. – С. 629–635. – Режим доступу: <https://doi.org/10.1109/QRS-C51114.2020.00107>. – Назва з екрана.
5. GoogleCloudPlatform. Microservices-Demo: sample cloud-first application with 10 microservices showcasing kubernetes, istio, and GRPC [Електронний ресурс] / GoogleCloudPlatform. – Режим доступу: <https://github.com/GoogleCloudPlatform/microservices-demo> (дата звернення: 01.11.2025). – Назва з екрана.

ВИКОРИСТАННЯ MIXTURE OF EXPERTS З LLM-АГРЕГАТОРОМ ДЛЯ ТОРГОВЕЛЬНОГО БОТА / MIXTURE OF EXPERTS WITH LLM AGGREGATION FOR TRADING BOT

Салата К.В. / Salata K.

Національний університет “Києво-Могилянська Академія” / National University of Kyiv-Mohyla
Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики

E-mail: kirill.salata@ukma.edu.ua

This research focuses on the architecture of intelligent trading bot, based on the Mixture of Experts (MoE) combined with a large language model (LLM)- based aggregator. Within this system several specialized expert models generate forecasts regarding trading decisions, and the LLM aggregator analyzes the results of these experts' work and makes a final decision. The proposed architecture is expected to improve the accuracy and explainability of trading decisions through dynamic expert selection and multimodal analysis, providing better adaptation to complex market conditions compared to traditional approaches.

Традиційні моделі машинного навчання досить часто стикаються з проблемою перенавчання та неможливістю урахування ринкових умов, через те що фінансові ринки характеризуються високою волатильністю та великою кількістю шуму у даних. Останнім часом для вирішення цих проблем все більше уваги звертають на архітектуру Mixture of Experts (MoE), яка містить суміш експертів, які можуть підвищити узагальнюючу здатність за рахунок використання декількох спеціалізованих експертних підсистем[1]. В останні роки активно досліджуються різні методи прогнозування часових рядів фінансових ринків. Найпопулярнішими є методи рекурентних нейронних мереж (RNN) та навчання з підкріпленням (RL). Однак, як показано в дослідженнях [2], [3], методи RNN мають труднощі в адаптації до мінливих умов і погану інтерпретованість прогнозів, структура LSTM не може розпізнати всі зміни на ринку, і проблема високої волатильності залишається невирішеною[2]. Припускають, що проблему високої волатильності можна вирішити за допомогою різних рішень, але ці рішення не можуть бути використані для впровадження в торгові системи, оскільки точність прогнозів значно варіюється в часі[4]. Тому проблеми високої волатильності та інтерпретованості залишаються відкритими. Це вимагає подальших досліджень архітектур для фінансових ринків. Застосування MoE вже продемонстрував перевагу над такими моделями в задачах фінансового прогнозування та алгоритмічного трейдингу за рахунок різнорідних експертів[1].

Наприклад, робота Mixture of Experts with Group Aggregation (MIGA) показала, що динамічне перемикання експертів, які спеціалізуються на різних стилях, значно покращує точність прогнозування порівняно з одиночними моделями[5]. Було запропоновано механізм групової агрегації, де експерти розбиваються на групи, всередині яких реалізовано механізм спільної уваги для обміну інформацією між експертами однієї групи. На вході відносно простий нейромережевий роутер, який переключає між групами експертів. Після роботи експертів результати агрегуються та робиться прогноз.

Враховуючи успіхи описаних підходів, в роботі пропонується архітектура торговельного бота, який буде розроблено на базі MoE та LLM-агрегатору. Новизна такого підходу у використанні LLM у якості модуля, який аналізує сигнали від різних моделей та вирішує чи треба входити у позицію на ф'ючерсах. Нижче описується архітектура запропонованої моделі та принципи її функціонування:

1. Модуль обробки вхідних даних - з різних джерел проводиться збір різних даних, які можуть впливати на ціну ф'ючерсів таких, як технічні індикатори, новини, цінові ряди та інші фактори які можуть впливати. Для кожного типу даних проводиться обробка після чого дані передаються експертним моделям.
2. Набір моделей-експертів - набір різних моделей машинного навчання навчених для аналізу різних аспектів ринку. Таким чином можна поєднувати експертів, наприклад, один експерт може бути рекурентною нейронною мережею для виявлення патернів в цінових рядах, а інший експерт для сентимент-аналізу. Для того аби підвищити ефективність, можна розбити експертів на групи по типам

даних та стратегій. Велика кількість різномірних експертів забезпечує більше повне покриття можливих ринкових сценаріїв.

3. Агрегатор рішень - ключова частина запропонованої архітектури, пропонується що LLM буде отримувати на вхід результати роботи експертів, інтерпретувати результати їх роботи та приймати рішення чи потрібно виконувати торговельну операцію. LLM виступає у якості мета-експерта, який оцінює впевненість різних експертів в залежності від ситуації на ринку. Такий агрегатор може враховувати також додаткові фактори, як новини, аналізуючи їх, розпізнавати приховані залежності та ризики, які можуть не бути очевидними для числових моделей. Також важливим фактором є те, що агрегатор може створити пояснення для того чому він прийняв певне рішення, що підвищує прозорість роботи системи.
4. Модуль виконання операції - виконує операцію на біржі після прийняття рішення. Коли агрегатор приймає рішення про відкриття позиції він передає модулю виконання операції інформацію про тип позиції, стартову ціну та очікувану ціну закриття позиції. До цього модулю також входить модуль контролю ризику, за допомогою якого контролюється розмір позиції в залежності від впевненості агрегатора у позиції, стану рахунка та заданого рівня ризику. Таким чином розраховується співвідношення прибутку до ризику, на основі якого розраховується розмір позиції. Після чого модуль відправляє на біржу запит на відкриття позиції.

Запропонована архітектура торговельного бота поєднує сильні сторони спеціалізованих моделей-експертів та великих мовних моделей. Використання МоЕ забезпечує диверсифікацію аналізу за рахунок великої кількості експертів, кожна група з яких навчена на своєму типі даних для підзадачі, що підвищує стійкість системи до різних ринкових умов. Використання LLM як агрегатора дає можливість системі динамічно обирати найбільш релевантних експертів, обґрунтовувати прийняте рішення. Таким чином архітектура системи є більш адаптивною порівняно з традиційними ботами заснованими на алгоритмічних стратегіях. Очікується, що така система матиме кращі результати в дохідності та пояснюваності рішень. В майбутньому планується провести експериментальну оцінку запропонованого підходу на історичних даних ф'ючерсного ринку та порівняти їх з результатами існуючих моделей. Така архітектура поєднує методи машинного навчання з елементами міркування природною мовою, що робить торгові рішення більш надійними та пояснювальними.

Список джерел:

1. Liu K. M., Lo M. C. LLM-Based Routing in Mixture of Experts: A Novel Framework for Trading //arXiv preprint arXiv:2501.09636. – 2025.
2. Chaudhary R. Advanced Stock Market Prediction Using Long Short-Term Memory Networks: A Comprehensive Deep Learning Framework //arXiv preprint arXiv:2505.05325. – 2025.
3. Bhuiyan M. S. M., Al Rafi M. D., Rodrigues G. N., Hossain Mir M. N., Ishaq A., Mridha M. F., Shin J. Deep learning for algorithmic trading: A systematic review of predictive models and optimization strategies // Array. – 2025. – № 100390. – ISSN 2590-0056. – DOI: 10.1016/j.array.2025.100390
4. Bai Y., Zhang X., Li Y., Wang J. A review of reinforcement learning in financial applications // Annual Review of Statistics and Its Application. – 2025. – Т. 12, № 1. – С. 209–232
5. Yu Z. et al. MIGA: Mixture-of-Experts with Group Aggregation for Stock Market Prediction //arXiv preprint arXiv:2410.02241. – 2024.

КОНТРОЛЬ ІНТЕНСИВНОСТІ ЕМОЦІЙ В УКРАЇНСЬКОМУ СИНТЕЗІ МОВЛЕННЯ ЧЕРЕЗ КРОС-МОВНЕ ПЕРЕНЕСЕННЯ ЗНАТЬ ТА АДАПТЕРИ LORA / EMOTION INTENSITY CONTROL IN UKRAINIAN TTS THROUGH CROSS-LINGUAL TRANSFER KNOWLEDGE AND LORA ADAPTERS

Іващенко Д.С. / Ivashchenko D.S.

*Національний університет «Києво-Могилянська академія» /
National University of "Kyiv-Mohyla Academy" (NaUKMA)*

*04655, м. Київ, вулиця Григорія Сковороди, 2, НаУКМА, Факультет інформатики
d.ivashchenko@ukma.edu.ua*

This paper presents a novel approach to controllable emotional text-to-speech synthesis for the Ukrainian language using cross-lingual emotion transfer and Low-Rank Adaptation (LoRA) adapters.

The proposed method leverages pre-trained multilingual TTS models and enables fine-grained control over emotional intensity through interpolation of latent representations. The paper investigates the effectiveness of LoRA adapters for emotion-specific fine-tuning, which provides a parameter-efficient way to inject emotional expressiveness while maintaining the naturalness of synthesized speech. The experimental setup includes adaptation of XTTS v2 model to Ukrainian using publicly available speech datasets, followed by emotion transfer from high-resource languages. This work addresses the critical gap in emotional speech synthesis in controllable emotional TTS for Ukrainian.

Емоційна виразність є критичною характеристикою природного мовлення, що робить синтез мовлення з контрольованими емоціями важливою задачею в галузі обробки мовлення. Для української мови, яка є малоресурсною в контексті технологій синтезу мовлення, розробка таких емоційних систем залишається актуальною. Сучасні системи синтезу українського мовлення [1] зосереджені переважно на нейтральному мовленні, не забезпечуючи контролю над емоційною складовою. Крім того, відсутність великих емоційно анотованих датасетів українського мовлення ускладнює тренування моделей, здатних синтезувати емоційно виразне мовлення.

Сучасні підходи до емоційного TTS базуються на глибоких нейронних мережах, зокрема на архітектурах типу VITS [2], YourTTS [3] та XTTS, які демонструють високу якість синтезу. Однак більшість цих моделей навчені на високоресурсних мовах (англійська, мандарин) і потребують адаптації для роботи з українською мовою. Метою даного дослідження є розробка методу контрольованого емоційного синтезу мовлення для української мови з використанням крос-мовного переносу знань та легких адаптерів LoRA.

Підхід заснований на трьох ключових компонентах: адаптації базової моделі до української мови, імплементації механізму контролю емоцій та використання LoRA адаптерів для ефективного донавчання моделей.

В якості базової моделі обрано XTTS v2 (Coqui TTS), яка є оновленою версією XTTS, багатомовною моделлю з підтримкою клонування голосу та перенесення стилю з аудіозразків. Модель базується на архітектурі GPT для автоматичного модуля обробки тексту, об'єднаний з генеративно-змагальним вокодером HiFi-GAN для синтезу аудіо з проміжних представлень. XTTS v2 навчена на 17 мовах, серед яких українська не представлена.

Адаптація здійснюється через донавчання моделі на українських мовленнєвих наборах даних, що містять текст, відповідний зразок мовлення та інформацію про мовця, таку як стать чи вік. Використовуються наступні відкриті ресурси: Ukrainian TTS dataset (голоси Lada, Mykuta) [4], набір даних M-AILABS та Mozilla Common Voice Ukrainian. Загальний обсяг тренувальних даних становить близько 30 годин високоякісного мовлення. Процес адаптації включає корегування текстового кодера для правильної обробки української фонетики та декодера для генерації природного мовлення.

Low-Rank Adaptation (LoRA) [5] використовується для ефективного навчання емоційно-специфічних адаптерів. Замість повного донавчання всієї моделі, LoRA додає низькорангові матриці до ваг шару уваги (attention layer). Для кожної емоції e тренується окремий LoRA адаптер A_e з рангом $r = 16$, що складає менше 1% від загальної кількості параметрів моделі. Це дозволяє ефективно змінювати емоційні представлення та комбінувати їх через зважену суму виходів шарів адаптерів.

Контроль емоційної інтенсивності реалізується через інтерполяцію латентних представлень. Для кожної цільової емоції e (радість, смуток, гнів) створюється референсне аудіо. Модель генерує умовні латентні вектори $Z_{neutral}$ для нейтрального мовлення та Z_e для емоційного. Контроль інтенсивності здійснюється через лінійну інтерполяцію:

$$z_{mixed} = \alpha \cdot z_e + (1 - \alpha) \cdot z_{neutral},$$

де $\alpha \in [0, 1]$ контролює інтенсивність емоції. Значення $\alpha = 0$ відповідає нейтральному мовленню, $\alpha = 1$ – максимальній емоційній інтенсивності.

Оскільки існує мало емоційно анотованих українських мовленнєвих датасетів, застосовується метод перенесення знань (transfer learning) з високоресурсних мов. Використовуються емоційні датасети для англійської мови (зокрема, ESD, RAVDESS) для попереднього навчання емоційних LoRA адаптерів. Завдяки багатомовності моделі XTTS,

латентні представлення емоцій є частково мовно-інваріантними, що забезпечує ефективне перенесення знань.

Оцінка якості синтезованого мовлення визначається за об'єктивними метриками (мел-кепстральне спотворення (MCD), частка помилково синтезованих слів (WER), класифікація емоційного мовлення моделями глибокого навчання) та суб'єктивними (усереднена суб'єктивна якісна оцінка (MOS) з 5 оцінювачами, суб'єктивний тест на розпізнавання емоцій).

Попередні експерименти показують, що адаптована модель здатна генерувати природне українське мовлення з MOS > 3.5. Очікується, що механізм контролю інтенсивності забезпечить плавний перехід між нейтральним та емоційними станами з точністю розпізнавання емоцій > 75%. LoRA адаптери демонструють ефективність у використанні пам'яті при збереженні якості емоційної виразності, порівнянної з повним донавчанням.

Запропонований підхід надає практичне рішення для розробки емоційно виразних систем синтезу мовлення для малоресурсних мов. Комбінація крос-мовного перенесення та LoRA шарів дозволяє ефективно адаптувати актуальні багатомовні моделі до української мови з контрольованою емоційною інтенсивністю. Подальші дослідження будуть зосереджені на розширенні набору емоцій, покращенні якості крос-мовного перенесення знань у моделях машинного навчання та удосконаленні підходу до моделювання емоційної інтенсивності.

Список використаних джерел

1. Ukrainian TTS using ESPnet framework // GitHub repository. 2021. URL: <https://github.com/robinhad/ukrainian-tts>.
2. Kim J., Kong J., Son J. Conditional variational autoencoder with adversarial learning for end-to-end text-to-speech // Proceedings of the 38th International Conference on Machine Learning (ICML). 2021. P. 5530-5540.
3. Casanova E., Weber J., Shulby C.D. et al. YourTTS: Towards Zero-Shot Multi-Speaker TTS and Zero-Shot Voice Conversion for Everyone // Proceedings of the 39th International Conference on Machine Learning (ICML). 2022. P. 2709-2720.
4. Casanova E. et al. XTTS: a Massively Multilingual Zero-Shot Text-to-Speech Model // arXiv preprint arXiv:2406.04904. 2024.
5. Hu E.J., Shen Y., Wallis P. et al. LoRA: Low-Rank Adaptation of Large Language Models // International Conference on Learning Representations (ICLR). 2022.

ВДОСКОНАЛЕННЯ ВІЗУАЛІЗАЦІЇ МОДЕЛЕЙ АРХІТЕКТУР ПРОГРАМ ДЛЯ ЗМЕНШЕННЯ КОГНІТИВНОГО НАВАНТАЖЕННЯ КОРИСТУВАЧА / IMPROVING VISUALIZATION OF SOFTWARE ARCHITECTURE MODELS TO REDUCE USER COGNITIVE LOAD

Франків О.О. / Frankiv O.O.

Національний університет «Києво-Могилянська академія» / National University of Kyiv-Mohyla Academy

04070, м. Київ, вул. Г. Сковороди 2, каф. інформатики, тел. (044) 425 60 64

E-mail: o.frankiv@ukma.edu.ua, тел. (044) 425 60 64

The main priority in developing automated visualization tools is to provide a clear and easily interpretable experience for the user. When visualizing complex and abstract entities such as software architecture, reducing the cognitive effort required to perceive information is essential, allowing the user to focus on the actual domain. In this paper, we examine practical

approaches to improving the UX and perceptual quality of the ADAR software-architecture visualization tool. Although graph-based models are intuitively suitable for representing connections, the quality of graph layout can be crucial for meaningful comprehension. At the same time, recent research shows that commonly used layout metrics may be misleading. From this perspective, user studies offer a more reliable basis for quality assessment. Furthermore, targeted improvements to the tool itself are expected to reduce cognitive load and enhance the overall usability.

На сьогодні проблема постійного аудиту та своєчасного виявлення вад програм під час розроблення залишається актуальною. Хоча для виявлення відносно простих проблем існує великий набір інструментів, складним завданням залишається розроблення нових інструментів, що дозволяють легко аналізувати архітектуру програми та виявити вади на цьому рівні. Додатковим викликом є природня потреба у мінімізації додаткових затрат ресурсів на використання відповідних інструментів.

Раніше запропонований програмний комплекс ADAR – Architecture Displayer in Augmented Reality – дозволяє автоматично будувати візуальну модель програм написаних мовою Swift у формі зваженого орієнтованого графу та відображати її у доповненій реальності (AR) за допомогою спеціального застосунку [1]. У графовій моделі типи даних (класи), їхні поля та методи відображаються у вигляді вузлів, а зв'язки між ними у вигляді направлених зважених ребер, що позначають напрям та інтенсивність.

Очевидним є те, що для достатньо великих програмних модулів графова модель буде налічувати велику кількість вузлів та ребер. Проблема швидкодії в даному контексті є зрозумілою і її можна вирішувати завдяки інтелектуальним оптимізаціям [2]. Однак значно складнішими є питання якості подання та забезпечення легкого сприйняття, що є критичними для будь-якої візуалізації, адже саме зменшення когнітивного навантаження при сприйнятті інформації є визначальним фактором.

Візуальна естетичність розміщення графа в просторі природньо визначає складність його сприйняття. На сьогодні існує велика кількість критеріїв естетичності та метрик, що дозволяють чисельно визначити якість подання. Більшість таких критеріїв, як-от мінімізація накладання вузлів чи перетинів ребер є інтуїтивно зрозумілими і очевидно ведуть до зменшення когнітивного навантаження користувача. Задоволення цих базових критеріїв не гарантує ідеального подання, тому існує також велика кількість метрик, покращення яких додатково роблять розміщення естетичнішим. Однак нещодавні дослідження показують, що часто використання багатьох метрик одночасно може бути суперечливим, хоча кожна окремо має сенс [3]. Окрім цього саме формальне задоволення набору критеріїв та покращення метрик не є достатнім, адже було проілюстровано, що подання графів можуть формально цілком відповідати критеріям і виглядати привабливо з точки зору визначення метрик, однак бути складними для сприйняття людиною. З огляду на це можна припустити, що оскільки збільшення кількості метрик та критеріїв не обов'язково гарантують створення простого для сприйняття подання, балансом між швидкодією та якістю візуалізації залишається застосування класичних алгоритмів, наприклад силового.

З іншого боку хоча результат виконання силового алгоритму для великих графів дозволяє отримати привабливе відображення в просторі, це не обов'язково означає простоту для розбору та інтерпретації даних. Тому ортогональним підходом до зменшення когнітивного навантаження є вдосконалення самого інструменту, що відображає певне подання графу.

Створення якісного користувацького досвіду, що буде нести мінімальне когнітивне навантаження, це вагома частина розроблення будь-якого інструменту. На сьогодні існує багато досліджень, що пропонують певні метрики для числового визначення інтуїтивності інтерфейсу та подачі інформації загалом [4]. Позитивні показники частини з них може забезпечити загальний підхід до візуалізації, для інших — створення гнучких налаштувань чи інструментів взаємодії.

Читабельність подання (layout readability) визначають за допомогою метрик співвідношення фону (white space ratio) та співвідношення перекриття міток (label overlap ratio). Відображення моделі архітектури в просторі за допомогою доповненої реальності природнім чином збільшує частку фону на зображенні, що може свідчити про хорошу композицію. З іншого боку надання користувачу можливості повертати та масштабувати модель в просторі, а також показувати і ховати описові мітки дозволяє не лише мінімізувати к-сть міток і, як наслідок, їхніх перекриттів, але й чітко контролювати кількість наданої інформації.

Кількість перекриттів (occlusion rate) визначає рівень просторового сприйняття. Класичний силовий алгоритм забезпечує мінімізацію, але не повне усунення перекриттів віртуальних об'єктів, особливо з урахуванням того, що де-факто користувач сприймає тривимірну модель як двовимірну проекцію. Однак можливість легко взаємодіяти з моделлю дозволяє користувачу змінювати цю проекцію залежно від своїх потреб і конкретної моделі, таким чином гнучко мінімізуючи кількість перекриттів у зоні інтересу.

Проблема використання кольорів є важливим фактором в оцінці складності сприйняття [4][5]. До прикладу семантична метрика “складність легенди” вираховується за кількістю унікальних кольорів, що використовуються на схемі. Для зменшення навантаження на користувача при пошуку вад проектування в інструменті візуалізації передбачено окремий біколірний режим. Таким чином всі елементи моделі є нейтральними, а ті, в яких система виявила вади, позначаються контрастним кольором, на відміну від звичайного режиму з повною інформацією, для відображення якої необхідно щонайменше десять кольорів.

Ще однією метрикою, що визначає легкість сприйняття є співвідношення контрастів (contrast ratio). Особливість її застосування полягає в тому, що співвідношення має бути в певному “достатньому” діапазоні залежно від контексту та тривалості використання інтерфейсу. Забезпечення правильного контрасту є комплексною проблемою, особливо в доповненій реальності, де освітлення середовища сильно впливає на контрастність доповненого зображення. Для забезпечення максимально комфортного сприйняття в програмі передбачено автоматичне вирівнювання контрасту віртуальних об'єктів з реальним середовищем.

Однак насамкінець варто зазначити, що будь-які метрики є спробою формально визначити результати емпіричних досліджень на реальних користувачах і допускають певні неточності зумовлені відмінностями між об'єктами досліджень. Тому більш реальну та якісну оцінку рівня когнітивного навантаження можна отримати саме з використанням відповідних замірів отриманих в результаті емпіричних досліджень для конкретної системи.

Список джерел:

1. Франків О. О., Глибовець М. М. Автоматизована візуалізація компонентів архітектури програми для мови Swift. *Кібернетика та системний аналіз*. 2024. Т. 60, вип. 6. С. 190–198. URL: <https://doi.org/10.1007/s10559-024-00737-9>.
2. Frankiv O. Machine learning in enhancing visualization of the spatial software architecture model. *Bulletin of Taras Shevchenko National University of Kyiv. Physical and Mathematical Sciences*. 2025. Т. 80, вип. 1. С. 164–173. URL: <https://doi.org/10.17721/1812-5409.2025/1.22>.
3. Wageningen S., Mchedlidze T., Telea A. C. Same Quality Metrics, Different Graph Drawings. *International Graph Drawing Conference 2025*. URL: <https://doi.org/10.48550/arXiv.2508.15557>.
4. A Survey on Measuring Cognitive Workload in Human-Computer Interaction / T. Kosch та ін. *ACM Computing Surveys*. 2023. Т. 55. URL: <https://doi.org/10.1145/3582272>.

5. Zhang R., Son O.-S. A Measurement Model for Visual Complexity in HCI: Focusing on Visual Elements in Mobile GUI Design. *Electronics*. 2025. Т. 14, вип. 5. С. 942. URL: <https://doi.org/10.3390/electronics14050942>.

ПРИСКОРЕНЕ МОДЕЛЮВАННЯ ЕПІДЕМІЙ У ВЕЛИКИХ МЕРЕЖАХ: ОПТИМІЗАЦІЯ МЕТОДУ ГІЛЛЕСПІ ТА ДВОШАРОВИЙ ПІДХІД / ACCELERATED EPIDEMIC SIMULATION IN LARGE-SCALE NETWORKS: OPTIMIZATION OF THE GILLESPIE ALGORITHM AND A TWO-LAYER APPROACH

Куриляк Ю.А.¹, Еммеріх М.Т.М.² / Kuryliak Y.A.¹, Emmerich M.T.M.²

¹Національний університет «Львівська політехніка» / Lviv Polytechnic National University

вулиця Степана Бандери, 12, Львів, Львівська область, 79000, Україна

²Університет Юваскуля / University of Jyväskylä

Agora, Mattilanniemi 2, 40100 Jyväskylä, Jyväskylä, 40014, Finland

E-mail: ¹yulian.a.kuryliak@lpnu.ua, ²michael.t.m.emmerich@jyu.fi

Abstract: This study addresses the challenge of accelerating epidemic simulations in large-scale complex networks through algorithmic and structural optimization. Traditional Gillespie-based stochastic simulations accurately reproduce epidemic dynamics but become computationally prohibitive for networks exceeding tens of thousands of nodes. To overcome this limitation, we build upon such efficiency techniques as local rate updates and ordered event-selection structures, which reduce the computational complexity of each simulation step from $O(n)$ to $O(\log(n))$. Building on these principles, we propose a two-layer (micro–macro) modeling framework: the micro layer simulates intra-community dynamics, while the macro layer captures inter-community infections using hazard-integral rates derived from mobility data and the epidemic states of metanodes. This hierarchical approach enables scalable and parallelizable simulations that preserve stochastic accuracy while substantially reducing computational cost, allowing realistic modeling of epidemic spread across millions of agents and multiple cities.

Epidemics have accompanied humanity throughout history. Each time a new wave of disease emerges, societies seek to understand how it spreads and what can be done to contain it. Historically, epidemic dynamics were modeled using differential equations based on the mean-field assumption — treating the population as a homogeneous system without accounting for its contact structure. Today, with increasing computational power and advances in simulation technology, it has become possible to reproduce infection dynamics in realistic, irregular contact networks that reflect individual-level interactions.

Contact networks are sparse, heterogeneous in degree, and exhibit strong local clustering and a pronounced community structure — groups of nodes densely connected within communities but weakly linked to others. Therefore, accurate modeling requires capturing explicit node-to-node connections and their temporal state transitions. Agent-based modeling achieves this by representing each node as an individual whose contacts define potential infection paths.

Such processes can be formulated as Continuous-Time Markov Chains (CTMC). To simulate them efficiently, the Gillespie algorithm provides an exact stochastic representation of system dynamics. Its key principles are: (I) only feasible transitions from the current state are considered; (II) only one event occurs at a time; and (III) the time to the next event follows an exponential distribution, after which the event is selected proportionally to its rate. This ensures accuracy and avoids state-space explosion, but computational complexity still limits the model size. As shown in [1], the classical Gillespie algorithm can simulate up to several tens of thousands of nodes, which remains insufficient for realistic population scales. For comparison, modern agent-based frameworks such as Covasim [2] implement stochastic transitions in discrete time with a daily step, which simplifies computations but sacrifices temporal precision.

The first optimization involves local rate updates. In the standard algorithm, every state change triggers recalculation of all possible transitions. In the improved implementation, only the node that changes state and its immediate neighbors are updated, dramatically reducing the number of operations while preserving accuracy.

The second optimization employs ordered data structures for rapid event selection. Whereas the classical approach scans all transition rates, the proposed sorted-list method achieves logarithmic-time selection [1]. An alternative solution is the binary-tree approach described in [3].

Despite these optimizations, simulating very large networks remains computationally demanding. To further enhance scalability, we introduce a two-layer model that hierarchically partitions the network into communities. At the micro layer, each community is simulated using the optimized Gillespie algorithm; at the macro layer, rarer interactions between communities are modeled using CTMC dynamics. This hierarchical approach, detailed in [4], drastically reduces the number of processed events while maintaining node-level granularity.

In the full-scale simulation, the underlying contact network is assumed to be static, and communities are connected through a limited number of bridge nodes that mediate inter-community infections. Once such a bridge node becomes infected, additional external infections cannot enter the same community until recovery occurs, effectively constraining cross-community transmission. In contrast, the two-layer abstraction treats each community as a metanode at the macro layer, where new infections are assigned to a random node within the community rather than fixed bridges. This eliminates topological bias, better reflects short-term human mobility, and makes the model behavior effectively more dynamic while preserving consistency with the underlying network structure.

Another advantage of the two-layer model is its parallel execution capability. Since most events occur within communities, each community can be simulated on a separate processor core, yielding near-linear speedup as the number of cores increases — particularly when inter-community connectivity is low. In this regime, synchronization between micro-layer simulations is infrequent because cross-community infections are rare, allowing communities to evolve almost independently. As inter-community connections become denser, synchronization must occur more often, which slightly reduces efficiency; however, the overall runtime still remains significantly lower than in a monolithic model. Moreover, epidemic activity peaks in different communities typically occur at different times, enabling dynamic reallocation of computational resources to the most active regions. Under limited hardware resources, a subset of communities outside the primary focus of analysis can be replaced with precomputed infection curves, preserving the accuracy of global dynamics while further reducing computational cost.

The two-layer simulator can also be interpreted as a simplified social-mobility model, where macro-layer nodes represent cities and edge weights correspond to intercity travel flows. With appropriate mobility data, the model enables analysis of how changes in movement patterns or local restrictions affect the rate and scale of infection spread. Australia offers a convenient case study due to its geographic isolation; its intercity mobility network can be reconstructed from mobile-operator-derived datasets [5].

In conclusion, the optimized Gillespie algorithm combined with the two-layer modeling framework achieves a practical balance between stochastic accuracy and computational scalability. It enables realistic simulations of epidemic dynamics in large, structured populations and can be readily integrated with demographic and mobility datasets. This framework also supports direct comparison between simulated and real-world epidemic data, paving the way for quantitative evaluation of intervention and containment policies.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Kuryliak, Y., Emmerich, M. T. M., Dosyn, D. Simulating epidemic peak dynamics on complex networks using efficient Gillespie algorithms. *Infection, Genetics and Evolution*, 2025, 105768. DOI: 10.1016/j.meegid.2025.105768

2. Kerr, C. C., Stuart, R. M., Mistry, D., et al. Covasim: An agent-based model of COVID-19 dynamics and interventions. *PLOS Computational Biology*, 2021, vol. 17, no. 7, e1009149. DOI: 10.1371/journal.pcbi.1009149.
3. Vo Hong Thanh, R. & Zunino, R. Tree-based search for Stochastic Simulation Algorithm. In: *Proceedings of the 27th Annual ACM Symposium on Applied Computing (SAC '12)*. Riva del Garda, Italy, 2012, pp. 144–151. DOI: 10.1145/2245276.2232001.
4. Kuryliak, Y. & Emmerich, M. T. M. Towards complexity reduction of large-scale epidemic simulation in two-scale networks. 2025, *CEUR Workshop Proceedings 4004*: pp. 327-336.
5. Du, Z., Yang, Y., Ertem, Z., et al. A review of human mobility: Linking data, models, and real-world applications. *Humanities and Social Sciences Communications*, 2025, vol. 12, no. 1, pp. 1–18. DOI: 10.1057/s41599-025-03160-7

ЗАСТОСУВАННЯ АЛГОРИТМІВ НАВЧАННЯ Q-МЕРЕЖ ДЛЯ ІТЕРАТИВНОЇ ЗАДАЧІ В'ЯЗНЯ

Ткач.Н.В.

Національний університет “Києво-Могилянська академія”

04655, м. Київ, вулиця Сковороди, 2, НаУКМА, Факультет інформатики. тел.(044) 426 60 64.

E-mail: n.tkach@ukma.edu.ua, факс (044) 426 60 64

This study is aimed at showcasing the performance of Deep Q-Networks (DQN) for the Iterated Prisoner’s Dilemma (IPD) with a compact episodic state embedding. The agent compresses the interaction context into a fixed-size vector and is trained against deterministic Axelrod strategies. Evaluation of normalized payoff, pairwise cooperation rate of strategies, and the learned behavior of the agent suggests the possibility of efficiently clustering existing strategies by latent learnable features. This may lead to advancements in both game theory and reinforcement learning. The limitations are outlined for future research, including recurrent-based and transformer-based policy-learning networks, stochastic opponents, and comparative analysis to the baseline performance.

Ітераційна дилема в’язня (IPD) є класичною моделлю співпраці^[1]. У даній роботі представлено результати агента, навченого на основі алгоритму навчання з підкріпленням глибинних Q-мереж (“deep Q-network”)^[3], здатного формувати контекстно-залежні рішення за рахунок стислого подання стану. Модель тренувалася проти 72 детерміністичних стратегій у 200-ходових епізодах, реалізованих згідно з турнірами Аксельрода [1] в однойменній бібліотеці.

DQN-агент приймає стислий стан, - 16-вимірний вектор поточної історії гри, що кодує ключові патерни взаємодії. Навчання здійснюється проти детерміністичної популяції у “сліпому” режимі - модель немає інформації про стратегію, проти якої навчається. Функція втрат $L(\theta)$ базується на

TD-помилці функції оцінювання $V_{\psi}(s)$. Q-функцію^[2] ми оцінюємо за допомогою багат шарового перцептрон, який приймає на вхід інформацію про стан s що містить контекст стратегії та останню дію a . Параметрами φ та θ інтерпретуємо як відповідні параметри нейронних мереж які надають оцінку поточного стану, та оцінку стратегії, яку агент має вибрати для мінімізації похибки часової різниці(TD).

$$\min_{\psi} \left(r + \gamma V_{\psi}(s') - V_{\psi}(s) \right)^2 \quad (1)$$

$$y^{DQN} = r + \gamma \max_{\theta} Q_{\theta}(s', a) \quad (2)$$

$$L(\theta) = E_{(s, a, r, s')_D} \left[\left(y^{DQN} - Q_{\theta}(s, a) \right)^2 \right] \quad (3)$$

Внаслідок тренування можна виокремити 4 кластери стратегій (рис.1) які згруповані за очікуваною винагородою та агента та їхнім рівнем кооперації. видно близько чотирьох кластерів стратегій. На рис.2 виведені результати різниці між очікуваним виграшом стратегії та результатами навченого агента.

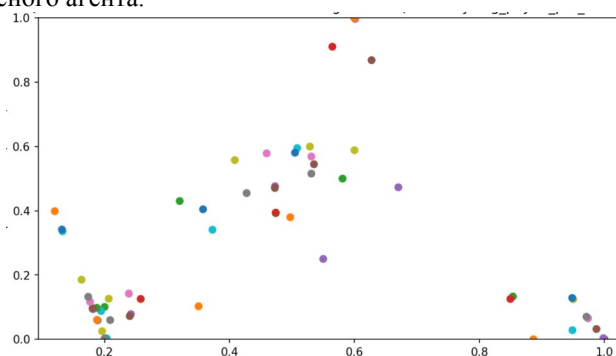


Рис.1 Розподіл стратегій у просторі їхньої здатності до кооперації з агентом та зваженого результату агента.

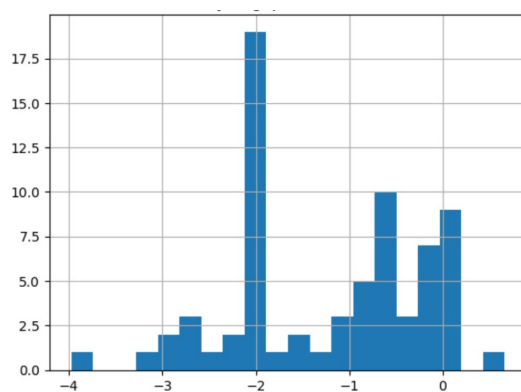


Рис.2 Різниця середнього очікуваного виграшу стратегій відносно навченого агента.

Стисле подання стану дозволяє розрізнити типи опонентів і адаптивно реагувати на їхні патерни поведінки і ефективно відтворювати їх під час фази тестування. Основними обмеженнями цього дослідження є детермінованість популяції, відсутність довготривалої пам'яті та нестационарність динаміки навчання. У подальших дослідженнях буде проведено порівняння з рекурентними та трансформер-моделями, введено стохастичних опонентів^[4] та популяційний тренінг. У підсумку, можна стверджувати що DQN із компактним поданням показує контекстно-залежну поведінку та виявляє кластерну структуру опонентів, формуючи основу для подальших досліджень кооперації.

ДЖЕРЕЛА

1. Axelrod, Robert, and William D. Hamilton. "The evolution of cooperation." *science* 211.4489 (1981): 1390-1396.
2. Mnih, Volodymyr, et al. "Human-level control through deep reinforcement learning." *nature* 518.7540 (2015): 529-533.
3. Roderick, Melrose, James MacGlashan, and Stefanie Tellex. "Implementing the deep q-network." *arXiv preprint arXiv:1711.07478* (2017).
4. Bertrand, Quentin, et al. "Q-learners Can Provably Collude in the Iterated Prisoner's Dilemma." *arXiv preprint arXiv:2312.08484* (2023).

ЕФЕКТИВНЕ НАВЧАННЯ СТРАТЕГІЙ КЕРУВАННЯ ДЛЯ РОБОТИЗОВАНИХ МАНІПУЛЯЦІЙ ЗА ДОПОМОГОЮ ДИСТИЛЯЦІЇ ЗНАНЬ/EFFICIENT POLICY LEARNING VIA KNOWLEDGE DISTILLATION FOR ROBOTIC MANIPULATION

Севергін О.В., Кузьменко Д.О., Швай Н.О. / Severhin O., Kuzmenko D., Shvai N.
Національний університет “Києво-Могилянська Академія” / National University of Kyiv-Mohyla
Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра математики
E-mail: oleksandr.severhin@ukma.edu.ua, Kuzmenko@ukma.edu.ua, n.shvai@ukma.edu.ua

The work focuses on the computational intractability of large-scale Reinforcement Learning (RL) models for robotic manipulation. While world-like models like TD-MPC2 demonstrate high performance in various manipulative tasks, their immense parameter count (e.g., 317M) hinders training and deployment on resource-constrained hardware. This research investigates Knowledge Distillation (KD) with a loss function specifically described in [1] and [2] as a primary method for model compression. This involves training a lightweight "student" model to mimic the behavior of a large, pre-trained "teacher" model. Unlike in supervised learning, distilling knowledge in RL is uniquely complex; the objective is to transfer a dynamic, reward-driven policy, not a simple input-output function.

Previous work in this area, specifically the TD-MPC-Opt study [3], established the feasibility of this approach using a static distillation coefficient. While successful in transferring capabilities to a 1M parameter student, a static coefficient presents a conceptual trade-off. A high, constant coefficient may overly bias the student towards imitation, stifling its own exploration and learning from direct environmental rewards. Conversely, a low coefficient may fail to provide sufficient guidance. This work extends the original study by hypothesizing that a dynamic distillation coefficient, which adapts over the training process, can lead to faster convergence and a more robust final policy.

Our methodology is centered on a composite loss function that governs the student model's training. This loss function has two primary components. The first is the student's original learning loss, which includes standard RL objectives such as reward loss, value loss, and consistency loss (a key feature of TD-MPC ensuring stable predictions). This component allows the student model to learn directly from its own environmental interactions. The second component is distillation loss. For this, we specifically target the teacher's learned value function. We chose this target because the value function encapsulates the teacher's long-term, high-level understanding of task success, providing a much denser and more stable training signal than distilling a volatile, high-variance policy directly. The loss is calculated as the mean squared error between the reward predictions of the teacher and the student for identical states and actions.

The entire process is balanced by the distillation coefficient, a critical hyperparameter that determines the weighting between self-learning and teacher imitation at any given point in training. The higher the coefficient the bigger effect teacher imitation has on the student model.

The central investigation of this paper is the behavior of this distillation coefficient. We moved beyond the static baseline to experiment with several adaptive schedules around 50,000 training steps. These included: Linear Decay (a high initial coefficient gradually decreasing to the baseline of 0.4, prioritizing guidance early); Cosine Decay (a smoother, non-linear version of decay); Increase (testing the inverse hypothesis of starting with self-learning and adding guidance later); and a Four Phase stepped schedule to test for sensitivity to abrupt changes.

We conducted our experiments on the MT30 dataset, distilling knowledge into a 1M parameter student model. The results demonstrate that dynamic coefficient strategies significantly outperform the static baseline.

Analysis of Results: The dynamic schedules showed a clear performance hierarchy. Both Linear Decay and Cosine Decay achieved the highest normalized score (8.64), surpassing the static Constant (0.4) coefficient's score of 8.39. This finding strongly supports our hypothesis: "front-loading" the teacher's guidance is the most effective strategy. At the same time same dynamic distillation coefficient showed worse results when they were starting with 1 and ending with 0 (lower than the baseline of 0.4) that means that not only "front-loading", but also a quite a strong "support" closer to end of training and KD is important.

The success of the decay schedules suggests that the student model benefits most from heavy imitation during the initial, random phases of its policy. As the student's own policy improves, the relative

importance of this guidance is reduced, allowing the student to fine-tune its behavior based on direct environmental feedback.

Conversely, strategies that delayed or provided erratic guidance underperformed. The "Increase" schedule, which delayed guidance, yielded a poor score of 8.09. The "Four Phase" schedule (8.32) and another "Decrease" variant (8.07) also failed to match the baseline. The failure of the "Increase" schedule is particularly informative, as it implies that withholding guidance initially is detrimental. The student fails to form a competent base policy on its own and cannot effectively leverage the teacher's complex knowledge when it is finally introduced. The results show that not all dynamic strategies are beneficial, but a clear pattern emerges: the most effective methods apply a high coefficient at the beginning of training and gradually decrease it.

Subsequent work will proceed in two main directions. First, we will validate these findings on more complex and diverse domains, specifically the ManiSkill 3 dataset. This benchmark includes tasks with complex object interactions and deformable objects, which will rigorously test the generalization of our dynamic distillation method in both single-task and multi-task regimes. We aim to verify if these decay schedules remain optimal or if greater task complexity necessitates a different approach.

Second, we will investigate a highly promising architectural direction: distilling knowledge into Mixture of Experts (MoE) models using a routing architecture [4]. Rather than compressing knowledge into a single dense student, we can distill it into a set of smaller, specialized "experts." This architecture is a natural fit for multi-task robotics, where one expert might learn "picking" motions and another "placing." We will explore using the teacher's knowledge to train not only the experts but also the routing mechanism that selects the appropriate expert for a given task. This could achieve higher performance and efficiency than a monolithic student.

Список джерел

1. Sheng R., Tao H., Wuyue Y. Tailored knowledge distillation with automated loss function learning. PLOS One. 2025. URL: <https://doi.org/10.1371/journal.pone.0325599>.
2. Romero A., Ballas N., Kahou S. E., Chassang A., Gatta C., Bengio Y. FitNets: Hints for Thin Deep Nets. arXiv preprint arXiv:1412.6550. 2014. URL: <https://arxiv.org/abs/1412.6550>.
3. Kuzmenko D., Shvai N. TD-MPC-Opt: Distilling Model-Based Multi-Task Reinforcement Learning Agents. arXiv preprint arXiv:2507.01823. 2025. URL: <https://arxiv.org/abs/2507.01823>.
4. Kuzmenko D., Shvai N. MoIRA: Modular Instruction Routing Architecture for Multi-Task Robotics. arXiv preprint arXiv:2507.01843. 2025. URL: <https://arxiv.org/abs/2507.01843>.

КООРДИНАЦІЯ МУЛЬТИАГЕНТНИХ LLM-СИСТЕМ / COORDINATION OF MULTI-AGENT LLM-BASED SYSTEMS

Діхтяр І.Ю., Нагірна А.М. / Dikhtiar I.Yu., Nahirna A.M.

Національний університет "Києво-Могилянська Академія" / National University of Kyiv-Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики
E-mail: i.dikhtiar@ukma.edu.ua, a.nahirna@ukma.edu.ua

Abstract. This research focuses on analyzing the current state and challenges in the coordination of multi-agent systems based on Large Language Models (LLMs). We review the evolution from single autonomous agents (like ReAct) to complex multi-agent architectures. The analysis of current approaches reveals a key limitation: a reliance on deterministic, predefined coordination rules. This paper identifies the lack of adaptive, learned coordination policies as a research gap. We conclude by highlighting Reinforcement Learning (RL) as a promising, though challenging, direction for developing dynamic meta-agents capable of overcoming these limitations.

Вступ. Розвиток великих мовних моделей (LLM) трансформував сферу штучного інтелекту, дозволивши створити автономних агентів, здатних до складних міркувань та планування. Такі підходи, як ReAct [1] та Reflexion [2], продемонстрували здатність одиночних агентів вирішувати багатоетапні завдання. Логічним наступним кроком стала побудова мультиагентних систем на основі LLM, де для вирішення комплексної проблеми залучається декілька спеціалізованих агентів. Однак такий перехід породжує нову фундаментальну проблему – необхідність ефективної координації їхньої спільної діяльності.

Аналіз стану проблеми. Ключовим проривом став ReAct [1], який запропонував інтегрувати етапи логічних міркувань (Reason) та виконання дій (Act). Агент, що діє за цим принципом, циклічно генерує "думку" (план або міркування), потім "дію" (наприклад, пошук в Інтернеті), отримує результат (Observation), і знову повертається до міркувань, коригуючи свій план. Іншим важливим підходом є Reflexion [2], який додав агенту механізм вербального самонавчання та довготривалу динамічну пам'ять. Після невдалої спроби агент аналізує траєкторію власних дій, генерує текстову "рефлексію" про причини помилки і зберігає її у пам'яті, щоб уникнути помилкових кроків у майбутньому.

Однак, при переході до мультиагентних сценаріїв, популярні практичні реалізації, такі як AutoGPT та CrewAI, хоч і будуються на цих ідеях, здебільшого покладаються на статичні, заздалегідь визначені ролі та фіксовані сценарії. Це знижує гнучкість системи та її здатність адаптуватися до непередбачуваних умов.

Більш просунуті архітектурні рішення, що з'являються, зокрема у вітчизняному науковому просторі, намагаються вирішити цю проблему шляхом введення централізованого управляючого модуля. Наприклад, у роботі [3] запропоновано архітектуру "LLMAgentNet" – "Модульну Мережу Взаємодіючих Агентів". Її мета – подолати обмеження масштабованості та залежності від конкретних інструментів. У цій архітектурі вводиться "спеціальний менеджер агента (Orchestrator)", який відповідає за "декомпозицію запитів", "вибір виконавців" та "узгодження потоків даних". Важливо, що взаємодія агентів у цій системі відбувається через RESTful API як універсальний комунікаційний інтерфейс. В іншому дослідженні [4] розглядається мультиагентна система для гетерогенного мультимарного середовища. Вона включає технічного, економічного та безпекового агентів, а їхню роботу координує "спеціалізований агент-координатор". Цей координатор використовує LLM для "обґрунтування прийнятих рішень" та "аналізу стану" середовища.

Попри значний архітектурний прогрес, глибокий аналіз усіх згаданих підходів (від CrewAI до "LLMAgentNet" [3]) виявляє спільне фундаментальне обмеження: механізми координації залишаються детермінованими або аналітичними. Опора LLMAgentNet на RESTful API робить його координацію процедурною, але не адаптивною – вона діє за чітко прописаними правилами обміну даними. Так само "агент-координатор" [4] використовує LLM пасивно – для обґрунтування рішення, яке, ймовірно, прийняте на основі інших метрик, а не для динамічного формування рішення. Навіть Reflexion [2], хоч і є адаптивним, використовує евристичний, а не оптимальний механізм самокорекції. Таким чином, ключова науково-технічна проблема залишається невирішеною: відсутність механізму, який би дозволив управляючому "метаагенту" адаптивно навчатися оптимальній політиці (policy) координації на основі досвіду, а не слідувати заздалегідь прописаним правилам чи евристикам. Сучасні системи не відповідають на питання: як оптимально обрати наступного агента? Коли варто створити нового агента, а коли видалити? Як сформувати команду, що мінімізує витрати та дублювання дій?

Висновки та напрями майбутніх досліджень. Огляд поточного стану розробки мультиагентних систем на основі LLM свідчить про чіткий перехід від одиночних агентів до координуваних систем. Ідентифікована прогалина – відсутність адаптивних політик управління – є головним стримуючим фактором для створення по-справжньому гнучких та автономних когнітивних архітектур.

Перспективним напрямом для подолання цього обмеження є застосування методів навчання з підкріпленням (Reinforcement Learning, RL) для навчання метаагента. Формалізація цього процесу як Марковського процесу прийняття рішень дозволить метаагенту формувати динамічні стратегії розподілу ролей та завдань. Водночас, цей напрям має власні виклики. Зокрема, стандартні підходи можуть призводити до "нестабільності поведінки" та "обчислювальної неефективності" [5]. Тому майбутні дослідження мають фокусуватися на більш досконалих методах RL, таких як "навчання на основі багатовимірної винагороди та диференційованої критики" [5], що дозволить створювати більш надійних та логічно послідовних автономних агентів.

Список використаних джерел

1. Zhao J., Chen Y., & Yu W. (2023). LLM-Based Multi-Agent Collaboration: Reasoning, Coordination, and Limitations. Proceedings of NeurIPS 2023 Workshop on Agentic AI.
2. Wang T., Liu S., & Zhang Q. (2024). Hierarchical Meta-Agents for Coordinated Reinforcement Learning. Journal of Artificial Intelligence Research, 78(4), 1123–1145.
3. Новаковський І., Яловега О. (2025). Здібності LLM-агентів та їх застосування у креативних і планувальних сценаріях. Препринт.
4. Урсацьєв О., Вольков О. (2025). Особливості застосування Deep RL у керуванні групами автономних систем (MARL). Матеріали конференції.
5. Wu Z., Li K., & Zhao Y. (2022). Challenges in Multi-Agent Reinforcement Learning: A Review of Coordination and Scalability Issues. IEEE Access, 10, 12867–12882.

РОЛЬ ФУЗИНГУ ТА СУЧАСНИХ ПІДХОДІВ У ВИЯВЛЕННІ ТА ЕКСПЛУАТАЦІЇ ВРАЗЛИВОСТЕЙ «НУЛЬОВОГО ДНЯ» В УМОВАХ ОБМЕЖЕНИХ РЕСУРСІВ / THE ROLE OF FUSING AND MODERN APPROACHES IN DETECTION AND EXPLOITATION OF ZERO-DAY VULNERABILITIES IN A RESOURCE-CONSTRAINED ENVIRONMENT

Ісмагілов А.І. / Ismahilov A.I.

Інститут програмних систем Національної академії наук України / Institute of Software
Systems National Academy of Sciences of Ukraine

03187, Київ, просп. Академіка Глушкова, 40 корп.5, тел.: (044) 526-55-07, e-mail:
ismagilov@ukr.net

During the current year, a significant number of incidents have been directly associated with the exploitation of zero-day vulnerabilities. The study identifies key trends in their detection and analyzes the role of fuzzing as the primary technique for discovering zero-day vulnerabilities in proprietary software, particularly in environments where access to source code is restricted or unavailable. Special emphasis is placed on the integration of decompilation, symbolic execution, and coverage feedback as an effective methodology for exploring new execution paths in binary programs.

В продовж року зафіксовано низку масштабних атак, заснованих на використанні невідомих раніше вразливостей «нульового дня» в популярних програмних продуктах. Вразливості даного класу традиційно становлять найбільшу загрозу інформаційній безпеці, оскільки їх експлуатація можлива до моменту випуску оновлень або публічного їх розкриття.

У липні 2025 року були виявлені декілька вразливостей «нульового дня» в «on-premises» версіях Microsoft SharePoint. Зловмисники отримали несанкціонований доступ до серверів, викрадали облікові дані та виконували віддалений код. Масові атаки охопили як комерційні структури, так і державні організації у різних країнах світу. В свою чергу Microsoft здійснила виявлення, офіційне підтвердження та випуск оновлень у період з 18 по 22 липня 2025 року [1].

Іншим прикладом стала вразливість у компоненті ImageIO операційних систем Apple, де обробка спеціально сформованого зображення могла спричинити переповнення пам'яті та подальшого виконання довільного коду. В окремих випадках експлуатація здійснювалася за принципом «zero-click», тобто без взаємодії користувача. Компанія Apple випустила термінові оновлення у серпні 2025 року [2].

Також у вересні 2025 року було виявлено вразливість CVE-2025-10585 у двигуні Chrome/Chromium. Некоректна інтерпретація типу об'єкта під час обробки спеціально сформованого JS/Wasm-коду призводила до пошкодження купи (heap corruption) і потенційного виконання довільного коду. Вразливість було виявлено командою Google Threat Analysis Group 16 вересня 2025 року, а стабільне оновлення випущено наприкінці вересня 2025 року [3].

Аналіз практик 2025 року свідчить, що більшість вразливостей «нульового дня» у пропрієтарному програмному забезпеченні, зокрема Adobe, Microsoft, Apple, Google, Oracle було виявлено за допомогою зовнішніх фузерів. Наразі фузинг залишається основним методом виявлення помилок типу «memory corruption» і широко використовується у програмах Bug Bounty.

Водночас великі вендори активно впроваджують фузинг у власні процеси розробки, прикладами є Google OSS-Fuzz, Microsoft SAGE, Apple fuzz pipelines, водночас саме незалежні

дослідники залишаються головним джерелом виявлення вразливостей «нульового дня» у реальних умовах експлуатації.

Цьогоріч спостерігається тенденція до комбінованого використання декількох методів аналізу програм без доступу до вихідного коду, а саме:

1. Декомпіляція - застосовується для перетворення двійкових даних у читабельну структуру, що дозволяє ідентифікувати потенційні точки входу, парсери, мережеві обробники, API-виклики тощо;
2. Символьне виконання - використовується для побудови логічних шляхів до складних умов виконання та генерації специфічних тестових випадків;
3. Зворотний зв'язок щодо охоплення - реалізується за допомогою систем динамічної інструментації, таких як QEMU, Intel PIN або DynamoRIO, що дозволяє вимірювати ступінь новизни тестів за охопленням коду [4].

В свою чергу процес виявлення вразливостей «нульового дня» в умовах обмежених ресурсів за допомогою фузингу включає такі етапи:

1. Ідентифікація точок входу через декомпіляцію.
2. Побудова шляхів виконання за допомогою символьного виконання.
3. Генерація і мутація тестових вхідних даних (корпус).
4. Використання coverage feedback для визначення «нових» шляхів виконання.
5. Кластеризація та мінімізація краш-кейсів, проведення тріажу та аналітичної перевірки [5].

Схематичне відображення процесу наведено на Рисунок 1.

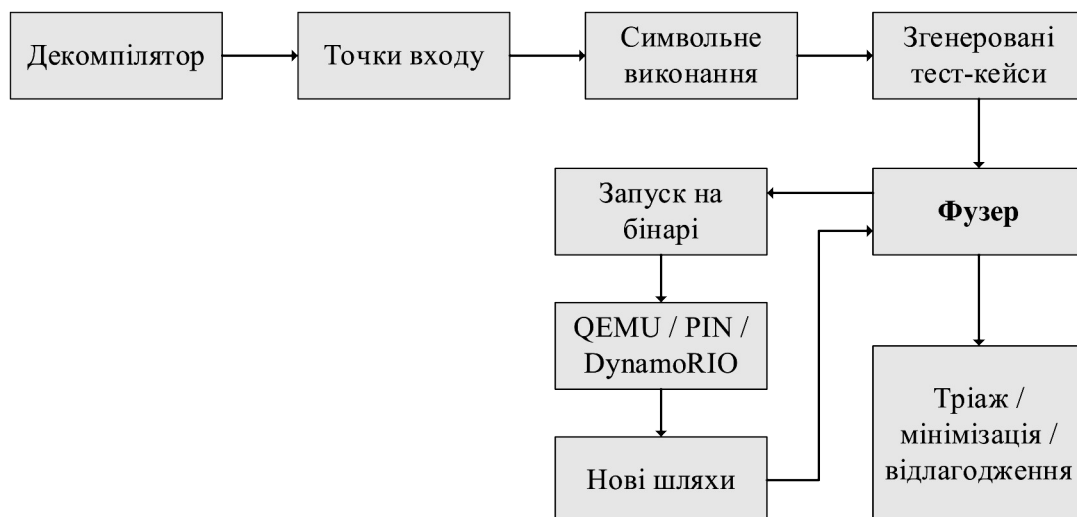


Рисунок 1. Процес виявлення вразливостей «нульового дня» в умовах обмежених ресурсів за допомогою фузингу

Підсумовуючи слід зазначити, що наразі фузинг є одним із найбільш ефективних інструментів виявлення вразливостей «нульового дня», особливо в умовах обмеженого доступу до вихідного коду та ресурсів. Максимальну ефективність забезпечує поєднання таких компонентів, як fuzz-harness, санітайзери (ASan/UBSan), coverage feedback, дисциплінований тріаж і відповідальне розкриття результатів.

Подальші дослідження мають бути спрямовані на оптимізацію фузингу в умовах обмежених обчислювальних ресурсів та використання методів машинного навчання та кластерного аналізу для автоматизації процесу класифікації знайдених вразливостей.

Список джерел

1. Microsoft SharePoint on-premises zero-day exploitation advisory. Microsoft Security Response Center, July 2025.
2. Security updates for ImageIO vulnerabilities. Apple Security Bulletin, August 2025.
3. CVE-2025-10585: Chrome heap corruption vulnerability. Google Threat Analysis Group Report, September 2025.

4. Shihao Jiangb , Yu Zhanga , Junqiang Lia , Hongfang Yua , Long Luo a , Gang Suna. A Survey of Network Protocol Fuzzing: Model, Techniques and Directions. URL: <https://arxiv.org/pdf/2402.17394> (Accessed: 11 November 2025).
5. Andrea Arcuri, Man Zhang, Susruthan Seran, Juan Pablo Galeotti, Amid Golmohammadi, Onur Duman, Agustina Aldasoro, Hernan Ghianni. Tool report: EvoMaster - black and white box search-based fuzzing for REST, GraphQL and RPC APIs. Automated Software Engineering (Autom. Softw. Eng.), Vol. 32, Article 4. 29 November 2024.

СЕКЦІЯ 2. ПРОГРАМНА ІНЖЕНЕРІЯ

РОЗПІЗНАВАННЯ ЕМОЦІЙНИХ СТАНІВ ПІД ЧАС ПЕРЕГЛЯДУ ВІДЕО НА ОСНОВІ ДАНИХ GSR, ПУЛЬСУ ТА ВІЗУАЛЬНОГО ВІДСТЕЖЕННЯ

Глибовець А.М., Хмель С.М. / Hlybovets A., Khmel S.

Національний університет "Києво-Могилянська академія" / National University of Kyiv-Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики

E-mail: a.glybovets@ukma.edu.ua, svitlana.khmel@ukma.edu.ua

This research focused on developing a prototype system for recognizing users' emotional states during video viewing through the integration of biometric and behavioural data. The system combines galvanic skin response (GSR), heart rate monitoring, and gaze tracking using the webgazer.js library to assess emotional arousal and engagement in real time. Biometric signals are collected via Arduino from MAX30102 and GSR Grove sensors and transmitted through the DSD TECH HM-10 Bluetooth Low Energy (BLE) module. A custom Chrome extension detects active video playback, tracks gaze coordinates, and synchronizes visual and physiological indicators into a unified JSON data structure. Experimental results confirm the system's feasibility for analysing user emotions and suggest potential applications in education, neuromarketing, and human-computer interaction.

Keywords: biometric reactions, emotional state, Arduino, MAX30102, GSR Grove, BLE, webgazer.js, human-computer interaction, BCI.

У сучасну епоху активної цифровізації зростає інтерес до систем, здатних розпізнавати емоційні стани людини у реальному часі. Такі системи мають важливе значення для розвитку людино-комп'ютерної взаємодії, адаптивних навчальних середовищ, маркетингових досліджень і нейроінтерфейсів. Особливо перспективним є підхід, що поєднує біометричні сенсори та аналітику поведінки користувача під час споживання мультимедійного контенту.

Метою даної роботи є створення прототипу системи розпізнавання емоційних станів користувача під час перегляду відео, який об'єднує дані гальванічної реакції шкіри (GSR), частоти серцевих скорочень (ЧСС) та напрямку погляду. Основою реалізації є мікроконтролерна платформа Arduino Uno, що забезпечує збір фізіологічних показників у режимі реального часу. Для вимірювання пульсу та рівня насиченості крові киснем використано сенсор MAX30102, а для фіксації електропровідності шкіри — модуль GSR Grove. Передача даних здійснюється через модуль DSD TECH HM-10 за технологією Bluetooth Low Energy (BLE), що забезпечує бездротову взаємодію між апаратною та програмною частинами системи.

Програмна складова реалізована у вигляді розширення для браузера Google Chrome, яке взаємодіє з апаратною частиною, приймає біометричні сигнали та синхронізує їх із поведінковими показниками користувача. Розширення автоматично виявляє активне відео на сторінці YouTube, відстежує його стан (позицію, тривалість, паузу, повноекранний режим), а також за допомогою бібліотеки webgazer.js визначає координати погляду користувача на екрані. Отримані біометричні та візуальні дані об'єднуються в уніфікований формат і зберігаються у вигляді структурованого файлу JSON. Такий підхід дозволяє формувати часові ряди емоційних та фізіологічних змін під час перегляду контенту.

Програмна складова складається з таких компонентів:

9. **Gaze Tracking:** відображає інтерфейс калібрування, визначає координати погляду користувача; виводить візуальний індикатор напрямку погляду; слухач `webgazer.setGazeListener` передає координати у `trackingData.updateGazeData`.
10. **Data Receiver:** підключається до BLE-модуля Arduino; приймає та розпізнає вхідні дані; синхронізує фізіологічні показники з координатами погляду.
11. **UI/UX Module:** кнопки запуску/зупинки відстеження, повідомлення про стан з'єднання.
12. **Messaging System:** забезпечує обмін даними між фоновим процесом, `popup`-вікном розширення та вкладкою браузера;
13. **Інтеграція та потік даних;** Arduino Uno зчитує дані з MAX30102 та GSR-сенсора, дані передаються через HM-10 BLE до комп'ютера користувача; Chrome Extension приймає ці дані та синхронізує з координатами погляду, які надходять від `webgazer.js`; інтегровані дані доступні для зберігання та подальшого аналізу.

Розроблена система побудована за модульним принципом. Кожен компонент — сенсори, модуль зв'язку, програмний інтерфейс та система обробки даних — функціонує автономно, що спрощує оновлення та розширення функціональності. Архітектура системи забезпечує масштабованість: можливе додавання нових типів сенсорів, інтеграція алгоритмів машинного навчання для автоматичної класифікації емоцій чи підключення альтернативних платформ, наприклад Raspberry Pi.

Під час експериментів було підтверджено працездатність прототипу. Система успішно зчитує та передає показники GSR, пульсу і SpO₂, синхронізує їх із координатами погляду та станом відео. У результаті формується повна картина фізіологічних реакцій користувача у процесі сприйняття медіаконтенту. Отримані дані можуть бути використані для побудови моделей оцінки емоційної залученості, візуалізації біологічного зворотного зв'язку або розроблення персоналізованих систем навчання.

Разом з тим, під час тестування виявлено низку обмежень. Зокрема, передача даних через BLE супроводжується затримками, що впливають на синхронізацію між сенсорами. Сенсор MAX30102 не є медичним приладом і демонструє похибки вимірювання, особливо за змінного освітлення чи неправильного положення пальця. Крім того, бібліотека `webgazer.js` має обмежену точність відстеження погляду, що ускладнює використання системи у неконтрольованих умовах.

Попри зазначені недоліки, розроблений прототип демонструє перспективність поєднання недорогих біометричних сенсорів, бездротової передачі даних та браузерних технологій для створення інтерактивних систем аналізу емоційних станів. Подальші напрями розвитку включають удосконалення механізмів калібрування сенсорів, підвищення стабільності зв'язку, застосування методів фільтрації шумів та використання алгоритмів машинного навчання для автоматизованої класифікації емоцій користувача.

Таким чином, представлена система є гнучкою платформою для інтеграції біологічних і поведінкових даних у процесі людино-комп'ютерної взаємодії. Її модульна архітектура, відкритість і простота реалізації роблять можливим широке застосування у сфері освітніх технологій, психологічних досліджень, медіа-аналітики та розвитку нейрокомп'ютерних інтерфейсів.

Список джерел

- [1] Chrome Developers. *Capabilities: Bluetooth* [Електронний ресурс] // Chrome Developers. – 2025. – Режим доступу: <https://developer.chrome.com/docs/capabilities/bluetooth> (дата звернення: 27.10.2025).
- [2] Instructables. *Guide to Using MAX30102 Heart Rate and Oxygen Sensor* [Електронний ресурс]. – 2025. – Режим доступу: <https://www.instructables.com/Guide-to-Using-MAX30102-Heart-Rate-and-Oxygen-Sens/> (дата звернення: 27.10.2025).
- [3] WebGazer.js. *WebGazer Eye Tracking Library* [Електронний ресурс]. – 2025. – Режим доступу: <https://webgazer.cs.brown.edu/> (дата звернення: 27.10.2025).

[4] Noldus. *What Is Galvanic Skin Response* [Електронний ресурс]. – 2025. – Режим доступу: <https://noldus.com/blog/what-is-galvanic-skin-response> (дата звернення: 27.10.2025).

[5] Google Developers. *Interact with BLE Devices on the Web* [Електронний ресурс]. – 2025. – Режим доступу: <https://developers.google.com/web/updates/2015/07/interact-with-ble-devices-on-the-web> (дата звернення: 28.10.2025).

ПРОБЛЕМИ ФРАГМЕНТАЦІЇ КОНТЕКСТУ В RAG СИСТЕМАХ / PROBLEMS OF CONTEXT FRAGMENTATION IN RAG SYSTEMS

Андрощук М.В. / Androshchuk M.

Національний університет “Києво-Могилянська Академія” / National University of Kyiv-Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики
E-mail: maxym.androshchuk@ukma.edu.ua

This research focuses on context fragmentation in Retrieval-Augmented Generation (RAG) systems, a critical issue where conventional document chunking severs logical connections, leading to poor response quality. It proposes a proactive strategy to manage context by overcoming the conflict between retrieval precision and generation completeness. Key solutions include structured indexing via Semantic Chunking and Parent-Child hierarchies, dynamic methods like Sentence-Window and Auto-Merging retrieval, and post-retrieval optimization using rerankers. A comparative analysis shows that an optimized RAG architecture significantly outperforms Large-Context LLMs (LC-LLMs) by avoiding pitfalls such as the "lost in the middle" problem. The paper concludes that strategic architectural investment in solving fragmentation is a more robust and scalable approach for enterprise applications than simply relying on larger LLM context windows.

Системи Retrieval-Augmented Generation (RAG) є поширеною архітектурною парадигмою, що інтегрує великі мовні моделі (LLM) із зовнішніми базами знань з метою підвищення фактологічної точності та мінімізації генерації необґрунтованої інформації ("галюцинацій")[2]. Функціонування цих систем обмежується основним архітектурним викликом — контекстною фрагментацією. Ця проблема виникає внаслідок процесу розбиття (chunking) вихідних документів на менші сегменти для їх ефективного індексування та векторного пошуку. Таке сегментування може призводити до розриву логічних, семантичних та структурних зв'язків, наявних у початковому документі. У результаті LLM отримує набір розрізнених фрагментів, що знижує когерентність генерованої відповіді, призводить до пропуску ключової інформації та включення нерелевантних даних ("шуму"). Особливо значущою є втрата каузальної цілісності, коли причинно-наслідкові зв'язки виявляються розподіленими між різними чанками, що ускладнює для моделі формування логічно обґрунтованих висновків [3].

Для вирішення цієї проблеми розробляються RAG-архітектури, що переходять від стандартних методів розбиття на частини фіксованого розміру до більш комплексних, контекстно-орієнтованих стратегій[5]. Одним із основних рішень є архітектура Parent Document Retrieval (PDR), яка розв'язує компроміс між вимогами до розміру чанків для пошуку та для генерації. В рамках PDR індексуються малі "дочірні" чанки для забезпечення високої точності семантичного пошуку, тоді як для генерації відповіді LLM надається асоційований з ними більший "батьківський" документ, що дозволяє відновити ширший контекст. Цей підхід часто доповнюється ієрархічним розбиттям, що враховує структурні елементи документа (розділи, параграфи), та гібридними стратегіями вибірки, які поєднують семантичний (векторний) та лексичний (на основі ключових слів) пошук для підвищення надійності вибірки.

Важливою складовою зменшення впливу контекстної фрагментації є постретривальна обробка, зокрема механізми переранжування (Re-ranking) [1]. На початковому етапі Retrieval система прагне забезпечити максимальну повноту (high recall) — тобто знайти якомога більше потенційно релевантних чанків. Проте це часто призводить до великої кількості “шумових” або слабо дотичних фрагментів. Щоб забезпечити якість подальшої генерації, може використовуватися переранжування: спеціальний етап, на якому відібрані фрагменти ранжуються за ступенем релевантності до запиту користувача. Найпоширенішим методом для цього є використання Cross-Encoder моделей, які обробляють запит та документ спільно, як єдину вхідну пару. На відміну від bi-encoder'ів, які оперують лише векторною подібністю між попередньо обчисленими ембедінгами, Cross-Encoder здатні захоплювати більш тонкі смислові взаємозв'язки, синтаксичні залежності та каузальні відношення між словами. Це забезпечує суттєво вищу точність фільтрації.

У практичних системах переранжування часто скорочує початковий набір у десятки або навіть сотні фрагментів до невеликої кількості найбільш релевантних, які потім подаються до

LLM. Новітні дослідження, такі як Rationale-Guided Selection (RGS), розвивають цю концепцію ще далі [4]. Замість простого відбору Top-K фрагментів за евристичними критеріями, RGS залучає саму LLM для пояснення, чому саме певний фрагмент є доказовим для відповіді. Це наближає систему до більш осмисленого, прозорого та стійкого вибору доказової бази, що особливо важливо у високонадійних доменах (медицина, право, наука).

Ефективне подолання контекстної фрагментації не може базуватися на одній техніці — воно потребує системного та багаторівневого підходу. Сучасні рішення комбінують кілька взаємодоповнюючих компонентів: структурно-орієнтоване розбиття (як у Parent Document Retrieval та ієрархічних моделях), багатоступеневу вибірку з гібридними стратегіями та потужне переранжування на основі Cross-Encoder. Разом ці елементи формують архітектурну екосистему, яка значно підвищує якість інформації, що подається на вхід LLM.

Наступним етапом розвитку є створення динамічних архітектур, здатних адаптувати стратегії chunking та retrieval у реальному часі — залежно від типу запиту, його складності та структури інформаційного простору [3]. Наприклад, для запитів, які потребують каузального ланцюжка, система зможе автоматично обирати більші батьківські блоки або звертатися до причинно-наслідкових графів, а для фактологічних запитів — оптимізувати пошук під максимальну точність.

Ключовим напрямом досліджень стає інтеграція структурованих знань, зокрема Knowledge Graphs та Causal Models. Вони дозволяють не просто вилучати фрагменти тексту, а будувати цілісні логічні моделі предметної області, що особливо важливо для складних аналітичних або наукових завдань. Додатковим викликом залишається обробка суперечливих даних у вилученому контексті — ситуацій, коли різні джерела подають взаємовиключну або неоднозначну інформацію. Розв'язання цієї проблеми відкриє шлях до більш стійких, надійних і пояснюваних RAG-систем нового покоління

Список джерел

1. Barnett S. et al. A Survey on Evaluation of Retrieval-Augmented Generation [Електронний ресурс]. 2024. arXiv:2405.07437. URL: <https://arxiv.org/abs/2405.07437> (дата звернення: 10.10.2025).
2. Gao Y. et al. Retrieval-Augmented Generation for Large Language Models: A Survey [Електронний ресурс]. 2023. arXiv:2312.10997. URL: <https://arxiv.org/abs/2312.10997> (дата звернення: 10.10.2025).
3. Ma Y. et al. The Integration of Causal-Counterfactual Reasoning into RAG [Електронний ресурс]. 2025. arXiv:2509.14435. URL: <https://arxiv.org/abs/2509.14435> (дата звернення: 10.10.2025).
4. Sial A. W. et al. METEORA: Replacing Re-ranking with Rationale-guided Selection to Enhance RAG for Sensitive Domains [Електронний ресурс]. 2025. arXiv:2505.16014. URL: <https://arxiv.org/abs/2505.16014> (дата звернення: 10.10.2025).
5. Townsend R. et al. A Systematic Review of Key Retrieval-Augmented Generation Methods, Challenges, and Future Directions [Електронний ресурс]. 2025. arXiv:2507.18910. URL: <https://arxiv.org/abs/2507.18910> (дата звернення: 10.10.2025).

ТЕХНОЛОГІЇ КОМУНІКАЦІЇ ТРАНСПОРТНИХ ЗАСОБИ / VEHICULAR COMMUNICATION TECHNOLOGIES

Яблонський В.Р., Афонін А.О. / Yablonskyi V., Afonin A.

Національний університет “Кієво-Могилянська Академія” / National University of Kyiv-Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики

E-mail: yablonskyivr@ukma.edu.ua, afonin@ukma.edu.ua

Intelligent Transportation Systems (ITS) and vehicular communication technologies are rapidly advancing to improve road safety. Modern vehicles can exchange data with each other, with infrastructure, and with networks or personal devices – a concept known as Vehicle-to-Everything (V2X). Such communication can be established either through decentralized ad-hoc networks (VANETs), or via centralized cloud servers. Various wireless technologies support V2X: the DSRC standard (IEEE 802.11p) pioneered short-range V2V/V2I communication, while the newer Cellular-V2X (C-V2X) offers both direct and network-assisted modes. Researchers are also exploring Wi-Fi based protocols, WiMAX, LoRa for different V2X scenarios. However, widespread adoption remains limited due to the need for additional on-board and roadside units, potential frequency conflicts, and challenges in security and network design. A shift toward a centralized infrastructure combined with decentralized elements is proposed.

У сучасному світі транспортні засоби є невід'ємною частиною повсякденного життя. Відбувається стрімкий розвиток не лише самих транспортних засобів, а й інфраструктури, яку вони використовують, а також технологій, що розширюють можливості транспорту та забезпечують його надійність і безпеку руху. Одним з напрямків, які активно досліджуються, є інтелектуальні транспортні системи (ITS) та комунікація транспортних засобів. Завдяки постійному обміну даними між транспортними засобами, інфраструктурою та людьми ITS можуть виконувати свої основні завдання: покращення дорожньої безпеки та зменшення кількості заторів. Досліджуються такі основні способи комунікації:

1. Vehicle-to-Vehicle (V2V) – коли два чи більше транспортні засоби спілкуються один з одним напряму та обмінюються даними в реальному часі.
2. Vehicle-to-Infrastructure (V2I) – коли транспортні засоби спілкуються з придорожною інфраструктурою, наприклад, світлофорами, шлагбаумами чи іншими елементами.
3. Vehicle-to-Network (V2N) – коли транспортні засоби спілкуються з хмарними серверами, які мають доступ до глобальної мережі Інтернет.

Окрім цього, існують й інші напрямки, такі як Vehicle-to-Pedestrian (V2P) для комунікації з пішоходами та Vehicle-to-Device (V2D) для комунікації з пристроями всередині чи поза транспортним засобом. Усі перелічені варіанти комунікації разом позначаються терміном Vehicle-to-Everything (V2X), який здатен забезпечити одразу декілька типів комунікації, а не лише один.

Загалом існує 2 підходи до побудови такої комунікації між транспортними засобами: використання децентралізованих Ad-Hoc мереж або використання глобальної мережі Інтернет.

Vehicle Ad-Hoc Network (VANET) насамперед розглядають для реалізації V2V та V2I взаємодії. Така мережа передбачає пряму взаємодію транспортних засобів між собою або з придорожною інфраструктурою з використанням додаткових пристроїв для бездротового зв'язку, які вбудовуються або під'єднуються до автомобіля. Цей підхід вимагає створення децентралізованої мережі, де кожен учасник є окремим вузлом. Таким чином можливий як прямий обмін повідомленнями між двома чи більше учасниками, так і передача повідомлень через інших учасників, що виступають у ролі ретрансляторів, якщо пряма комунікація неможлива. Це забезпечує швидкий обмін повідомленнями напряму між учасниками і не потребує додаткових серверів, але створює проблеми, пов'язані з безпекою - як у контексті перехоплення даних, так і спотворення інформації чи спаму.

При використанні глобальної мережі комунікація зазвичай відбувається через хмарний сервер, який здійснює обробку та контроль повідомлень. У такому разі комунікація типу V2V чи V2I відбувається не напряму, а через глобальну мережу Інтернет, тобто транспортний засіб спілкується із сервером, а сервер спілкується з іншим транспортним засобом чи елементом дорожньої інфраструктури. Перевагою такого підходу є централізований контроль, що здатен забезпечувати безпеку комунікації на всіх рівнях. Водночас для цього підходу потрібен доступ до глобальної мережі, і через додаткові вузли і при низькій якості з'єднання він може бути повільнішим, ніж пряма комунікація.

Усі ці види та варіації комунікації можуть бути реалізовані лише завдяки надійним технологіям бездротового зв'язку. Для забезпечення цих цілей досліджуються і тестуються різноманітні засоби і підходи. Далі ми розглянемо основні з них і проаналізуємо їх основні переваги та недоліки.

Dedicated Short-Range Communication (DSRC) тривалий час залишався основним механізмом для забезпечення зв'язку між транспортними засобами та побудови мереж VANET. DSRC - це модифікований стандарт Wi-Fi (802.11p), створений для комунікації між автомобілями. Він працює на діапазоні частот 5,9 ГГц. DSRC може забезпечувати пряму V2V та V2I комунікацію на невеликих відстанях – до 300 м (і до 1 км за сприятливих умов) [1].

Cellular V2X (C-V2X) – відносно нова технологія, яка дедалі більше привертає увагу дослідників. C-V2X може працювати у двох режимах. Перший – це пряма взаємодія (V2V або V2I) із застосуванням технології PC5 sidelink, яка, як і DSRC, працює на діапазоні частот 5.9 ГГц. Другий – взаємодія через інфраструктуру стільникового зв'язку [3]. За умови використання хмарних серверів і підключення інших елементів до глобальної мережі Інтернет цей підхід забезпечує повний спектр взаємодії типу V2X.

Окрім DSRC та C-V2X, ведуться дослідження й експерименти з низкою інших технологій. Наприклад, вивчається можливість використання різних протоколів сімейства Wi-Fi (802.11) для комунікації в умовах руху транспортних засобів. Протокол 802.11p, що лежить в основі DSRC, розроблено на базі стандарту 802.11a [2]. Отже, інші представники сімейства Wi-Fi також заслуговують на увагу дослідників. WiMAX (802.16) – ще одна технологія

бездротового зв'язку, що може застосовуватись для комунікації на великих відстанях (до 50 км). Подібно до цього, технологія LoRa здатна забезпечувати стабільний зв'язок на відстані до 5 км [4]. Крім того, дослідники зосереджують увагу й на технологіях короткого радіусу дії, таких як Bluetooth, ZigBee, NFC. Ці технології розглядають у контексті комунікації транспортного засобу з пристроями, які розташовані всередині нього (V2D), а також з тими, що його оточують (V2P, V2I).

Незважаючи на велику кількість технологій, що пройшли польові випробування, і на те, що дослідження у сфері комунікації транспортних засобів тривають уже досить давно, ці технології досі не набули значного поширення. Це пояснюється багатьма факторами. Один із них полягає в тому, що для більшості технологій потрібні додаткові пристрої, які треба встановлювати як у транспортні засоби, так і в інфраструктуру. Наприклад, для технологій DSRC, C-V2X, LoRa необхідні бортові пристрої (OBUs) у кожному транспортному засобі для забезпечення комунікації типу V2V, а також додаткові стаціонарні пристрої вздовж дороги (RSUs) для взаємодії з інфраструктурою. Утім, такі рішення, у випадку з технологіями DSRC та LoRa, забезпечують лише децентралізовану взаємодію. Щоб отримати доступ до глобальної мережі, ці RSU необхідно підключити до Інтернету. В таких умовах потрібне дуже чітке узгодження дій між виробниками транспорту і муніципальними органами, адже кожна технологія має свою специфікацію. До того ж, різні технології можуть конфліктувати між собою: наприклад, у DSRC та C-V2X перетинаються частотні діапазони, що може створювати перешкоди при одночасній роботі. Інша проблема полягає в тому, що наукова спільнота наразі переважно зосереджена на розбудові децентралізованих мереж. Це загострює зазначені вище проблеми та породжує додаткові труднощі при проектуванні мереж і забезпеченні їхньої безпеки. Отже, зміна фокусу з повністю децентралізованої моделі на використання централізованої інфраструктури з децентралізованими елементами може докорінно змінити загальну ситуацію і прискорити розвиток усієї галузі інтелектуальних транспортних систем. У цьому контексті виникає потреба в проектуванні і розробці відкритої системи та універсального відкритого протоколу, здатних забезпечити єдиний механізм обміну повідомленнями та централізовані засоби безпеки. Це дозволить розробляти застосунки для ITS без прив'язки до конкретних технологій зв'язку чи виробників.

Список джерел: (до 5 джерел, оформлений за стандартом ДСТУ/БАК)

- [1] Kenney J. B. Dedicated Short-Range Communications (DSRC) Standards in the United States [Електронний ресурс] / John B. Kenney // Proceedings of the IEEE. — 2011. — Т. 99, № 7. — С. 1162–1182. – Режим доступу: <https://doi.org/10.1109/jproc.2011.2132790> (дата звернення: 27.10.2025).
- [2] Ott J. Drive-thru internet: IEEE 802.11b for "automobile" users [Електронний ресурс] / J. Ott, D. Kutscher // IEEE INFOCOM 2004, Hong Kong, PR China. – [Б. м.]. – Режим доступу: <https://doi.org/10.1109/infcom.2004.1354509> (дата звернення: 27.10.2025).
- [3] Wang X. An Overview of 3GPP Cellular Vehicle-to-Everything Standards [Електронний ресурс] / Xuyu Wang, Shiwen Mao, Michelle X. Gong // GetMobile: Mobile Computing and Communications. – 2017. – Т. 21, № 3. – С. 19–25. – Режим доступу: <https://doi.org/10.1145/3161587.3161593> (дата звернення: 27.10.2025).
- [4] Zadobrischi E. Enhancing Scalability of C-V2X and DSRC Vehicular Communication Protocols with LoRa 2.4 GHz in the Scenario of Urban Traffic Systems [Електронний ресурс] / Eduard Zadobrischi, Ștefan Havriliuc // Electronics. – 2024. – Т. 13, № 14. – С. 2845. – Режим доступу: <https://doi.org/10.3390/electronics13142845> (дата звернення: 27.10.2025).

ДО ПИТАННЯ УЗАГАЛЬНЕНОГО МЕТАПРОГРАМУВАННЯ В C++ / TOWARDS THE ISSUE OF TEMPLATE METAPROGRAMMING IN C++

Бублик В. В., Трохимчук А. А. / Boublik V., Trokhymchuk A.

Національний університет “Кієво-Могилянська Академія” кафедра мультимедійних систем /
National University of Kyiv-Mohyla Academy
04655, Київ, вул. Григорія Сковороди, 2,
факультет інформатики,

E-mail: boublik@ukma.edu.ua, a.trokhymchuk@ukma.edu.ua

The paper explores C++ metaprogramming as a form of declarative functional programming that originates from Alonzo Church's lambda calculus. Paper describes the imperative approach based on iteration with the functional style centered around recursion. The core of

C++ metaprogramming, the template system, is analyzed. Key techniques such as *Substitution Failure Is Not An Error* (SFINAE) and template specializations are explored in detail. The methods enable compile-time computations, turning the template language into a Turing-complete functional programming language within C++ itself. As a practical demonstration, the paper presents a modern solution to the "Abstract Factory" design pattern problem, originally formulated by A. Alexandrescu, showcasing the power of template metaprogramming in C++.

Ключові слова: метапрограмування, мова шаблонів, C++, SFINAE.

Програмування історично розвивалося у двох основних стилях: імперативному та декларативному (зокрема, функціональному). Ітерація є ключовим засобом організації обчислень в імперативному програмуванні, яке, за своєю суттю, є програмуванням над комірками пам'яті. На противагу цьому, рекурсія є основним інструментом функціонального (декларативного) програмування, яке не вимагає явного керування пам'яттю та операцій присвоєння.

Мовою C++, яка традиційно розглядається як мова імперативного програмування, концепції функціонального стилю реалізовані через механізм метапрограмування шаблонами. Це так зване «статичне метапрограмування», де всі обчислювальні процеси відбуваються на етапі компіляції, до початку виконання основної програми. Такий підхід дозволяє значно зменшити навантаження на етапі виконання коду, переклавши обчислення на час компіляції. Власне, програма компілюється один раз, а виконується тисячі чи мільйони разів, тому такий підхід дозволяє зекономити значну кількість ресурсів.

Розглянемо класичну задачу обчислення найбільшого спільного дільника (НСД). Традиційна рекурсивна реалізація мовою C++ виконує обчислення під час роботи програми:

```
int gcd(const int m, const int n) { return (n == 0) ? m : gcd(n, m % n); }
```

Проте, використовуючи метапрограмування шаблонами, ці обчислення можна перенести на етап компіляції. Для цього створюється основний шаблон та його часткова спеціалізація, що слугує умовою завершення рекурсії:

```
// Основний шаблон
template<int M, int N> struct GCD {
    static consteval int gcd() {
        return GCD<N, M%N>::gcd();
    }
};
// Часткова спеціалізація для завершення рекурсії
template<int M> struct GCD<M, 0> {
    static consteval int gcd() { return M; }
};
```

У такому випадку виклик `int n = GCD<168, 392>::gcd()` змусить компілятор самостійно обчислити результат, і в об'єктний код потрапить лише готовий результат. Це і є суть статичного метапрограмування: програма, написана розробником, генерує і виконує кінцевий код на етапі компіляції, жодних накладних витрат під час виконання програми. В результаті мова шаблонів C++ перетворилася з простого засобу узагальненого програмування для створення контейнерів на повноцінну функціональну мову, повну за Тюрінгом [1], завдяки двом ключовим механізмам: SFINAE та спеціалізації.

SFINAE («*Substitution Failure Is Not An Error*») [2] — це правило, згідно з яким невдала підстановка шаблонного параметра не є помилкою компіляції. Замість цього відповідний шаблон просто виключається з розгляду. Це дозволяє створювати перевантажені функції, які компілятор обиратиме залежно від характеристик типів, що підставляються. Наприклад, можна створити два конструктори для класу рядків, один з яких буде обрано, якщо довжина рядка дозволяє розмістити його у внутрішньому буфері (реалізація *Small String Optimization*), а інший — якщо потрібно динамічно виділяти пам'ять. SFINAE є основою для реалізації умовної логіки на етапі компіляції.

Спеціалізація шаблонів — це механізм, що дозволяє перевизначити поведінку шаблону для конкретних типів або значень. Спеціалізація буває повною, коли надано всі шаблонні параметри (по суті, спеціалізація одного конкретного випадку), і частковою. Яскравим прикладом є `std::vector<bool>`, який є частковою спеціалізацією `std::vector<T>`. Замість того, щоб зберігати кожне булеве значення в окремому байті (що призвело б до 700% надлишкових витрат пам'яті), спеціалізована версія використовує бітові поля, зберігаючи 8 значень в одному байті.

Спеціалізація також відіграє роль умови зупинки в рекурсивних метафункціях. Розглянемо метафункцію для обчислення чисел Фібоначчі:

$$Fibonacci_n = Fibonacci_{n-1} + Fibonacci_{n-2}, n=2,3,\dots$$

Основний шаблон реалізує рекурсивне правило, а спеціалізації для 0 та 1 надають кінцеві значення для зупинки рекурсії.

```
template<int n> struct Fibonacci {
    static consteval int fib() {
        return Fibonacci<n - 1>::fib() + Fibonacci<n - 2>::fib();
    }
};
template<> struct Fibonacci<1> { static consteval int fib() { return 1; } };
template<> struct Fibonacci<0> { static consteval int fib() { return 0; } };
```

Ефективність таких обчислень можна підвищити за допомогою кешування, зберігаючи вже обчислені результати у статичних членах класу, що дозволяє оптимізувати алгоритмічну складність з експоненційної до лінійної.

Сила метапрограмування полягає у вирішенні реальних задач. У 2001 році Андрей Александреску [3] сформулював задачу реалізації шаблону проектування «Абстрактна Фабрика» з єдиним шаблонним методом `Create<T>`. На той час елегантне рішення було неможливим, і доводилося використовувати громіздкі макроси препроцесора. Сучасне метапрограмування дозволяє вирішити цю задачу за допомогою структури даних списку типів (`type_list`). Така фабрика може перевіряти на етапі компіляції, чи входить певний тип до списку продуктів, які вона може створювати, і чи можливо його сконструювати з наданих аргументів, використовуючи SFINAE.

```
type_list<Triangle, Rectangle, Circle> factory{};
Triangle* t = factory.create<Triangle>(3, 4, 5);
```

Такий підхід робить фабрику «дружньою до шаблонів» і дозволяє легко її розширювати, не змінюючи існуючий інтерфейс.

Таким чином, метапрограмування шаблонами в C++ виводить нас за рамки чисто імперативного програмування, дозволяючи виконувати складні обчислення на етапі компіляції, генерувати ефективний код та реалізовувати складні патерни проектування елегантно та безпечно з точки зору типів.

Список джерел:

1. Veldhuizen T. L. C++ Templates are Turing Complete, University of Indiana Technical Report, 2003.
2. Vandevorde D. C++ Templates: The Complete Guide / David Vandevorde and Nicolai M. Josuttis, Addison-Wesley, Boston, MA, 2002
3. Alexandrescu A. Modern C++ design: generic programming and design patterns applied / Andrei Alexandrescu. Addison-Wesley Professional, 2001. 352 p.

DATA STREAMING PIPELINE FOR THE QUADCOPTER FLIGHT CONTROL STACK

T. Zavalij, N. Shakhovska, V. Iatsyshyn

Lviv Polytechnic National University

79000, Lviv, Kn. Romana str., 5, Department of Artificial Intelligence

E-mail: taras.i.zavalii@lpnu.ua, nataliya.b.shakhovska@lpnu.ua, volodymyr.p.yatsyshyn@lpnu.ua

Abstract: Modern UAV and UGV technologies are highly competitive dynamic field of applied research and development. Mainstream software stacks being used include PX4, ArduPilot, Betaflight, as well as ROS2 – a modular computing environment for processing sensor data streams, running perception and control algorithms, as well as software-in-the-loop simulations. In our research, ROS2 integration with Betaflight for on-board sensor data streaming is implemented and tested.

Keywords: *data streaming, digital twin, ROS2, UAV, Betaflight.*

Introduction

The ultimate goal in unmanned aerial vehicles (UAV) research is to enable high-speed perception, mapping and planning, which requires software-in-the-loop and hardware-in-the-loop simulation and training. It will allow full autonomy in multiple applications – surveillance, military, disaster recovery, agriculture and logistics.

Robot operating system (ROS2) is one of the tools for achieving the goal. It's data streaming model consists of topics and nodes as publishers/subscribers, allowing one-to-many distributed communication between system components. We focused on the integration of ROS2 middleware with the mainstream flight control stack to enable data collection and sensor fusion. Our expected outcome was to understand the limitations, data rates, processing latency and standard deviation that can be achieved with the off-the-shelf hardware.

Related work

The open-source tooling and the ecosystem has greatly improved, fuelling wide spectrum of research in UAV flight simulation, perception and control. Two popular flight control software stacks, both PX4 and Ardupilot, are targeted at RnD, industrial and commercial use. Survey [1] has brief introduction into the ecosystem, and there are multiple examples of PX4 integration with companion computers [2], [3].

Betaflight-based flight control stack gained it's reputation as affordable, user-friendly and effective in Ukraine, though not optimal for autonomous applications. However, it was shown in [4], that despite limited support for MAVLink, Betaflight integration with companion computer is possible via MSP protocol. Another example is indoor racing platform running Betaflight with NVIDIA Orin NX companion, ROS2 middleware, sending control commands and reading IMU data via MSP [5].

The gap in Betaflight's integration options presents a good research opportunity, as design and implementation of autonomous flight control algorithms is a crucial task today.

Results

An optimal companion computer in terms of price/productivity in our task was the Orange Pi Zero 3 series. Flight controller GOKU F405S was running Betaflight 4.5.2. Implementation steps needed for our experimental set up on the Companion include but not limited to:

- configure ROS2 docker container with USB serial link to flight controller,
- research MSP protocol structure and message format,
- implement MSP bridge node to poll Betaflight and publish to ROS2 topics,
- insert timestamps into the data stream for latency measurement,
- implement data sink node for ROS2 bag storage and MQTT streaming.

We found the upper rate limit of Betaflight's MSP implementation to be at 100 Hz. Increasing request rate did not increase response rate over this limit. Read rate measurements in Table 1 were obtained at the bridge node using standard ROS2 topic monitoring tools. File write latency was calculated at the Data sink node by inspecting timestamps in ROS2 Bag file and comparing them with original read timestamps. Barometer and IMU packets mean write latency was 2.47 and 8 ms respectively, with standard deviation of 1.6 and 9.5 ms. P95 latency was around 5 ms for barometer and 26 ms for IMU datastream. Note, that IMU packet is larger, as it contains both accelerometer and gyroscope sensor readings.

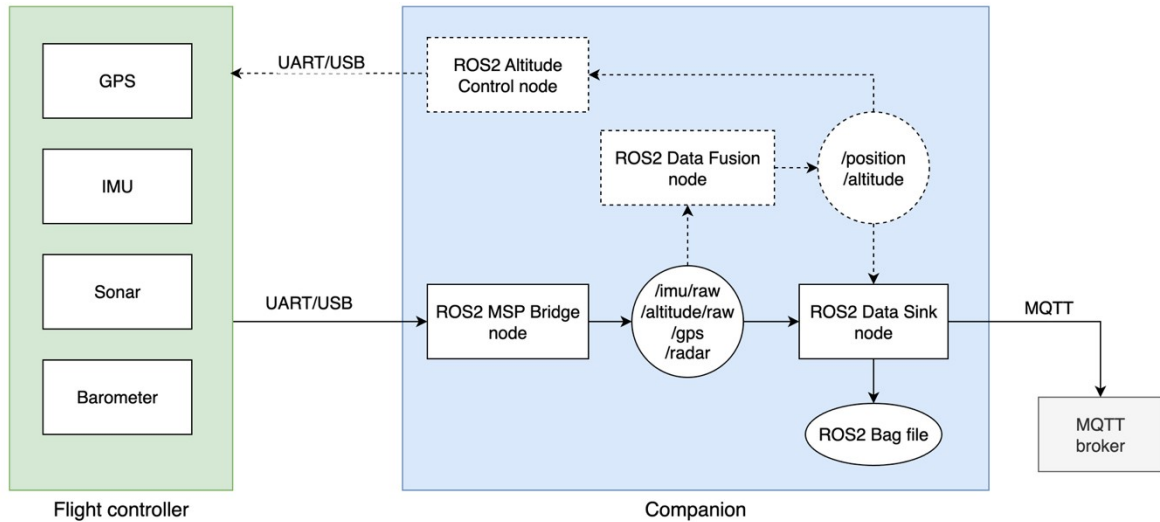


Figure 1. ROS2–MSP integration schema for Betaflight telemetry and control.

Finally, the MQTT data sink was implemented and tested over Wi-Fi connection. Due to the networking overhead the latency was higher: P95 latency of 27 ms for barometer and 125 ms for the IMU. In this case, the MQTT broker (EMQX) on the laptop was used to receive data from the companion computer for analysis, visualization and storage.

Sensor	MSP read rate, Hz	MSP read rate, min, ms	MSP read rate, max, ms	MSP read rate σ , ms	File write latency, mean, ms	MQTT write latency, mean, ms
IMU	99.4	0.1	5.2	0.57	8	105.7
Barometer	99.6	0.1	3.5	0.71	2.47	17.7

Table 1. Metrics for sensor data streams over MSP and MQTT protocols.

Conclusion

Data streaming integration with popular flight control stack was implemented and tested. We showed that MSP protocol in Betaflight has rate limitation. Using pull mode for querying single sensor, we achieved the maximum frequency of up to 100 Hz and measured the latencies for MQTT target. Future research will focus on sensor data fusion for reliable navigation in complex environments, as well as running quadcopter control algorithms on companion computer.

References

1. N. Aliane, “A Survey of Open-Source UAV Autopilots” *Electronics*, vol. 13, no. 23, p. 4785, Jan. 2024, doi: 10.3390/electronics13234785.
2. L. Meier, D. Honegger, and M. Pollefeys, “PX4: A node-based multithreaded open source robotics framework for deeply embedded platforms,” in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, May 2015, pp. 6235–6240. doi: 10.1109/ICRA.2015.7140074.
3. J. Galvez-Serna, F. Vanegas, S. Brar, J. Sandino, D. Flannery, and F. Gonzalez, “UAV4PE: An Open-Source Framework to Plan UAV Autonomous Missions for Planetary Exploration,” *Drones*, vol. 6, no. 12, p. 391, Dec. 2022, doi: 10.3390/drones6120391.
4. D. Sazonov, “FPV autonomous operation with Betaflight and Raspberry Pi,” ILLUMINATION. Accessed: Sept. 16, 2025. [Online]. Available: <https://medium.com/illumination/fpv-autonomous-operation-with-betaflight-and-raspberry-pi-0caeb4b3ca69>
5. M. Bosello *et al.*, “Race Against the Machine: a Fully-annotated, Open-design Dataset of Autonomous and Piloted High-speed Flight,” Feb. 24, 2024, *arXiv*: arXiv:2311.02667. doi: 10.48550/arXiv.2311.02667.

Генеративний фреймворк для побудови візуально-текстових датасетів на основі онтологій

Чоловський С.О., Здырко В.В. / Cholovskyi S. Zdyrko V.

Національний університет “Києво-Могилянська Академія” / National University of Kyiv-Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра інформатики

E-mail: s.cholovskyi@ukma.edu.ua, v.zdyrko@ukma.edu.ua

This work presents a general framework for generating VQA (Visual Question Answering) datasets across arbitrary knowledge domains. Logically complex questions are derived from OWL-ready formatted ontologies, and correct answers are obtained using SPARQL queries. The diversity of generated questions is enhanced through paraphrasing with a large language model. Relevant scenes are generated using Stable Diffusion with CLIP-score-based post-filtering. We believe this hybrid approach enables efficient creation of high-quality, semantically rich datasets.

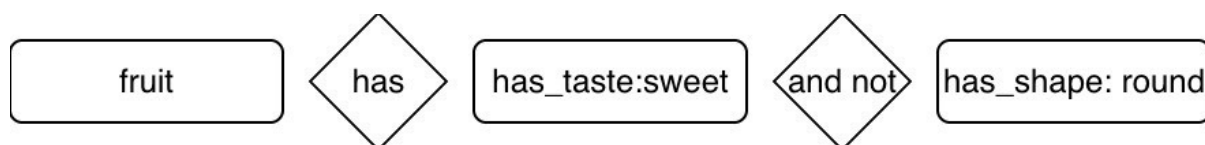
Одним з напрямків комп’ютерного зору є візуально-лінгвістична обробка зображень, він включає в себе, зокрема, такі задачі як опис(captioning), відповіді на питання за зображенням(далі VQA - visual question answering), та загальне розуміння(visual common sense reasoning). Задача VQA полягає в тому, щоб на основі зображення надати коротку відповідь на розгорнуте питання (в common sense задачах відповідь може бути довільною). Постановка задачі VQA є більш простою ніж у задачі опису зображень яка полягає в тому, щоб з заданим зображенням будувати його текстовий опис. Збір даних для задачі Visual Question Answering (VQA) є складним і ресурсоємним процесом, оскільки вимагає ретельно узгоджених наборів зображень, запитань і відповідей. Одним із перспективних напрямів часткової автоматизації цього процесу є використання онтологічних моделей та методів логічного виведення, що забезпечує можливість автоматичного формування запитань із наперед визначеними правильними відповідями. Для підвищення варіативності як змісту запитань, так і візуальних сцен, доцільним є застосування генеративних нейронних мереж, здатних синтезувати нові текстові формулювання запитань та створювати синтетичні зображення, які містять релевантну інформацію для відповіді.

Ми пропонуємо підхід що дозволяє автоматично будувати VQA -датасети на основі онтологій. За основу взято підхід описаний в[1]. Для побудови питань онтологія має відповідати формату owlready2. Першим кроком є побудова табличного представлення онтології, для чого обробляється ієрархія класів, кортежі відношень та пари об’єктів. Формування списку словників сутність відношення з ієрархії, для кожного рівня ієрархії. Перетворення словника в таблицю з п’яти стовпців класи об’єктів, класи атрибутів, сутності, атрибути та відношення. Для генерації запитань із таблиці випадковим чином обираються концепти, ролі та сутності, щоб сформувати блок речення. питання формуються шляхом поєднання блоків логічними операторами. Правильна відповідь формується на основі SPARQL запиту. Результатом є пара первинне питання, та набір атомарних сутностей, що є відповідями на це питання. Нехай Q — множина усіх можливих первинних питань, сформованих алгоритмом, а O — множина усіх об’єктів (сутностей) у розгорнутій онтології (базі знань). Результат роботи блоку R формалізується як множина пар:

$$R = \{ (q, L) \mid q \in Q, L \subseteq O \}$$

Де:

- q — це **первинне питання** (string/текст), що сформоване за шаблоном(рис.1).
- L — це **множина об’єктів-відповідей** (set), де кожен об’єкт $o \in L$ задовольняє **усі логічні умови**, виражені у питанні q .



Pattern: What [concept] + <logic operation> + [attributes] + <logic operation> + [attributes]

Example: What fruit has sweet taste and not round shape?

рисунок 1. Схема типового шаблону, із вибором випадкового параметру з таблиці, з довільним логічним оператором. Приклад сформованого початкового питання.

Попри те, що автоматично згенеровані питання абсолютно коректні, вони не варіативні. Ми використовуємо велику мовну модель для перефразування первинного питання q і отримання q_{para} . Цей прийом вже успішно використовується з image captioning[3]. Наведемо приклад згенерованих варіацій (для онтології з оригінальної статті). Первинне питання “Which food on photo has cylindrical shape and has nutritious fats” та його варіанти “Which food in the photo is cylindrical in shape and contains nutritious fats?” та “Which cylindrical food in the photo is rich in healthy fats?”.

Для генерації зображень ми використовуємо мережу StableDiffusion. Для вибору об’єктів, що мають бути на зображенні, з L випадковим чином обирається набір I (**included**) від одного до чотирьох атомарних об’єктів, решту об’єктів позначимо як E (**excluded**). До набору I також включаємо зашумлюючі об’єкти $N \subseteq L^C$. Шаблон запиту генерації “{general scene description} with {objects from I }, {objects from N }.” Для фільтрації результатів ми відсікаємо CLIP-score згенерованого зображення з кожною з ключових частин нашого запиту та для усіх об’єктів з E . Таким чином забезпечується відповідність зображення домену та наявність відповіді при збереженні певної варіативності.

В результаті виконання отримуємо триплет: $(q_{para}, image, I)$ та проміжні результати (L, P, q, N, E) . Для подальшого оцінювання описаним методом побудовано набір з 1000 зображень, питань та відповідей до них. Дані будуть оцінені експертами, а для зображень також будуть обчислені такі кількісні показники як FID(відстань сприйняття за Фреше), IS(inception score) та MUSIQ.

Таким чином, запропонований фреймворк включає в себе інструменти для генерації та розширення VQA-дасетів. Розділення етапів генерації дозволяє при потребі гнучко замінити їх необхідними для конкретної задачі. Наприклад, використання LLM дозволяє отримати запитання новою мовою, без зміни логіки побудови питань та базової онтології. Варто відмітити, що сучасний стан розвитку генеративних моделей накладає певні обмеження на варіативність отриманих даних. Якість перефразування деградує зі зростанням складності та довжини питань. Дифузійні моделі демонструють обмежену здатність до генерації складних сцен, тому було додано механізм, що забезпечує узгодженість згенерованих зображень із відповідними питаннями та відповідями.

Джерела

1. LoRA: a logical reasoning augmented dataset for visual question answering / Jingying Gao [et al.] // Proceedings of the 37th International Conference on Neural Information Processing Systems. – Red Hook, {NY}, {USA}, 2023. – P. 30579–30591.
2. VQA: Visual Question Answering [Electronic resource] / Aishwarya Agrawal [et al.]. – Mode of access: <https://doi.org/10.48550/arXiv.1505.00468> (date of access: 02.11.2025). – Title from screen.
3. Xiao C. Multimodal Data Augmentation for Image Captioning using Diffusion Models [Electronic resource] / Changrong Xiao, Sean Xin Xu, Kunpeng Zhang // Proceedings of the 1st Workshop on Large Generative Models Meet Multimodal Applications. – [S. l.], 2023. – P. 23–33. – Mode of access: <https://doi.org/10.1145/3607827.3616839> (date of access: 02.02.2024). – Title from screen.

ПОРІВНЯЛЬНИЙ АНАЛІЗ DJANGO I EXPRESS.JS: КРИТЕРІЇ ВИБОРУ ТА СФЕРИ ЗАСТОСУВАННЯ У WEB-РОЗРОБЦІ / COMPARATIVE ANALYSIS OF DJANGO AND EXPRESS.JS: CRITERIA FOR SELECTION AND APPLICATION AREAS IN WEB DEVELOPMENT

Печкурова Е.Я., Печкурова О.М./ Piechkurova E., Pyechkurova O.

Національний університет “Києво-Могилянська Академія” / National University of Kyiv-Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра мультимедійних технологій

E-mail: piechkurova@ukma.edu.ua, pyechkurova@ukma.edu.ua

The article is dedicated to the analysis and justification of the selection of frameworks for web application development. It examines the types of web applications and the main criteria for choosing a framework, such as technology stack, tools, security, and flexibility. A comparison of two popular frameworks, Django and Express.js, is conducted, describing their architectural features, advantages, and disadvantages. The article emphasizes the importance of the appropriate choice of framework depending on the project type, performance and security requirements, and developer qualifications. The article may be useful for developers and researchers considering the choice of a technological platform for web development.

У сучасній web-розробці використання фреймворків є невід’ємною складовою ефективною реалізації web-застосунків. Фреймворки забезпечують розробникам набір інструментів і шаблонів для спрощення створення, підтримки та масштабування проєктів. Основна мета їх використання полягає в автоматизації повторюваних задач, покращенні організації коду та підвищенні безпеки застосунків.

Типи web-застосунків та вимоги

Web-застосунки реалізують клієнт-серверну архітектуру, де клієнтська частина взаємодіє через браузер, а сервер обробляє логіку і дані. Різновиди web-застосунків включають SPA (single-page applications), які забезпечують плавну динамічну взаємодію користувача без перезавантажень сторінок, та PWA (progressive web apps), які підтримують офлайн-режими та встановлення на пристрій. Незалежно від типу, усі застосунки мають спільні завдання: маршрутизація, автентифікація, взаємодія з базою даних, захист від атак.

Критерії вибору фреймворку

Основні критерії оцінки фреймворків для веб web-розробки:

- Походження і технологічний стек: мова програмування, відкритість і ліцензія, активність спільноти.
- Інтегровані інструменти: маршрутизація, ORM, шаблонізатори, механізми безпеки.
- Гнучкість: підтримка архітектурних патернів, можливість кастомізації, контроль структури проєкту.
- Безпека: вбудовані засоби захисту від атак (CSRF, XSS, SQL injection), автентифікація та авторизація.
- Практичне застосування: популярність, сфери використання, підтримка в індустрії, документація.

Django — це фулстек web-фреймворк на Python, створений з ідеєю «batteries included», що означає наявність великої кількості вбудованих інструментів для швидкої розробки. Він застосовує архітектуру MVT (Model-View-Template), підтримує власну ORM для роботи з реляційними базами даних, має вбудовані засоби маршрутизації та шаблонізації (Django Template Language). Безпека реалізована на високому рівні зі стандартним захистом від типових web-атак, посиленням механізмом автентифікації, хешуванням паролів і системою дозволів. Django ідеально підходить для застосунків із типовою архітектурою, де потрібна висока швидкість розробки, масштабованість і безпека.

Express.js — мінімалістичний фреймворк для Node.js, написаний на JavaScript. Він забезпечує основу для побудови бекенду, зосереджуючись на маршрутизації і middleware для обробки HTTP-запитів. Express.js не містить вбудованих ORM, системи безпеки чи шаблонізатора, надаючи розробнику максимальну свободу вибору архітектури і сторонніх компонентів. Безпека реалізується через підключення сторонніх бібліотек, наприклад Helmet.js чи Passport.js. Фреймворк підходить для проєктів, що потребують гнучких, кастомізованих рішень та контролю за стеком технологій, добре підтримує асинхронну обробку запитів і завоював широку популярність у індустрії.

Вибір фреймворку має базуватися на характеристиках конкретного проєкту, кваліфікації розробників та вимогах до функціональності й безпеки. Django забезпечує ґрунтовний набір

готових рішень і високу безпеку, підходить для проєктів класичної архітектури. Express.js надає гнучкість і свободу вибору, але вимагає більшого рівня технічної компетенції та уваги до підбору сторонніх інструментів.

1. Choma D., Chwaleba K., Dzieńkowski M. The Efficiency and Reliability of Backend Technologies: Express, Django, and Spring Boot // IAPGOŚ, 4/2023. — P. 73–78. — DOI: 10.35784/iapgos.4279. — Available online: <http://doi.org/10.35784/iapgos.4279> (дата звернення: 1.11.2025).
2. BitByteSoft. Express JS vs Django: A Comparison of Development Approaches and Performance. Available at: <https://bitbytesoft.com/express-js-vs-django-comparison> (accessed November 1, 2025).
3. IEEE Xplore Digital Library. A Comparative Analysis of Node.js and Django for Scalable Web Applications. IEEE. Режим доступу: <https://ieeexplore.ieee.org> (дата звернення: 1.11.2025).
4. Abhishek Jha. Node.js vs Django: A Performance and Scalability Comparison. Режим доступу: https://d197for5662m48.cloudfront.net/documents/publicationstatus/256355/preprint_pdf/f6990ad1813ca8ba2683cb9ee1b01228.pdf (дата звернення: 11.11.2025).
5. Coursera. Node.js vs Django: What's the Difference? 2025, 23 липня. Режим доступу: <https://www.coursera.org/articles/node-js-vs-django> (дата звернення: 1.11.2025)

РОЗРІЗНЕННЯ БЕЗКОШТОВНИХ ЗАСТОСУНКІВ І ПРОБНИХ ВЕРСІЙ ПЛАТНИХ MACOS-ЗАСТОСУНКІВ НА ПУБЛІЧНИХ МАРКЕТПЛЕЙСАХ / DISTINGUISHING FREE APPS AND TRIAL VERSIONS OF PAID MACOS APPS ON PUBLIC APP MARKETPLACES

Літвінчук З.В., Франків О.О, Петелев Є.Р, Кривоблоцький С.І, Стулова Н.С / Litvinchuk Z.V., Frankiv O.O, Peteliev Y.R., Krivoblotsky S.I., Stulova N.S

Національний університет «Києво-Могилянська академія» / National University of Kyiv-Mohyla Academy

04070, м. Київ, вул. Г. Сковороди 2, каф. інформатики, тел. (044) 425 60 64

E-mail: z.litvinchuk@ukma.edu.ua, o.frankiv@ukma.edu.ua, zhenya.peteliev@macpaw.com, krivoblotsky@macpaw.com, nata.stulova@macpaw.com

The lack of transparency in app monetization on public marketplaces misleads users and complicates policy enforcement. The paper presents an algorithm to distinguish truly free macOS apps from those that involve in-app purchases by classifying apps into free, freemium, and paid categories. The research was based on developing a hybrid method that combined technical analysis of Mach-O binaries, embedded frameworks, and receipt file with language-based processing of local bundle resources and external marketplace text content. It yielded a feature set for a supervised machine learning classifier. To evaluate the suggested solution, we used a dataset consisting of 1,219 apps and benchmarked supervised classifiers. The tree-based ensembles performed best: Random Forest and Gradient Boosting achieved ~90% accuracy, precision, recall, and F1-score on an 80/20 train–test split. The findings contribute to classification systems automation, enhancing applicability and transparency in the software distribution ecosystem.

Стрімкий розвиток маркетплейсів програмного забезпечення привів до їх комерціалізації. Межа між справді безкоштовними застосунками та продуктами з платними обмеженнями розмивається. У цих умовах прозорість монетизації та коректність маркування стають критично важливими: різноманіття моделей монетизації (підписки, внутрішні покупки, пробні періоди), непослідовне розкриття платних функцій ускладнюють ручну перевірку й підвищують ризики для користувачів

і модерації. Це зумовило потребу в автоматизованому підході до категоризації застосунків за моделлю монетизації.

Класифікація застосунків є актуальною незалежно від операційної системи. У [1] виконано класифікацію Android-застосунків методами статичного аналізу текстових рядків із декомпільованих APK та перевірки дозволів, ознаки доповнено даними з Android Market. Для iOS [2] і macOS [3, 4] поєднують статичний і динамічний аналіз: із виконуваних файлів Mach-O виокремлюють використані фреймворки, класи та методи, а під час виконання відстежують доступ до даних і мережевий трафік. Для macOS додатково застосовують структурний аналіз Mach-O.

Для здійснення дослідження застосунки було поділено на три групи:

3. *free*: доступ до повного функціоналу без додаткових плат;
4. *freemium*: базові можливості залишаються безкоштовними, в той час як розширені функції можуть бути розблоковані через внутрішні покупки або підписки;
5. *paid*: застосунок доступний лише після одноразової покупки.

В окремих випадках ці моделі можуть комбінуватись для максимізації бізнес-вигоди.

Реалізований алгоритм складається з трьох послідовних етапів: технічного аналізу, лінгвістичного аналізу та підсумкової класифікації на основі виокремлених ознак.

Під час технічного аналізу опрацьовуються складові пакета macOS-застосунку: виконуваний файл у форматі Mach-O, вбудовані фреймворки разом з їхніми залежностями та, за наявності, файл MASReceipt із даними про покупки в Mac App Store. Основний бінарний файл і фреймворки перевіряються на наявність платіжної логіки: визначаються платіжні фреймворки (StoreKit, сторонні SDK), виокремлюються відповідні класи й методи, аналізуються символні таблиці. У текстових сегментах здійснюється пошук адрес API платіжних сервісів та URL-адрес сторінок оплати за допомогою регулярних виразів і словника доменів; той самий аналіз застосовується до вбудованих фреймворків. Якщо застосунок завантажено з Mac App Store, MASReceipt розшифровується для отримання записів щодо покупок і підписок. На виході формується набір технічних ознак: бінарні індикатори наявності платіжних фреймворків, класів і методів; лічильники адрес API та URL-адрес оплати/підтвердження транзакцій; агреговані показники з MASReceipt.

На етапі лінгвістичного аналізу опрацьовуються локальні ресурси пакета (Info.plist, файли локалізацій, UserDefaults, logs) та матеріали маркетплейсів (описи й відгуки на Mac App Store і MacUpdate), а також контент веб-сторінок політики конфіденційності та умов використання. Зовнішні тексти збираються автоматизовано: для Mac App Store - через iTunes Search API з доповненням вмістом HTML-сторінки; для MacUpdate опис і відгуки агрегуються методом web scraping із відтворенням пагінації. Попередня обробка передбачає токенізацію, лематизацію та розбиття CamelCase-ідентифікаторів. Далі, використовуючи перелік платіжних ключових слів, підраховуються їхні частоти для кожного джерела, формуючи джерелозалежний вектор мовних ознак (ключове слово \times джерело).

Результати технічного й лінгвістичного аналізу об'єднуються в єдину числову таблицю - вхідні дані класифікатора. Датасет зібрано з Mac App Store та MacUpdate, він налічує 1219 macOS-застосунків і поділений у співвідношенні 80/20 на тренувальну та тестову частини. Первинну розмітку виконано з використанням даних агрегатора AppFigures. Розподіл класів: 39,2 % - free, 58,8 % - freemium, 2 % - paid.

Для класифікації використано моделі машинного навчання з учителем: Decision Tree Classifier, Support Vector Machine з лінійним і поліноміальним ядрами, k-Nearest Neighbors, Gaussian Naïve Bayes, Random Forest Classifier, Extremely Randomized Trees та Histogram-based Gradient Boosting. Якість оцінювалась за Accuracy, Precision, Recall і F1 як зважені середні. Найвищі результати показали деревні ансамблі: Random Forest і Histogram-based Gradient Boosting продемонстрували близько 90 % за Accuracy, Precision і Recall та F1 на рівні 0,89–0,90. Поліноміальна версія Support Vector Classifier перевершила лінійну, що відповідає потребі враховувати нелінійні залежності у

змішаному просторі технічних та мовних ознак. Базові деревоподібні та ймовірнісні моделі, а також SVM із лінійним ядром поступилися ансамблям, особливо на малочисельному класі *paid* із низьким *recall*.

Практичне розв'язання проблем непрозорої монетизації дозволяє надійно розмежовувати справді безкоштовні застосунки (*free*) від *freemium* і *paid* та коректно маркувати їх на маркетплейсах. Рішення орієнтоване на інтеграцію у процеси модерації та забезпечення відповідності нормативним вимогам, щоб автоматично виявляти суперечності між заявленою моделлю й фактичною поведінкою застосунку, знижуючи ризики для користувачів і платформи.

Використані джерела:

1. Sanz B., Santos I., Laorden C., Ugarte-Pedrero X., Garcia Bringas P. On the automatic categorisation of Android applications // *2012 IEEE Consumer Communications and Networking Conference (CCNC)*. 2012. P. 149–153. DOI: [10.1109/CCNC.2012.6181075](https://doi.org/10.1109/CCNC.2012.6181075)
2. Bhatt A. J., Gupta C., Mittal S. iABC-AL: Active learning-based privacy leaks threat detection for iOS applications. *Journal of King Saud University – Computer and Information Sciences*. 2021. Т. 33, № 7. P. 769–786. <https://doi.org/10.1016/j.jksuci.2018.05.008>
3. Pajouh, H.H., Dehghantanha, A., Khayami, R. et al. Intelligent OS X malware threat detection with code inspection. *Journal of Computer Virology and Hacking Techniques*, 14. 2018. P. 213–223. <https://doi.org/10.1007/s11416-017-0307-5>
4. Burgardt C. A. P. Malware detection in macOS using supervised learning : dissertation (Master's degree in Computer Science). Recife : Federal University of Pernambuco, 2022.

МОДЕЛЮВАННЯ АСИНХРОННИХ ПРОЦЕСІВ ТА УПРАВЛІННЯ СТАНОМ ЗА ДОПОМОГОЮ МЕРЕЖ ПЕТРІ

Давиденко А.М.

Національний університет «Києво-Могилянська академія»

04655, м. Київ, вулиця Григорія Сковороди, 2, НаУКМА, Факультет інформатики,
andrii.davydenko@ukma.edu.ua

This work explores the formal modeling of asynchronous systems, using classic Place/Transition (P/T) nets and data-driven Coloured Petri Nets (CPNs). Formal models are presented for three distinct processing patterns: a P/T net for a simple mutual exclusion queue, a CPN for context-based concurrent processing, and a CPN for timestamp-based state management.

Моделювання сучасних інформаційних систем, особливо асинхронних та розподілених, вимагає формалізмів, здатних адекватно описувати паралелізм та керувану даними логіку [4, 5]. У цій роботі демонструється застосування мереж Місць/Переходів (М/П) та Кольорових мереж Петрі (КМП) для моделювання попередньо визначених трьох класів розподілених систем [1].

Другий клас розподілених систем моделюється класичною мережею М/П [3]. Ця система моделює просту чергу, де вхідні повідомлення обробляються послідовно. “Маркер-замок” гарантує, що в кожний момент часу може оброблятися лише одне повідомлення (рис. 1, зліва). Мережа М/П визначається як кортеж $N = (P, T, W, M_0)$. Множина місць $P = \{p_{queue}, p_{lock}, p_{proc}\}$, де p_{queue} – буфер необроблених повідомлень, p_{lock} – доступність обробника (м'ютекс), p_{proc} – стан активного оброблення повідомлення. Множина переходів $T = \{t_{start}, t_{finish}\}$, де t_{start} – подія початку оброблення повідомлення, t_{finish} – подія завершення оброблення повідомлення. Вагова функція дуг $W : (P \times T) \cup (T \times P) \rightarrow N$ визначає їх кратність. $W(p_{queue}, t_{start}) = W(p_{lock}, t_{start}) = W(t_{start}, p_{proc}) = 1$, $W(p_{proc}, t_{finish}) = 1$, $W(t_{finish}, p_{lock}) = 1$, $W(a, b) = 0$ для всіх інших пар (a, b) . Початкове маркування $M_0 : P \rightarrow N$ визначає початковий стан мережі. $M_0(p_{queue}) = k$, де $k \geq 0$ – початкова кількість повідомлень. $M_0(p_{lock}) = 1$, $M_0(p_{proc}) = 0$.

Третій клас являє собою чергу оброблення на основі контексту, яка моделюється кольоровою мережею Петрі (КМП). Це модель системи, де повідомлення з різними контекстами можуть оброблятися паралельно, але повідомлення з однаковим контекстом обробляються послідовно (рис. 1, по центру). Кольорова мережа Петрі визначається як кортеж $CPN = (\Sigma, P, T, C, G, E, I)$ [2]. Набори кольорів (Σ) – множина типів даних, що використовуються в мережі. $CTX = S$ (множина всіх рядків, що представляють контексти повідомлень), $\Sigma = \{CTX\}$. Множина місць $P = \{p_{queue}, p_{lock}, p_{proc}\}$. Множина переходів $T = \{t_{start}, t_{finish}\}$. Функція наборів кольорів $C: P \rightarrow \Sigma$, що відображає кожне місце у його тип (колір). $C(p_{queue}) = C(p_{lock}) = C(p_{proc}) = CTX$. Функція охоронних умов $G: T \rightarrow Expr$, що відображає переходи у булеві вирази. У цій моделі всі умови тривіальні: $G(t_{start}) = G(t_{finish}) = true$. Функція виразів дуг $E: A \rightarrow Expr$ (де A – множина дуг), що визначає мультимножини маркерів для переміщення. Оголосимо змінну $var c: CTX$, тоді $E(p_{queue}, t_{start}) = E(p_{lock}, t_{start}) = E(t_{start}, p_{proc}) = E(p_{proc}, t_{finish}) = E(t_{finish}, p_{lock}) = 1'c$. Правило спрацювання вимагає, щоб перехід t_{start} знайшов таке прив'язування (binding) для c , яке одночасно задовольняє обидві вхідні дуги. Початкове маркування задається функцією ініціалізації $I: P \rightarrow CTX_{MS}$, де CTX_{MS} – множина мультимножин над CTX . $I(p_{queue}) = \sum_i 1'ctx_i$ (довільна мультимножина k контекстів повідомлень). $I(p_{lock}) = \sum_{j \in CTX_{active}} 1'ctx_j$ (один “маркер-замок” для кожного можливого активного контексту). $I(p_{proc}) = \emptyset$.

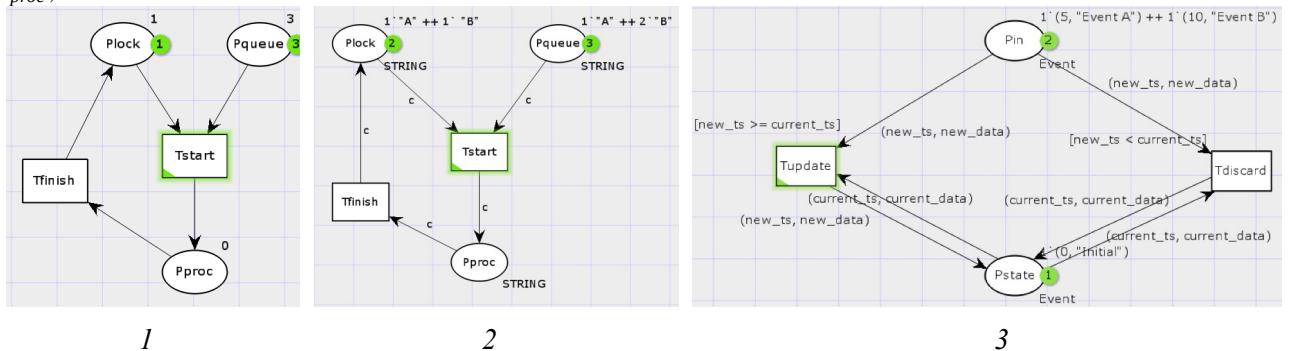


Рис. 1. Моделі мереж Петрі: 1 – модель М/П для послідовного оброблення; 2 – модель КМП для послідовного оброблення в рамках контексту; 3 – модель КМП для збереження найновішого стану

Перший клас моделюється КМП, що представляє монітор стану, де стан системи визначається подією (дані та часова мітка) з найновішою часовою міткою (рис. 1, справа). Вхідні події з більш ранніми часовими мітками відкидаються. Набори кольорів (Σ) : $TS = N$ (множина натуральних чисел для часових міток), $Data = S$ (множина рядків для даних події), $Event = TS \times Data$, $\Sigma = \{TS, Data, Event\}$. Місця $P = \{p_{\square}, p_{state}\}$, де p_{\square} – черга вхідних, ще не оброблених подій, а p_{state} – поточний стан системи, що містить рівно один маркер події. Множина переходів $T = \{t_{update}, t_{discard}\}$, де t_{update} – подія заміни поточного стану новішою подією, $t_{discard}$ – подія відкидання старої вхідної події. Функція наборів кольорів $C(p_{\square}) = C(p_{state}) = Event$. Для визначення охоронних умов спочатку оголосимо змінні, що використовуються у виразах дуг та охоронних умовах: $var new_{ts}, curr_{ts}: TS$; $var new_{data}, curr_{data}: Data$; Охоронні умови визначають ключову логіку моделі: $G(t_{update}) = [new_{ts} \geq curr_{ts}]$, $G(t_{discard}) = [new_{ts} < curr_{ts}]$. Функція виразів дуг $E(p_{\square}, t_{update}) = E(p_{state}, t_{update}) = E(t_{update}, p_{state}) = E(p_{\square}, t_{discard}) = E(p_{state}, t_{discard}) = E(t_{discard}, p_{state}) = 1'(new_{ts}, new_{data})$. Початкове маркування $I(p_{\square}) = \sum_{i=1}^k 1'(ts_i, data_i)$ (довільна мультимножина k вхідних подій). $I(p_{state}) = 1'(0, Initial)$ – маркер з найменшою можливою часовою міткою, що гарантує оновлення при першій реальній події.

Висновки

Аналіз моделей показує, що мережі Петрі дозволяють змодельовати різні класи розподілених систем. М/П мережі ефективні для моделювання простого управління ресурсами та взаємного виключення, проте для логіки, залежної від даних, необхідна виразність КМП.

Подальша робота може включати верифікацію властивостей цих моделей (наприклад, обмеженості, життєздатності та досяжності).

Список літератури

1. Давиденко А. Забезпечення порядку оброблення повідомлень у розподілених системах. *Наукові записки НаУКМА*. 2025. Т. 7 : Комп'ютерні науки. С. 58–62. URL: <https://doi.org/10.18523/2617-3808.2024.7.58-62>.
2. Kristensen L. M., Jensen K. *Coloured Petri Nets: Modeling and Validation of Concurrent Systems*. Springer Berlin / Heidelberg, 2009. 384 p.
3. Murata T. Petri nets: Properties, analysis and applications. *Proceedings of the IEEE*. 1989. Vol. 77, no. 4. P. 541–580. URL: <https://doi.org/10.1109/5.24143> (date of access: 30.10.2025).
4. Peterson J. L. *Petri Net Theory and the Modeling of Systems*. Englewood Cliffs, N.J : Prentice-Hall, 1981. 290 p.
5. Reisig W. *Understanding Petri Nets: Modeling Techniques, Analysis Methods, Case Studies*. Springer, 2013. 260 p.

СЕКЦІЯ 3. ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНІ СИСТЕМИ ТА ТЕХНОЛОГІЇ

ІЄРАРХІЧНА МОДЕЛЬ РОЛЕЙ У CRM-СИСТЕМАХ / HIERARCHICAL ROLE MODEL IN CRM-SYSTEMS

Демидов М.¹, Бабенко Ю.¹, Бабенко М.² / Demydov M.¹, Babenko Y.¹, Babenko M.²

¹Київський національний університет імені Тараса Шевченка / Taras Shevchenko National University of Kyiv

04116, Київ, вул. Богдана Гаврилишина, 24, факультет інформаційних технологій, кафедра кібербезпеки та захисту інформації

²Дніпровський державний технічний університет / Dniprovsky State Technical University 51918, м. Кам'янське, вул. Дніпробудівська, 2, факультет комп'ютерних технологій та енергетики, кафедра програмного забезпечення систем

E-mail: maksimdemidov05@gmail.com, yurii.babenko@knu.ua, mvbab130973@gmail.com

The growing adoption of CRM systems increases the need for effective user access control mechanisms. This paper analyzes the hierarchical role-based access control (RBAC) model in the context of CRM platforms, emphasizing its role in ensuring data protection and centralized access management. A comparative overview of the OAuth Scopes approach is also presented, highlighting their conceptual differences, advantages, and appropriate areas of application. Although OAuth Scopes provide flexible and temporary permissions for API interactions, the hierarchical role model remains a fundamental basis for stable, scalable, and structured internal access management in CRM systems.

Ієрархічна модель ролей у CRM – це система, у якій ролі користувачів мають «вертикальні» взаємовідносини: користувач із вищою роллю може переглядати дані підлеглих, але не навпаки [1]. Основні складові цієї моделі включають: користувача, роль, профіль, а також дозволи та рівні видимості (дивись рисунок 1).

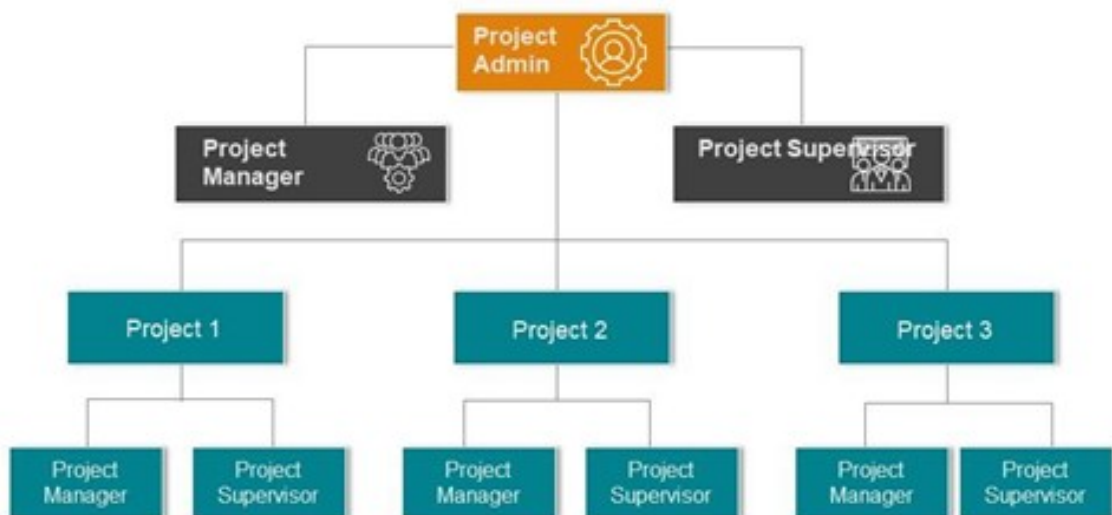


Рисунок 1 – Схема роботи ієрархічної моделі ролей

Ролі визначають, які записи (наприклад, угоди або контакти) користувач може переглядати. Наприклад, регіональний менеджер має доступ до даних своїх підлеглих у межах регіону, але не до інформації інших відділів. Профіль визначає, які дії користувач може виконувати з цими записами – створення, перегляд, редагування, видалення, імпорт чи експорт у різних модулях CRM (ліди, контакти, угоди тощо). Дозволи та видимість встановлюють базовий рівень доступу в межах організації.

Наприклад, певний модуль може бути “Public” (публічний) – доступний усім, або “Private” (приватний) – лише власнику та його керівникам у структурі).

У сучасних інформаційних системах існує кілька підходів до управління доступом користувачів. Найпоширенішими є OAuth Scopes [2, 3] та RBAC (Role-Based Access Control) [3–5] – вони різняться за принципами дії та сферою застосування:

- OAuth Scopes – це рядки дозволів, прив’язані до токенів доступу, які визначають, що саме додаток може робити з даними користувача. Вони тимчасові, більш гнучкі та використовуються переважно для зовнішніх інтеграцій з API, де потрібні короткострокові дозволи.

- RBAC – надає права доступу на основі ролей користувача, групуючи дії відповідно до службових обов’язків. Це статичний підхід, орієнтований на довготривалий внутрішній доступ у корпоративних системах.

Наведемо в таблиці 1 порівняння характеристик обох підходів до управління доступом користувачів.

Таблиця 1 – Порівняння підходів до управління доступом

Характеристика	OAuth Scopes	RBAC
Тривалість дії	Тимчасова (залежить від токена)	Постійна (на основі ролі)
Рівень деталізації	Високий (для конкретних ресурсів)	Загальний (для всієї ролі)
Гнучкість	Висока, налаштовується динамічно	Нижча, потребує ручного оновлення
Типове застосування	Зовнішні API-інтеграції	Внутрішній доступ користувачів

Попри появу нових методів авторизації, ієрархічна модель ролей (RBAC) залишається важливою, адже вона забезпечує чітку структуру доступу, відображає організаційну ієрархію (наприклад, “адмін → менеджер → користувач”), та дозволяє централізовано керувати правами великої кількості працівників.

У системах, де рівень доступу визначається службовими обов’язками, RBAC спрощує адміністрування, підвищує прозорість і забезпечує стабільне довготривале управління доступами.

Хоча OAuth Scopes надають високу точність і безпеку контролю доступу, їхня роль обмежується керуванням конкретними дозволами в межах ресурсів чи API. Вони ефективні для короткострокових технічних сценаріїв, але не замінюють стабільну ієрархічну структуру, яку забезпечує RBAC.

Ієрархічна модель гарантує послідовність, централізований контроль і логічну підпорядкованість прав – це особливо важливо для великих організацій і внутрішніх систем.

Таким чином, ієрархічна модель ролей залишається ключовим елементом сучасних систем безпеки, оскільки забезпечує чітку структуру доступу, що відображає реальну організаційну ієрархію. Вона спрощує управління правами користувачів, знижує ризик помилок у доступах і підвищує контроль над інформацією. Завдяки зрозумілості та масштабованості, ця модель і надалі залишається ефективною основою для побудови безпечних корпоративних систем.

Список джерел

1. Marketing P. C. A complete guide to Roles Hierarchy and Access Controls for sales organisations. *CRM Blog | Pepper Cloud*. URL: <https://blog.peppercloud.com/roles-hierarchy-and-access-controls-for-sales-organisations/> (дата звернення: 20.10.2025).

2. Using OAuth 2.0 scopes vs. permissions for app authorization. *Using OAuth 2.0 scopes vs. permissions for app authorization*. URL: <https://www.aserto.com/blog/scopes-vs-permissions-authorization> (дата звернення: 20.10.2025).

3. Bennett T. OAuth Scopes vs RBAC: Key Differences. *Blog*. URL: <https://blog.dreamfactory.com/oauth-scopes-vs-rbac-key-differences> (дата звернення: 20.10.2025).

4. What is Role-Based Access Control (RBAC)? Examples, Benefits, and More. *Enterprise IP and DLP Software | Digital Guardian*. URL: <https://www.digitalguardian.com/blog/what-role-based-access-control-rbac-examples-benefits-and-more> (дата звернення: 20.10.2025).

5. RBAC vs. ABAC Access Control: What's the Difference? - DNSstuff. *Software Reviews, Opinions, and Tips - DNSstuff*. URL: <https://www.dnsstuff.com/rbac-vs-abac-access-control> (дата звернення: 20.10.2025).

INFORMATION PRACTICE THROUGH DIGITAL ETHICS: WORLDVIEW ASSESSMENT OF AI RELIABILITY

Bagrationi I. O.

Batumi Shota Rustaveli State University

Batumi 6010 / + 995 599 947 668

irma.bagrationi@bsu.edu.ge

Bagratishvili A. N.

Batumi Shota Rustaveli State University

Batumi 6010 / + 995 597 741 855

anri.bagratishvili@gmail.com

Annotation

The paper is devoted to the research and evaluation of ethical regularities of Artificial Intelligence as a certain digital existence in the context of worldview parameters. A critical review of the main provisions and principles surrounding the issue is given, finding and developing the progressive criteria in them and establishing a unique theoretical-worldview concept corresponding to modern requirements. The relationships fixed between the thinking of the fields of humanitarian and exact knowledge around the issue are shown in an original way; a kind of attempt is presented to understand and evaluate the system-conceptual model developed on the basis of moral criteria in the digital world of organized management in a new way.

The discursive resource of the axiological approach and the potential of constructing reflexive conclusions with the dialectical method are utilized and used. It is indicated what probable moral consequences can result from the solution of the intellectual dilemma established in modern digital practice, more specifically, the alternative task of the protocol and software of the creative-argumentative algorithm associated with it. In particular, the latter can only be considered in the format of an admissible possibility or present it in the form of an implemented cybernetic reality. The expected difficulties of interaction between natural and artificial intelligences and appropriate control mechanism paradigms are described and analyzed, namely, the concepts of the ethical value of the artifact are demonstrated - within the framework of essential and principled reliability/unreliability.

Introduction: Challenges in the process of developing moral criteria over time give rise to the need and necessity to review the traditional principles accompanying worldview knowledge. Based on the research question, it is interesting for us to find out what to expect in the field of worldview understanding and evaluation of Artificial Intelligence - from the aspect of ethical dimension. It is significant that two main aspects are still not clearly defined in the field of scientific discourse:

- What is the main function of applied ethics in the practical implementation of constructed knowledge about Artificial Intelligence?
- How well will the system of ethical values work following the software operation of the algorithmic mechanism of the simulated intelligence?

As is generally known, Ethics is a field of humanitarian knowledge, thus it conveys interdisciplinary knowledge and skills on one-world topics and uses open concepts, but when it comes to the concept of Artificial Intelligence, which acts as human intelligence in a certain field, and then Ethics moves from the category of humanitarian thinking to the line of exact sciences. And this is because the computer, as an Artificial Intelligence, must clearly define how to behave, and this is done on the basis of moral considerations, but the same spiritual and moral competences in turn, must already be

recognized in the context of natural intelligence. However, “purely figurative-ideological progress and feat are a direct result of the spiritual abilities of the morals” [1].

The Theoretical Part: As a classic example, the authors mention and reference to the phenomenon of driving means of transport. Naturally, a question arises: how should Artificial Intelligence behave when a critical situation arises? The fact is that even at the global level this has not yet been decided, however in an increasingly algorithm-driven society, the question is not whether we can trust our technology, but how we can assure it deserves our trust. We give an opinion “never put the question about the way information could express “the objective will of the world” and “aspiration for the progress”, when it fails to reflect technological processes, because they are objective and not subjective-psychological phenomena” [2], or in other words, philosophical knowledge is also working on it now, and so is the designer of Artificial Intelligence: The computer must know exactly how to behave, that is, the ethical calculation of the software operation must be embedded in it at the algorithm level, or more concretely and precisely, trustworthy AI advocates for ethical, transparent, and successful AI systems, guaranteeing that we can rely on them to make smart and equitable decisions. Let's suppose that “the manufacturer of one of the well-known car brands had the idea to transfer the driving control to a person in case of a crisis situation. But this cannot become an acceptable option; because it is absolutely clear that the computer mechanism reacts much faster” [3]. And, consequently, no normal person would buy for himself such Artificial Intelligence that he could not trust in principle and rely on and, more importantly, that could not essentially replace the natural capabilities of real intelligence.

Thus, the main goal of the research is to find out to what extent it is possible to axiologically understand and evaluate the multifaceted possibilities of Artificial Intelligence, on the one hand, humanitarian, and, on the other hand, within the framework of mutual understanding and cooperation of exact scientific knowledge. The main task of the research, in its essence content, involves answering the question: to what extent the human being is ready to ethically absorb/adopt and morally manage the countless digital variations and unknown combinations of Artificial Intelligence they have created.

One could postulate that the specific configuration of the education system includes some different types of AI models related to both analytical and generative AI mechanisms – “more recent debates began to inspect the information role of social, cultural, and economic conditions in regulating judgments” [4] An in-depth perception and understanding of digital technologies and modeling them in the practical skills is effective in identifying and solving applied educational problems. The ethical dilemma of digital practice is based on the formation of axiological ideas and value principles of transformation in the informational context of the evolution and mass dissemination of the potential of AI. “It is here we can begin to see how information technology can impact moral values” [5] Formation and activation of the ability of transferable competence it is very steadfastly based on the creativity of software resource used by AI.

Conclusion: From the above-mentioned we may conclude:

Attempts to construct standard moral-ethical patterns and templates that predict the use of AI in a safe and reliable context are defined and developed at the level of theoretical-worldview discourse.

Maneuvering with basic concepts and ideas of Applied Ethics, in the process of continuous development of digital technologies, as a regulatory factor of Artificial Intelligence, is argued and evaluated. It was found that in response to the communicative challenges of the Artificial Mentality - both in the objective risky situation and in the subjective dangerous conditions - the balancing of the algorithmic complex is reliably revealed by the principle of moral assurance.

References:

1. Bagrationi Irma, Didmanidze Ibraim, Cultural Studies in Konstantine Kapaneli's Philosophy: Historical Types of Culture and Organotropism, the Journal “Cross-Cultural Studies: Education and Science (CCS&ES)”, Volume 10, Issue I, Publisher: “Beyer Thomas Robert”. Middlebury, 2025. - Pp. 103-112. DOI: [10.24412/2470-1262-2025-1-103-112](https://doi.org/10.24412/2470-1262-2025-1-103-112) - URL: <https://j-ccses.org/wp-content/uploads/2025/03/Irma-Bagrationi.pdf>
2. Bagrationi Irma, Philosophy of Konstantine Kapaneli, Third Edition, Monograph (Part Three: “Synopsis” in English), Publishing House: “Batumi Shota Rustaveli State University”, 2018. – Pp. 39-48. <https://drive.google.com/file/d/1HPJTqJtPLuFhqkVmHy313LqgTPCUyRrx/view>

3. Didmanidze Ibraim, Bagrationi Irma, Yamkova Olga, Zaslavskiy Volodymyr, (Un) Dependable Artificial Intelligence: Digital Practice and Ethical Worldview Evaluation, Proceedings: 14th International IEEE Conference_DESSERT'2024. Greece, Athens. Publisher: "IEEE", 2025/08. DOI: [10.1109/DESSERT65323.2024.11122197](https://doi.org/10.1109/DESSERT65323.2024.11122197) – URL: <https://ieeexplore.ieee.org/abstract/document/11122197>
4. Didmanidze Ibraim, Bagrationi Irma, Information Paradigms of Art from the History of Social Aesthetics, Proceedings of 9th SWS International Scientific Conference on Arts and Humanities_ISCAH 2022, Volume 9, Issue I, Publisher: "SGEM WORLD SCIENCE Scholarly Society", Vienna, 2022. - Pp. 45-47. doi.org/10.35603/sws.iscah.2022/s07.06 <https://drive.google.com/file/d/1OpFvINdxpwD3LGyVpQaTDluK10MCSgZG/view?usp=sharing>
5. Sullins, John P., Information Technology and Moral Values, [Stanford Encyclopedia of Philosophy](#), First published February 12, 2012; Substantive revision November 9, 2018. - URL: <https://plato.stanford.edu/entries/it-moral-values/>

**INFORMATION EDUCATION:
DIGITAL KNOWLEDGE AND ONLINE LANGUAGE LEARNING**

Didmanidze I. Sh.

Batumi Shota Rustaveli State University
Batumi 6010, + 995 599 276 600
ibraim.didmanidze@bsu.edu.ge

Didmanidze T. I.

Batumi Shota Rustaveli State University
didmanidzetengiz@gmail.com

Annotation

The presented paper outlines approaches to worldview researching language education in on-line environments and survey research that has been conducted to date. Is it worthwhile to use the Internet in Language Education? How does Language Education use change in on-line environments? What are the best ways to incorporate e-mail or the World Wide Web into the classroom? We can use several approaches to learn and research about such issues. One is to talk to fellow lecturers and professors. Another is to try practical things out in the classroom and see how they work. And another is to conduct and share scientific-theoretical research.

The paper first briefly addresses some general questions of approach that relate to some educational and academic researches of technical issues. Then it looks in more detail at particular approaches to researching the on-line classroom. Research conducts to better understand the research role and responsibility of the Internet in online language education and learning generally involves five overlapping areas: linguistic features of language teaching, interaction between practical and theoretical problems, attitude, and context and language impact.

The authors of the paper hope that the references cited in this work provide a useful starting point for readers interested in conducting worldview research on on-line learning. Worldview research studies in the area of the Internet will likely yield not general conclusions but rather specific insights into the impact of using particular on-line activities in particular circumstances. An accumulation of such specific insights over time, however, may allow us as teachers to continually improve our understanding of the cognitive role of the Internet in language education and teaching.

Introduction: It is generally known, that scientific research on online language learning may eventually serve broader purposes than providing pedagogical guidance. The large amount of second language interaction occurring on-line and the ease with which this interaction can be electronically archived and analyzed, makes on-line communication a potentially important source for investigating a wide range of issues related to second language acquisition and use. In due time, the power of computers and the Internet may help us not only increase our understanding of the on-line classroom but also unlock broader mysteries of the language learning process.

Research that investigates the linguistic features of on-line communication might ask questions such as these: Is e-mail communication more similar to speaking or to writing? What are the

lexical, syntactic and functional features common to on-line communication? A growing body of first language research has begun to address these questions. Here it needs to be noticed, that the types of educational research fall along the following main continua, among others: experimental versus natural, quantitative versus qualitative, etic versus emic and product versus process. Experimental research is based on the goal of trying to find generalized answers to specific questions. To achieve that, as many variables as possible are held constant so that one or two specific questions can be addressed in tightly controlled experimental situations.

The Theoretical Part: First of all, we need to understand the traditional lexical meaning of the concept of "Digital Education" which is often found in pedagogical literature and regulatory documentation: "the term ["digital education"](#) is understood as ["education"](#) in the [digital age](#), referring to all nuances of meaning of the concept of [education"](#) [3]. Based on our specific goal, it is greatly interesting for us, that product-based research attempts to look at the results of a particular educational situation. Process-based research attempts instead to analyze the processes that take place in the classroom or other learning situations. Often these are combined into process/product research. According to recent research, in epistemological world view context, some "educational procedures exhaust the possibilities of operating with ready-made information fixed in the memory of the machine" [2].

Action research, another important category, exists not on a continuum but as a type in itself. Action research is research carried out by practitioners in order to improve their own practice. It can involve any of the types of research described above. In action research in education, one or more teachers think about a change they would like to bring about in their classroom. Moreover, the universal distribution proposed by digital practice "encodes all possible mathematical knowledge and, when updated based on empirical observations, will in principle converge to the optimal scientific model of our world" [1]. Following the collection and analysis of the data, the lecturer teacher disseminates information to colleagues (for example, in in-service training sessions, newsletter or journal articles, or informal discussions), makes further curriculum decisions based on what has been learned, and plans a new cycle of action research to continue the learning process.

We can apply any of the above approaches to understanding on-line language learning. In many cases, educational research will fall somewhere between the various approaches. It is important, though, to think about which general approach matches your beliefs and is suitable to address the research questions chosen. Research conducted to better understand the role of the Internet in language teaching generally involves five overlapping areas: linguistic features, interaction, attitude, context, and language impact. Research that investigates the linguistic features of on-line intercultural communication might ask questions such as these: Is e-mail communication more similar to speaking or to writing? What are the lexical and functional features common to on-line communication? Does a unique electronic language exist? Thus, data technology sector includes solutions for data management, and products or services that are based on data generated by both human and machines" [5]

Conclusions and Results: The approved way of expressing our findings is to observe that the situated theoretical worldview in Internet space utilizes three overlapping orders of monitoring in the effort to stay on top of what happens around them:

- Routine online monitoring: This refers to the informal, everyday e-mail information and cognitive research work that is part and parcel of other online activities.
- Occasional online monitoring: This order of online educational monitoring activities is prompted and conditioned by particular intercultural communication, encounters or settings, such as foreign language study worldview meetings, and formal and informal online collaborations.
- Intentional online information-seeking: This type of online educational monitoring is carried out in pursuit of a particular known end, such as a single specific cognitive academic decision-making.

One of the key results of our research is that the online worldview activities of sense-making and sense-giving are mostly conducted dialogically, that is through serious conversations with others, carried out in a variety of online forums and opportunities. These conversations constitute one of the central types of research work and are accomplished through diverse internet online media, mainly face-to-face intercultural conversations and e-mail and internet telephone service.

Research studies in this area will likely yield not general conclusions but rather specific insights into the impact of using particular on-line activities in particular circumstances. An accumulation of such specific insights over time, however, may allow us as lecturer teachers to continually improve our understanding of the cognitive role and responsibility of the Internet in language education.

References:

6. Adriaans Pieter, Information, In Electronic "[Stanford Encyclopedia of Philosophy](#)" _First published February 26, 2012; Substantive revision November 1, 2023. - URL: <https://plato.stanford.edu/entries/information/>
7. Bagrationi Irma, For the Ethical Problems of Digital Education in Decision-Making Process, Proceedings of International Scientific Conference: PDMU-2023-XXXVIII, Kiev, 2023. pp. 16-18. - URL: http://www.pdmu.univ.kiev.ua/PDMU_2023/PDMU-2023_End.pdf
8. Didmanidze I., Bagrationi I., Ulanov V., Matrosova N., Chargazia G., The Ethical Transformations of the Technological Systems For Digital Education Management, Proceedings of International Scientific Conference: DTMS'2020, New York, 2021. - Article № 69, pp. 1-7. DOI: <https://doi.org/10.1145/3446434.3446456>
9. Didmanidze I., Bagrationi I., the Issue of Student Distance Communication and Collaboration, the Journal "Cross-Cultural Studies: Education and Science (CCS&ES)", Volume 3, Issue 1, Vermont, 2018, Pp. 6-19. - URL: <http://j-ccses.org/wp-content/uploads/2018/06/Issue-1-2018.pdf>
10. Murrell Paul, Introduction to Data Technologies, First Edition, UK: "Chapman & Hall", 2009. - URL: <https://www.stat.auckland.ac.nz/~paul/itDT/itdt-2019-03-06.pdf>

СИСТЕМА FINGERPRINTING ДЛЯ ВИЯВЛЕННЯ БОТІВ І ШКІДЛИВОЇ АКТИВНОСТІ У ВЕБСИСТЕМАХ / FINGERPRINTING SYSTEM FOR DETECTING BOTS AND MALICIOUS ACTIVITY IN WEB SYSTEMS

Думан А.¹, Бабенко Ю.¹, Бабенко М.² / Duman A.¹, Babenko Y.¹, Babenko M.²

¹Київський національний університет імені Тараса Шевченка / Taras Shevchenko National University of Kyiv

04116, Київ, вул. Богдана Гаврилишина, 24, факультет інформаційних технологій, кафедра кібербезпеки та захисту інформації

²Дніпровський державний технічний університет / Dniprovsky State Technical University 51918, м. Кам'янське, вул. Дніпробудівська, 2, факультет комп'ютерних технологій та енергетики, кафедра програмного забезпечення систем

E-mail: alina.duman@knu.ua, yurii.babenko@knu.ua, mvbab130973@gmail.com

Bots and malicious activity increasingly distort web traffic and harm security and business outcomes. We propose a fingerprinting design that merges deep client and network stacks with behavioral micro-signals and human-interaction checks, aiming for resilient deployment in real-world web systems. The developed method captures client- and server-side signals, forms a normalized feature vector, and leverages an LSTM-based network to identify automated and malicious traffic in real time. This work-in-progress reframes fingerprinting from user identification to bot/malicious-activity detection, combining deep multi-source parameters with human-interaction checks and an LSTM-based core under privacy-first normalization to enable resilient, low-latency decisions in web systems. The allow/step-up/block policy at the traffic edge reduces the load on the backend and improves the quality of analytics data. We expect a reduction in false positives at a fixed detection level and increased resistance to VPN/proxy and cross-browser spoofing.

Вебплатформи все частіше стають мішенню автоматизованих атак: credential stuffing/ АТО, склік і накрутки, сканування, масові реєстрації [1]. За даними Imperva Bad Bot Report, останніми роками автоматизований трафік уперше перевищив людський і сягнув 51%, а частка «поганих» ботів зросла до 37% загального інтернет-трафіку (проти 32% роком раніше) [2].

Атаки стають складнішими: 55% бот-атак мають середню/високу витонченість, 44% просунутого бот-трафіку спрямовано на API, а АТО зросли на 40% за рік (на 54% від 2022); популярні ухили – резиденційні проксі (21%) і маскування під Chrome (46%).

Це напряму б'є по бізнес-метриках і безпеці: спотворюється аналітика й ціноутворення, зростають втрати від шахрайства, ризики витоків і регуляторні штрафи (GDPR до 4% глобального обороту тощо). Отже, потрібні стійкі до спуфінгу засоби відсікання ботів, що працюють без залежності від cookie, менш вразливі до VPN/проксі/інкогніто й придатні для продакшну з низькою затримкою – саме це ми і пропонуємо з використанням fingerprinting методу.

Browser/device fingerprinting – це підхід до ідентифікації та розпізнавання клієнтів у вебсередовищі без stateful-ідентифікаторів (на кшталт cookies). Він спирається на сукупність стабільних або квазістабільних ознак браузера, операційної системи, апаратної платформи та мережевого оточення, що разом формують «цифровий відбиток» сеансу [3–5].

Після поширення практик блокування/обмеження cookies було показано, що комбінації клієнтських параметрів (User-Agent, набір і порядок заголовків, властивості navigator/screen, шрифти, Canvas/WebGL тощо) забезпечують достатню унікальність для стійкої ідентифікації. Цей «безкукісний» підхід поступово доповнили мережевими сигналами (зокрема, параметрами TLS-рукописання) та тонкими рендер-відмінностями.

На практиці виділяють щонайменше три шари:

1. Браузерно-клієнтський: UA, порядок/наявність заголовків, navigator/screen, Canvas/WebGL, список шрифтів, часовий пояс/локаль, індикатори headless/емуляції.

2. Мережевий: версії TLS/ALPN/ciphers, JA3/подібні хеші, варіабельність RTT/TTFB, стабільність ASN, евристики проксі/VPN.

3. Поведінковий (за наявності): мікроінтервали взаємодій (клавіатура/миша/тач), скрол-патерни, «бурстовість» подій.

Боти в більшості спуфлять окремі параметри (наприклад, User-Agent), однак крос-шарова узгодженість множини сигналів відтворюється значно складніше та нестабільна в часі. Саме тому fingerprinting є корисним антифрод/antibot інструментом: він надає ранній сигнал ще на периферії трафіку, дозволяючи застосовувати політики allow / step-up / block до виконання бізнес-логіки.

Fingerprinting у контексті antibot/antifraud дає ранній сигнал ризику на периферії трафіку, не потребує cookies і дозволяє приймати рішення з низькою затримкою до виконання бізнес-логіки; поєднання клієнтських і мережових характеристик ускладнює достовірну емуляцію середовища ботами та зберігає стабільність ідентифікації в часі. Водночас точність знижується у гомогенних середовищах і за умов anti-fingerprinting; частина сигналів залежить від JavaScript, а «життєвий цикл» відбитка обмежений оновленнями та використанням VPN, що ускладнює зв'язання з репозиторієм. Ці обмеження ми пом'якшуємо мультимодальністю (клієнтські, мережеві та поведінкові мікросигнали з перевітками людської взаємодії) і privacy-first нормалізацією, поєднуючи нейромережевий скоринг із короткогоризонтним аналізом послідовностей; далі окреслимо наявні методології.

У практиці виділяють кілька родин технік цифрового відбитка. JavaScript-based ознаки (набір/порядок заголовків, navigator/screen, timezone/locale, список шрифтів, headless-індикатори) дають базову ентропію; CSS/DOM та Canvas/WebGL використовують тонкі відмінності рендеру; Audio API – мікророзбіжності аудіотракту; історично застосовувалися plugin-based (Flash/Java), нині майже неактуальні; hardware/software-based включають драйвери/GPU/версії стеків; TLS/JA3/ALPN формують мережеві сигнатури клієнта незалежні від JS; доповнюють їх поведінкові мікросигнали (інтервали натискань/рухів, скрол-патерни) як короткогоризонтний шар. Типова схема збору – легка клієнтська бібліотека (JS) + серверний/edge-сервіс, що нормалізує параметри у вектор і співставляє з репозиторієм. У нашій роботі ми комбінуємо ці родини у мультимодальний дизайн (клієнтські+мережеві+поведінкові),

додаємо перевірки людської взаємодії та застосовуємо LSTM-ядро з privacy-first нормалізацією для ухвалення рішень allow / step-up / block у реальному часі.

Підсумовуючи, мультимодальний fingerprinting із поєднанням клієнтських, мережових і поведінкових сигналів забезпечує раннє й низьколатентне відсікання ботів без залежності від cookies. LSTM-ядро та короткогоризонтний аналіз послідовностей підвищують чутливість до headless/емуляцій і «low-and-slow» патернів. Політика allow / step-up / block на периферії трафіку зменшує навантаження на бекенд і покращує якість даних аналітики. Очікуємо зниження хибнопозитивних спрацювань за фіксованого рівня виявлення й підвищення стійкості до VPN/проксі та крос-браузерного спуфінгу.

Список джерел

1. Zakutynskiy I., Kalishuk O., Iavich M., Nebylytsia V., Yehunko V. AI-based user identification method for web services. URL: <https://www.researchgate.net/publication/392391397> (дата звернення: 20.10.2025).

2. Imperva Bad Bot Report. URL: <https://www.imperva.com/resources/resource-library/reports/bad-bot-report/> (дата звернення: 20.10.2025).

3. Zhao R. Toward the flow-centric detection of browser fingerprinting. *Computers & Security*. 2023. P. 103642. URL: <https://doi.org/10.1016/j.cose.2023.103642> (дата звернення: 20.10.2025).

4. Tracking your browser with high-performance browser fingerprint recognition model / W. Jiang et al. *China Communications*. 2020. Vol. 17, № 3. P. 168–175. URL: <https://doi.org/10.23919/jcc.2020.03.014> (дата звернення: 20.10.2025).

5. A Survey of Browser Fingerprint Research and Application / D. Zhang et al. *Wireless Communications and Mobile Computing*. 2022. Vol. 2022. P. 1–14. URL: <https://doi.org/10.1155/2022/3363335> (дата звернення: 20.10.2025).

КАЛІБРУВАННЯ ПАРИ РОБОТИЗОВАНИХ МАНІПУЛЯТОРІВ У СХЕМІ «ВЕДУЧА–ВЕДЕНА» / CALIBRATION OF A LEADER–FOLLOWER PAIR OF ROBOTIC MANIPULATORS

Волинець Є.А. / Y. Volynets

Національний університет “Киево-Могилянська Академія” / National University of Kyiv-
Mohyla Academy

04655, Київ, вул. Григорія Сковороди, 2, факультет інформатики, кафедра мережних
технологій

E-mail: ye.volynets@ukma.edu.ua

This work addresses a systems-level reliability gap in leader–follower manipulation: nominally identical commands do not yield identical states. We present a reproducible joint-level calibration that estimates per-joint direction and zero offset from controlled rotations, applies robust angle normalization to avoid wrap-around, and standardizes a small pose set for validation and control. Empirical evaluations indicate that previously significant pose errors are reduced to minor, task-tolerant deviations.

Для багатоланкових роботизованих маніпуляторів із розумними сервоприводами (Dynamixel, Feetech) типовими джерелами похибок є механічні неточності та зазори, люфт, некоректне нульове базування суглобів і порушення орієнтації привода під час складання кінематичних пар. Навіть за однакових алгоритмів керування дві зовні однакові системи можуть суттєво відрізнятися за фактичними координатами, що ускладнює масштабування та обслуговування. У прикладних сценаріях (телеоперування, навчальні стенди) ведучою є дешева, полегшена рука, з якою безпосередньо працює оператор; веденою — дорога, промислова рука, що виконує роботу з об'єктами та має точно відтворювати рухи ведучої. Через відмінності в

кінематичних ланцюгах і жорсткості вузлів навіть малі зсуви нульового базування спричиняють помітну розбіжність траєкторій, тож міжсистемне калібрування є критично необхідним.

Щоб усунути ці відхилення, запропоновано універсальну програмну процедуру калібрування у схемі «ведуча–ведена», яка визначає та компенсує кутові зсуви нульових положень і узгоджує напрям обертання кожного ступеня вільності. Реалізацію виконано мовою Python у вигляді окремого скрипта, що обмінюється даними з маніпуляторами через послідовні інтерфейси (UART/TTL через USB-перетворювач) і формує YAML-параметри калібрування для подальшого використання у системі керування.

Алгоритм калібрування пари «ведуча–ведена» для кожного суглоба визначає знак напрямку D та кутове зміщення нуля S у градусах, щоб ведена рука відтворювала траєкторію ведучої. На першому кроці обидва маніпулятори проводять до однакової початкової пози, уникаючи кутів сервоприводів поблизу $0^\circ/360^\circ$ (з порогом $\approx 20^\circ$), і за потреби механічно зміщуючи ближче до $\sim 180^\circ$. Далі для кожного суглоба оператор послідовно обертає його на обох маніпуляторах приблизно на 45° в однакових напрямках, фіксуючи знак фактичного обертання; якщо знаки збігаються — напрям вважається узгодженим, якщо ні — для цього суглоба задають інверсію напрямку ($D = -1$). Після визначення напрямку обчислюють кутову різницю між суглобами двох рук у стартовій позі, приводять її до «звичного» діапазону ($-180^\circ, 180^\circ$) без стрибків через $0^\circ/360^\circ$ окружності і округлюють до найближчого кроку (практично — до кратних сорока п'яти градусів), щоб зменшити вплив люфту та шуму; за потреби різницю додатково стабілізують усередненням по кількох позах. Окремо для веденої руки знімають реперні пози — «початкові», «робоча» — для подальшого використання в конфігурації. Результатом є набір параметрів для кожного суглоба: чи потрібно інвертувати напрям і яку кутову поправку застосувати, плюс зафіксовані опорні пози; ці дані далі використовуються, щоб ведена (дорога) рука точно відтворювала рухи ведучої (дешевої) попри відмінності в механіці.

Результати записуються до конфігураційного файлу у форматі YAML:

```
joint_1:
  offset_deg: -90
  direction: -1
joint_2:
  offset_deg: 45
  direction: 1
...
```

Під час ініціалізації ці зсуви застосовуються як компенсація, а напрям обертання узгоджується зі знаком ведучої руки.

Реалізований підхід використовує офіційний Dynamixel SDK для доступу до сервоприводів (UART/TTL) та модуль взаємодії з кінематичними ланцюгами; забезпечено синхронізоване опитування обох маніпуляторів, нормалізацію кутів із урахуванням меж обертання, виявлення потреби інверсії напрямку та генерацію YAML-файлу з параметрами компенсації. Застосування такої процедури забезпечило узгоджене відтворення рухів між ведучою (дешевою) та веденою (промисловою) руками попри відмінності в механіці; фактичні відхилення зведено до рівня,

прийнятного для надійного копіювання траєкторій, а тривалість підготовки нової системи помітно скорочено порівняно з ручним налаштуванням.

Список джерел:

1. **DYNAMIXEL SDK: Overview** [Електронний ресурс]. — Режим доступу: https://emanual.robotis.com/docs/en/software/dynamixel/dynamixel_sdk/overview/
2. DYNAMIXEL Protocol 2.0 — Communication Manual [Електронний ресурс]. Режим доступу: <https://emanual.robotis.com/>
3. **ROBOTIS-GIT. DynamixelSDK** : репозиторій програмного забезпечення [Електронний ресурс]. — GitHub. — Режим доступу: <https://github.com/ROBOTIS-GIT/DynamixelSDK>
4. **Hugging Face. LeRobot — Documentation** [Електронний ресурс]. — Режим доступу: <https://huggingface.co/docs/lerobot/en/index>

АВТОМАТИЧНЕ ФОРМУВАННЯ ОНТОЛОГІЇ ТОВАРІВ НА ОСНОВІ АНАЛІЗУ ДАНИХ ЕЛЕКТРОННОЇ КОМЕРЦІЇ

Жежерун О. П., Колесніков А.О.

Національний університет «Києво-Могилянська академія»

вул. Сковороди 2, м. Київ, 04070, Україна, anton.kolesnikov@ukma.edu.ua

The article presents a system for automatic generation of product ontology based on analysis of heterogeneous data from multiple e-commerce sources. The system architecture and algorithm for concept extraction from natural language texts without manual synonym dictionary creation are described. The system generated an ontology with 486 concepts and 1216 relationships with F1=95.2% extraction accuracy. The system uses a four-layer hybrid architecture with transformer embeddings (gte-small, 384-dimensional) and HNSW indexing (M=16, efSearch=16). Experimental deployment on 700,000 products from 34 sources in four languages showed F1=95.2% concept extraction accuracy at 13 products per second processing speed. Main advantages: no need for large labeled datasets, automatic multilingual processing without translation dictionaries, ability to supplement ontology with new concepts without retraining. The system can be adapted for other domains: medicine, finance, logistics.

Keywords: ontology engineering, knowledge base, product ontology, natural language processing, transformer embeddings, semantic matching, big data.

Після створення концепції Semantic Web онтологія стала синонімом рішення проблем розуміння природної мови комп'ютерами [1]. Проте ручне створення онтологій потребує значних інтелектуальних ресурсів та швидко застаріває. Щоб знайти рішення, з'явився напрям онтологічної інженерії, який вивчає шляхи автоматизації генерування знань з тексту [2,3].

У процесі роботи розглянуто задачу автоматизованої генерації онтології товарів з використанням гетерогенних даних з 34 джерел електронної комерції чотирма мовами. Побудовано систему, яка формує онтологію з 486 концептів без ручного створення словників.

Традиційні підходи потребують залучення експертів та ручного створення правил, що не масштабується при роботі з мільйонами товарів [3,4]. Сучасні методи можна поділити на три класи: словникові системи (потребують ручних синонімів), правилові системи (не масштабуються), системи на основі машинного навчання (потребують великих розмічених датасетів) [5].

МЕТОДОЛОГІЯ

Для виокремлення концептів використано трансформерні нейронні мережі, які генерують векторні представлення текстів у багатовимірному просторі. На відміну від класичних підходів (стемінг, лематизація), трансформери захоплюють семантичний зміст та автоматично виявляють близькість концептів у різних мовах без словників перекладу.

У роботі використано модель gte-small (384-вимірні вкладення) з прискоренням на GPU. Для швидкого пошуку застосовано алгоритм HNSW [6], який буде багаторівневий граф навігації зі складністю $O(\log N)$.

Онтологію побудовано у вигляді ієрархічної структури з класами: Brand (59 брендів з 847 варіаціями), ProductType (16 типів з ієрархією), Color (337 відтінків у 12 родин), Material (47 матеріалів з 213 синонімами), Size (23 стандарти зі 156 зв'язками конвертації EU/US/UK/CM), Gender (4 категорії), Product (абстрактний клас з зв'язками до всіх класів).

Система має архітектуру конвеєра (рис. 1) з п'ятьма етапами. Модулі написано мовою Python з використанням sentence-transformers, hnswlib, MongoDB.

Етап 1. Збір даних з 34 джерел (JSON формат, 4 мови: англійська 60%, італійська 25%, іспанська 10%, французька 5%).

Етап 2. Попередня обробка – витягування та очищення тексту, формування запиту.

Етап 3. Генерація вкладень – завантаження моделі gte-small, генерація вкладень для канонічних концептів, побудова HNSW-індексів ($M=16$, $efConstruction=200$, $efSearch=16$, обсяг 4,2 МБ).

Етап 4. Гібридний пошук – чотиришарова архітектура:

- Шар 1: Правила (<1мс, 98% довіра, 45% товарів)
- Шар 2: Псевдоніми (~1мс, 95% довіра, 35% товарів)
- Шар 3: Семантика (~8мс, 70-90% довіра, 18% товарів)
- Шар 4: Валідація (~80мс, 60-80% довіра, 2% товарів)

Адаптивний оптимізатор пропускає шари за довірою: $>0,98 \rightarrow$ шари 3-4 пропускаються (80% товарів), $>0,90 \rightarrow$ шар 4 пропускається (95% товарів).

Етап 5. Доповнення онтології – додавання концептів, встановлення зв'язків, пошук синонімів, ієрархічні зв'язки. Для кольорів: текстова класифікація у 12 родин, за впевненості $<0,7$ – візуальна валідація через ΔE у LAB-просторі. Багатомовні еквіваленти через семантичну близькість (Blu \leftrightarrow Blue \leftrightarrow Azul \leftrightarrow Bleu).

РЕЗУЛЬТАТИ

Датасет: 700 000 товарів з 34 джерел, 4 мови. Обладнання: Apple M1 Max (10-ядерний CPU, 32-ядерний GPU, 32 ГБ RAM). Програмне забезпечення: Python 3.10, sentence-transformers, hnswlib, MongoDB.

Результат генерації. Система згенерувала онтологію: 59 брендів (847 варіацій), 16 типів (ієрархія до 3 рівнів), 337 кольорів (12 родин), 23 розміри (156 конвертацій), 4 статі, 47 матеріалів (213 синонімів). Загалом 486 концептів та 1216 зв'язків.

Для оцінки точності створено еталонну розмітку на 1000 товарів з усіх 34 джерел.

Концепт	Кількість	Зв'язків	Точність	Повнота	F1-міра
Brand	59	847	97,50%	96,80%	97,10%
ProductType	16	128	96,80%	95,40%	96,10%
Color	337 (12)	89	95,20%	93,80%	94,50%
Gender	4	64	95,70%	94,20%	94,90%
Size	23	156	94,30%	92,70%	93,50%
Material	47	213	91,80%	89,40%	90,60%
Всього	486	1216	95,20%	93,70%	94,50%

Таблиця 1. Результати автоматичної генерації онтології

Продуктивність: 581 товар за 44 секунди (13 тов/с). Пам'ять: HNSW 4,2 МБ, концепти 28 МБ, модель 133 МБ (165 МБ загалом).

Головна перевага системи – відсутність потреби у розмічених датасетах та можливість роботи з новими концептами без перенавчання. Система автоматично виявляє еквіваленти у різних мовах через семантичну близькість ($\cosine > 0,95$), що дозволяє інтегрувати товари від міжнародних постачальників без словників перекладу.

ВИСНОВКИ

Розроблена система автоматично генерує онтологію товарів без ручного створення правил та словників синонімів. Експериментальне розгортання на 700 000 товарів з 34 джерел чотирма мовами показало точність $F1=95,2\%$ при швидкості 13 товарів за секунду.

Система згенерувала онтологію з 486 концептів та 1216 зв'язків. Гібридна чотиришарова архітектура з адаптивним вибором забезпечує баланс між швидкістю та точністю.

Головні переваги: відсутність потреби у розмічених датасетах, автоматична багатомовність без словників перекладу, доповнення онтології без перенавчання. Систему можна адаптувати для інших галузей: медицина, фінанси, логістика.

ЛІТЕРАТУРА

1. Berners-Lee T. The Semantic Web. Scientific American. 2001. Vol. 284, no. 5. P. 34–43.
2. Biemann C. Ontology Learning from Text: A Survey of Methods. 2005. URL: https://www.researchgate.net/publication/200044378_Ontology_Learning_from_Text_A_Survey_of_Methods (дата звернення: 15.11.2024).
3. Жежерун О. П., Репкін М. С. Автоматична генерація онтологій на основі статей українською мовою. Наукові записки НаУКМА. Комп'ютерні науки. 2022. Том 5. С. 12–15. DOI: 10.18523/2617-3808.2022.5.12-15. URL: https://www.researchgate.net/publication/369350039_Automatic_Generation_of_Ontologies_Based_on_Articles_Written_in_Ukrainian_Language (дата звернення: 15.11.2024).
4. Chen Q., Lin J., Zhang Y. et al. Towards Knowledge-Based Personalized Product Description Generation in E-commerce. KDD 2019. URL: <https://arxiv.org/abs/1903.12457> (дата звернення: 15.11.2024).
5. Papadakis G., Efthymiou V., Thanos E. et al. An analysis of one-to-one matching algorithms for entity resolution. The VLDB Journal. 2023. Vol. 32. P. 1369–1400. URL: <https://doi.org/10.1007/s00778-023-00791-3> (дата звернення: 15.11.2024).
6. Malkov Y., Yashunin D. Efficient and robust approximate nearest neighbor search using HNSW graphs. IEEE TPAMI. 2020. Vol. 42, no. 4. P. 824–836. URL: <https://doi.org/10.1109/TPAMI.2018.2889473>