

ВІДГУК

офіційного опонента на дисертаційну роботу
Кравчука Олександра Анатолійовича
на тему «Метод оцінювання ризиків життєвого циклу моделей
штучного інтелекту геопросторового сприйняття для інформаційних
підсистем роїв безпілотних літальних апаратів»,
представлену на здобуття ступеня доктора філософії
в галузі знань 12 – Інформаційні технології
за спеціальністю 122 – Комп'ютерні науки

Актуальність теми дисертації

ШІ-моделі геопросторового сприйняття є важливою складовою інформаційних підсистем автономних і ройових комплексів. Вони перетворюють дані спостереження на описи об'єктів, зон, перешкод і можливих маршрутів. Помилки такого перетворення можуть впливати на окремого агента, а через механізми обміну й координації – на колективний стан рою.

Оцінювання моделі лише за показниками точності на етапі тестування не охоплює всього спектра загроз. Ризики можуть виникнути ще під час добору даних, формулювання задачі, визначення архітектури, навчання й валідації. Після розгортання додаються доменний зсув, сенсорні деградації, зміна умов середовища, обмеження зв'язку та помилки взаємодії оператора із системою.

У цих умовах потрібен процес, який дає змогу не лише скласти перелік ризиків, а й пов'язати їх із життєвим циклом, визначити пріоритети, перевірити стійкість оцінок і використати результати для планування тестування та валідації. Саме цій проблемі присвячено дисертацію О. А. Кравчука. Обрана тема є актуальною для комп'ютерних наук, інженерії програмного забезпечення, багатоагентних систем і застосування штучного інтелекту в критичних інформаційних підсистемах.

Ступінь обґрунтованості наукових положень, висновків і рекомендацій

Наукове завдання дисертації полягає в розробленні методу комплексного оцінювання ризиків, що виникають на різних стадіях життєвого циклу ШІ-моделей геопросторового сприйняття для інформаційних підсистем роїв БпЛА, з урахуванням можливості поширення в ройовій системі локальних помилок сприйняття через взаємодію агентів і

формування внаслідок цього системних ризиків порушення колективної поведінки рою.

Метою дослідження є забезпечення обґрунтованого визначення пріоритетних ризиків, що виникають на різних стадіях життєвого циклу ІІІ моделей геопросторового сприйняття для інформаційних підсистем роїв безпілотних літальних апаратів, та напрямів зниження цих ризиків шляхом розроблення методу комплексного оцінювання, який охоплює їх ідентифікацію, 2 класифікацію і кількісне пріоритизування, перевірку стійкості результатів оцінювання та врахування системних наслідків поширення локальних помилок сприйняття для колективної поведінки рою.

Об'єктом дослідження є процес комплексного оцінювання ризиків, що виникають на різних стадіях життєвого циклу ІІІ-моделей геопросторового сприйняття для інформаційних підсистем роїв БпЛА.

Предметом дослідження є метод комплексного оцінювання цих ризиків з урахуванням можливих системних наслідків поширення локальних помилок сприйняття, а також моделі, критерії, показники та процедури, які забезпечують його реалізацію і перевірку результатів оцінювання.

В дисертаційній роботі автор послідовно розв'язує поставлені завдання:

- у першому розділі проаналізовано стандарти управління ризиками, документи щодо ризиків ІІІ, підходи до безпечності автономних систем і дослідження машинного навчання для роїв БпЛА. На цій основі визначено вимоги до методу;

- у другому розділі розроблено структуру методу. Життєвий цикл подано через стадії селекції знань, аналізу знань, навчання та експлуатації моделі. Для кожної стадії визначено людські й технологічні фактори. Сформовано ядро таксономії для керованого пріоритизування та розширений реєстр для детального аналізу. Це є обґрунтованим компромісом між повнотою опису і складністю попарних порівнянь.

- третій розділ присвячено оцінювання відносної значущості ризиків та проведена експериментальна і симуляційна перевірка їх проявів. Експериментальна частина містить два блоки. У першому на наборі з 30 аерофотозображень досліджено помилки п'яти мультимодальних ІІІ-систем загального призначення. У другому за допомогою графової моделі оцінено поширення локальної помилки між агентами. Результати не використовуються для прямого числового підтвердження ваг МАІ, а формують частковий спостережуваний профіль. Це методологічно правильне розмежування різних шкал;

- у четвертому розділі наведено практичні рекомендації щодо застосування методу комплексного оцінювання ризиків.

Обґрунтованість результатів також підтримується відкритими репозиторіями з конфігураціями, програмною реалізацією, первинними матеріалами та звітами. Водночас частина результатів має характер апробації процедури, а не завершеної експлуатаційної валідації. Автор прямо вказує ці межі.

Наукова новизна одержаних результатів

Наукова новизна дисертації полягає в таких трьох положеннях:

1. Розроблено метод комплексного оцінювання ризиків, що виникають на різних стадіях життєвого циклу ШІ-моделей геопросторового сприйняття для інформаційних підсистем роїв БпЛА. Відмінність методу полягає в поєднанні структурування і класифікації ризиків за стадіями життєвого циклу, кількісного пріоритизування, аналізу стійкості результатів оцінювання та врахування можливих системних наслідків локальних помилок сприйняття.

2. Розроблено таксономію ризиків, відмінністю якої є пов'язування ризиків зі стадіями селекції знань, аналізу знань, навчання й експлуатації моделі та їх розподіл за людськими і технологічними факторами.

3. Удосконалено процедуру кількісного пріоритизування ризиків на основі МАІ шляхом використання кількох профілів змодельованих експертних суджень і моделювання Монте-Карло з контролем узгодженості матриць попарних порівнянь.

У формулюванні дисертації цей третій результат позначено як удосконалення процедури. Експериментальні показники, очікуваний і спостережуваний профілі та результати графової симуляції є складовими реалізації й перевірки методу, а не окремими пунктами наукової новизни.

Практичне значення одержаних результатів

Практичне значення полягає в можливості застосування розробленого методу для відтворюваного оцінювання ризиків, визначення їх відносної пріоритетності та планування перевірок ШІ-моделей геопросторового сприйняття.

До конкретних прикладних результатів належать ядро таксономії з 20 ризиків для первинного оцінювання і розширена таксономія з 71 ризику для ведення деталізованого реєстру. Для ризиків наведено механізми виникнення, індикатори раннього попередження, заходи зниження та

способи перевірки. Практичне значення мають також сценарні процедури МАІ, експериментальні й симуляційні протоколи та рекомендації щодо тестування, валідації, моніторингу, контролю невизначеності й обмеження каскадного поширення помилок.

Результати можуть використовуватися для визначення пріоритетних напрямів спрямування ресурсів тестування й валідації. Водночас дисертація не містить завершеної оптимізаційної моделі розподілу цих ресурсів і не подає кількісної оцінки ефективності заходів після впровадження. Тому практичне значення полягає у підтримці прийняття рішень, а не в автоматичному формуванні плану випробувань.

Оцінка змісту та структури дисертації, її завершеності й дотримання принципів академічної доброчесності

Загальний обсяг дисертації становить 185 сторінок, основний текст – 112 сторінок. Робота складається зі вступу, чотирьох розділів, висновків, списку з 93 джерел і чотирьох додатків. Вона містить 9 рисунків та 17 таблиць.

У вступі наведено всі основні характеристики дослідження. Мета пов'язана з визначенням пріоритетних ризиків і напрямів їх зниження. Вісім завдань охоплюють аналітичний, методичний, експериментальний, симуляційний та практичний компоненти.

Перший розділ формує предметне підґрунтя роботи. Розглянуто архітектуру ройових систем, функції геопросторового сприйняття, обмеження моделей ШІ, стандарти й підходи до управління ризиками. Визначено прогалину, яку автор пов'язує з відсутністю єдиного методу, що поєднує життєвий цикл, пріоритизування та системні наслідки локальної помилки.

Другий розділ є основним методичним розділом. У ньому побудовано ієрархію «стадія – фактор – ризик», визначено глобальні ваги, правила контролю узгодженості та профілі P1–P4. Окремо описано обмеження змодельованих суджень і логіку зіставлення очікуваного та спостережуваного профілів.

Третій розділ містить результати застосування процедури. Для ядра таксономії визначено два ризики зі стійкою міжпрофільною пріоритетністю. Експеримент А виявив типові помилки просторово-геометричного аналізу, а експеримент В показав можливість поширення локальної хибної оцінки. Додаткове оцінювання 71 ризику підтвердило збереження основних

висновків на рівні взаємопов'язаних груп, хоча не зберегло спільної п'ятірки найпріоритетніших ризиків на рівні окремих пунктів.

Четвертий розділ переводить результати у практичні рекомендації. Вони охоплюють організацію оцінювання, контроль невизначеності й конфабуляцій, перевірку просторових висновків і обмеження каскадного поширення помилок.

Робота є логічно завершеною. Загальні висновки відповідають поставленим завданням і не виходять за межі основних результатів.

У дисертації визначено особистий внесок здобувача в розроблення методу, таксономії, програмного забезпечення, експериментів і симуляцій. Для праць зі співавторами наведено розподіл внесків. Джерела оформлено в бібліографічному списку, а посилання подано в тексті. Позитивно слід оцінити окреме повідомлення про використання ChatGPT для пошуку, первинного опрацювання літератури та мовної перевірки. Автор зазначає, що перевіряв і відредагував відповідні матеріали та несе відповідальність за науковий зміст.

Повнота викладення основних результатів у наукових публікаціях

Основні результати представлено у чотирьох наукових статтях: двох публікаціях у фахових виданнях України категорії «Б» та двох статтях у виданнях, що індексуються Scopus. Одне з цих видань належить до першого квартиля за SCImago Journal and Country Rank. Матеріали дисертації також апробовано у чотирьох публікаціях за матеріалами конференцій.

Тематика праць відповідає змісту дисертації. Публікації висвітлюють ройові системи, застосування ІІІ, процедуру оцінювання ризиків і її використання для геопросторового сприйняття під час моніторингу критичної енергетичної інфраструктури.

Зауваження та дискусійні положення

1. Змодельовані профілі, використані в дисертації, стали основою для перевірки працездатності процедури та її чутливості до різних фахових перспектив. Формування додаткових профілів суджень із залученням фахівців із машинного навчання, мережної взаємодії, експлуатації рою та безпеки дозволило б доповнити проведену перевірку практичними експертними оцінками.

2. Одержані результати пріоритизування ризиків створюють змістовну основу для визначення черговості подальших перевірок. Ще одним

можливим кроком для розширення практичного застосування методу могло б стати поєднання ваг ризиків із вартістю і тривалістю перевірки, доступністю даних та очікуваною інформативністю випробування. Це дозволило б формувати план перевірок за умов обмеженого бюджету або часу.

3. У роботі запропоновано ієрархічне групування ризиків, що дає змогу зменшити розмір окремих матриць попарних порівнянь. Водночас практична апробація методу могла б бути ще змістовнішою, якби поряд із результатами оцінювання було визначено орієнтовний час роботи експертів, загальну кількість порівнянь, частку повторних оцінок і можливий вплив втоми учасників.

4. Вагомою практично-орієнтованою складовою четвертого розділу є рекомендації щодо зниження визначених ризиків. Їх наочність додатково підсилило б порівняння значень визначених у роботі показників до та після застосування одного із заходів зниження відповідного ризику, наприклад, обмеження впливу центрального агента.

5. У дисертації для експериментальної та симуляційної перевірки обґрунтовано обрано два пріоритетні ризики зі стадій аналізу знань та експлуатації. Хоча для демонстрації застосування запропонованого методу цього цілком достатньо, його додаткова апробація на прикладі одного з ризиків, що виникають на стадіях селекції знань або навчання моделі, дала б змогу ще повніше представити можливості методу на різних етапах життєвого циклу ІІІ-моделі.

Наведені зауваження не зменшують наукову та практичну цінність дисертаційної роботи і стосуються переважно напрямків подальшого розвитку дослідження.

Загальний висновок

Дисертація Кравчука Олександра Анатолійовича «Метод оцінювання ризиків життєвого циклу моделей штучного інтелекту геопросторового сприйняття для інформаційних підсистем роїв безпілотних літальних апаратів» є завершеною кваліфікаційною науковою працею. У ній розв'язано актуальне наукове завдання. Наукові положення, висновки і рекомендації мають належне обґрунтування; дисертація містить заявлені положення наукової новизни та результати, що мають практичне значення.

За актуальністю, змістом, рівнем наукових результатів, повнотою публікацій і дотриманням принципів академічної доброчесності робота відповідає вимогам Порядку присудження ступеня доктора філософії та

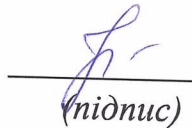
скасування рішення разової спеціалізованої вченої ради закладу вищої освіти, наукової установи про присудження ступеня доктора філософії, затвердженого постановою Кабінету Міністрів України від 12 січня 2022 р. № 44.

Кравчук Олександр Анатолійович заслуговує на присудження ступеня доктора філософії в галузі знань 12 – Інформаційні технології за спеціальністю 122 – Комп'ютерні науки.

Офіційний опонент:

професор кафедри інженерії програмного забезпечення в енергетиці
навчально-наукового інституту атомної та теплової енергетики
Національного технічного університету України
«Київський політехнічний інститут імені Ігоря Сікорського»

доктор технічних наук, професор

 Наталія ФЕДОРОВА
(підпис)

Підпис Федорової Н.В. засвідчує:

Генеральний секретар

*Національного технічного університету України
«Київський політехнічний інститут імені Ігоря Сікорського»*



Валерій Халевко