УДК 004.056:629.3.083
DOI https://doi.org/10.32782/2663-5941/2026.1.2/41

Скоп А.С.
https://orcid.org/0009-0005-6715-4906
Державний університет інформаційно-комунікаційних технологій

Востріков С.О.
https://orcid.org/0009-0008-8425-8872
Державний університет інформаційно-комунікаційних технологій

АНАЛІЗ ЗАГРОЗ DATA POISONING У ВІЗУАЛЬНО-ІНЕРЦІАЛЬНИХ СИСТЕМАХ

Швидке розгортання сервісів самокерованих таксі посилює залежність транспортної інфраструктури від точності сенсорних даних, що формують основу візуально-інерціальної навігації. Будь-які спотворення цих даних, штучно внесені під час збору, передавання чи донавчання моделей, здатні впливати на рішення системи, провокувати відхилення від маршруту або спричиняти втрату локалізації. Атаки типу data poisoning стають усе більш реалістичним інструментом впливу через можливість маніпулювати елементами міського середовища, інерціальними сигналами або навчальними вибірками, не маючи фізичного доступу до транспортного засобу.

Метою дослідження є комплексний аналіз загроз data poisoning у візуально-інерціальних системах, визначення ключових векторів атак для сервісів самокерованих таксі й оцінка того, який рівень контролю над поведінкою транспортного засобу може отримати зловмисник. У роботі застосовано аналіз процесів порушення сенсорної інтеграції, моделювання можливих сценаріїв атак на камери та інерціальний вимірювальний блок, а також порівняння підходів до виявлення спотворених даних.

Отримані результати демонструють, що візуально-інерціальні системи залишаються вразливими до навіть незначних спотворень у потоках даних. Модифіковані об'єкти міського середовища здатні спричиняти хибні зміщення в оцінці руху, а похибки інерціального сигналу провокують накопичувальний дрейф, малопомітний на стартових етапах. Встановлено, що комбіновані атаки призводять до найбільш тривалих і складних для виявлення порушень, оскільки створюють ілюзію внутрішньої узгодженості між сенсорними каналами. На основі цих спостережень запропоновано сукупність організаційних і технічних заходів: контроль походження даних, перевірка синхронізації сенсорів, очищення вибірок перед донавчанням і сценарне тестування моделей у складних умовах.

Дослідження підтверджує, що загрози data poisoning становлять реальну небезпеку для сервісів самокерованого таксі, адже дають змогу впливати на поведінку транспортного засобу без фізичного контакту з його обладнанням. Сформовані в роботі висновки створюють підґрунтя для розроблення багаторівневих механізмів захисту, які можуть бути інтегровані в системи моніторингу даних, навігаційні модулі та інструменти виявлення аномалій у реальному часі. Перспективи подальших досліджень пов'язані з моделюванням багатокомпонентних атак, адаптивною оцінкою надійності сенсорних потоків і створенням симуляційних середовищ для тестування поведінки автономних транспортних засобів у складних міських сценаріях.

Ключові слова: *сенсорний ф'южн, загрози даних, автономне таксі, навмисні спотворення, стійкість моделей, візуально-інерціальні системи.*

Постановка проблеми. Швидке поширення самокерованих таксі створює ситуацію, в якій автономні транспортні системи опиняються в зоні підвищеного ризику, адже їхня робота безпосередньо залежить від цілісності даних, що надходять із візуально-інерціальних сенсорів. Порушення цих даних шляхом *data poisoning* здатне призвести до хибної роботи навігаційних алгоритмів, втрати локалізації або формування помилкових карт оточення. Навіть невелике викривлення вхідних даних може призвести до некоректної реакції інтелектуальної системи та створення небезпеки для користувачів транспортної послуги [1, с. 9–11]. Уразливість до таких атак є особливо

суттєвою для компаній, що працюють у сфері надання послуг самокерованого таксі, оскільки такі системи функціонують у складних міських середовищах із високою щільністю об'єктів.

Природа *data poisoning* полягає в цілеспрямованому внесенні зловмисником спотворених даних до наборів, що використовуються моделями глибинного навчання, або до сенсорних потоків, які надходять у режимі реального часу. Уразливими є навіть сучасні інтелектуальні транспортні системи, оскільки високоточні моделі дедалі частіше спираються на багатокомпонентні джерела інформації, зокрема сигнали інерціального вимірювального блока (Inertial Measurement Unit, IMU), який фіксує прискорення та кутові швидкості руху, а також візуальні ознаки [2]. На практиці це може проявлятися як підміна маркерів орієнтирів, спотворення текстур на статичних об'єктах або модифікація сегментів дорожньої розмітки, які автономний автомобіль використовує для корекції руху. Подібні втручання вже демонстрували здатність змушувати моделі комп'ютерного зору неправильно інтерпретувати дорожні знаки та змінювати траєкторії руху, що підтверджують експериментальні дані про якісні збої поведінкових моделей у разі навмисного викривлення вхідної інформації [3, с. 99–101].

Для умовної компанії, що експлуатує парк самокерованих таксі, ключовими є вектори атаки, які дають змогу зловмиснику впливати на роботу візуально-інерціальної системи з різним рівнем контролю над результатом. Одним із прикладів є розміщення змінених або контрастно підсилених зображень на білбордах чи стендах біля дороги, адже такі цілеспрямовані сигнали можуть викликати помилки SLAM-алгоритмів і локальне зрушення системи навігації, що описано у праці С. Ванга (S. Wang) та ін. щодо маніпуляцій, здатних призвести до деградації автономних моделей [4]. Іншим вектором є вплив на IMU-дані через електромагнітні випромінювання або вібраційні стимули, які здатні порушувати синхронізацію сенсорного ф'южну (sensor fusion). Зловмисник, залежно від ресурсів, може мати на меті як тимчасову зупинку транспортного засобу, так і спрямоване відхилення маршруту. У літературі виокремлено і складні сценарії, за яких атака створює тривалі зміни в роботі моделі, впливаючи на навчальні дані та формуючи хибні шаблони поведінки [5, с. 210–212].

Загроза *data poisoning* посилюється комплексністю візуально-інерціальних систем і відсутністю повної прозорості в процесах їхнього навчання, що ускладнює виявлення маніпуляцій. Системи, що працюють із потоковими даними, мають вразливість до точкових втручань, які здатні спричинити каскадне порушення логіки прийняття рішень. Для компанії, що надає послуги автономного таксі, це означає формування критичного ризику. Навіть одноразова атака з незначним рівнем контролю може призвести до локальної втрати навігації, а більш складні втручання – до навмисного відхилення траєкторії, зупинки автомобіля або створення небезпечних ситуацій у транспортному потоці. Саме тому аналіз загроз *data poisoning* у візуально-інерціальних системах набуває особливої актуальності, оскільки визначає основу для формування ефективних механізмів протидії і стійкої роботи автономних транспортних послуг.

Аналіз останніх досліджень і публікацій. Проблематика *data poisoning* у системах штучного інтелекту (ШІ) загалом представлена досить широко, проте аспекти, пов'язані саме з візуально-інерціальними комплексами самокерованих транспортних систем, досліджено неповно. На загальному рівні уразливості ШІ описують О. Неретін та В. Харченко, які підкреслюють ризики викривлення даних і демонструють, як це впливає на поведінку складних алгоритмів [1]. Також значний внесок у розуміння системних ризиків роблять Г. Гайдур, С. Гахов та О. Скибун, які аналізують стан кібербезпеки критичної інфраструктури з використанням ШІ і показують, що порушення цілісності потокових даних може спричиняти каскадні відмови в складних системах [6].

У технічному напрямі важливими є дослідження стійкості систем розпізнавання образів до спрямованих втручань. Так, О. Омельченко та А. Шелестов показують, що навіть невеликі зміни візуальних елементів можуть спричиняти небажані зміщення в результатах моделей [3]. У свою чергу, О. Ящик та ін. звертають увагу на недостатність механізмів контролю цілісності сенсорних потоків і демонструють, що наявні моделі вразливі до локальних втручань у дані [7]. Проте питання інтегрованого впливу на камеру та інерціальний вимірювальний блок (IMU) залишається розкритим частково. О. Скіцько та ін. систематизували загальні загрози застосування ШІ в різних сферах і підкреслили, що моделі є чутливими до невидимих змін у вхідних даних [8]. Хоча їх дослідження не орієнтоване на автономний транспорт, автори слушно зауважують, що джерелом ризику часто стає не алгоритм, а маніпуляція даними, яка робить атаки *data poisoning* особливо небез-

печними. Додатково у методологічному аспекті важливими є результати дослідження К. Завражного та А. Кулик, які обґрунтовують необхідність оцінки впливу ШІ на інформаційну безпеку через призму стійкості даних [9].

Міжнародний пласт літератури фокусується насамперед на атаках у транспортних інтелектуальних системах. У статті С. Ванга (S. Wang) та ін. запропоновано адаптивні схеми нападів, що здатні поступово знижувати точність автономних моделей і дають зловмиснику контроль над рівнем впливу від незначного збою до втрати навігації [4]. Оглядова стаття Ф. Ванга (F. Wang) та ін. доповнює транспортний контекст, розкриваючи широкий спектр атак на інтелектуальні транспортні системи, зокрема отруєння робочих сенсорних потоків [10]. Автори підтверджують, що такі втручання можуть призводити до помилкової навігації. Натомість Ю. Джянг (Y. Jiang) та ін. підкреслюють накопичувальний характер *data poisoning*, який може проявитися через певний час після атаки [11]. Важливо, що ці дослідження підтверджують реальність сценаріїв, коли зловмисник, використовуючи доступні ресурси, може свідомо впливати на поведінку автономного авто.

Окремі роботи формують контекст ризиків автономного транспорту. Наприклад, К. Гросс (K. Grosse) та А. Алахі (A. Alahi) показують, що навіть незначні маніпуляції з оточенням можуть спричиняти критичні порушення траєкторій руху [12]. Хоча автори не аналізують *data poisoning* як окремий тип атаки, їхні висновки демонструють чутливість автопілотних систем до малопомітних змін входів. На складності захисту сенсорних потоків у повітряному таксі наголошують Б. Тахір (B. Tahir) і М. Тарік (M. Tariq) [13], що підтверджує аналогічні виклики для наземних сервісів. Додатковий фаховий контекст подають Дж. Вілд (G. Wild) та ін., які досліджують мобільні та легкі моделі ШІ в авіаційних застосуваннях [14]. Науковці показують, що сенсорні системи залишаються чутливими до контекстних маніпуляцій навіть у контрольованому середовищі, що опосередковано підтверджує високі ризики для наземних автономних платформ, які працюють у значно більш варіативних умовах.

Попри значну кількість напрацювань, у літературі залишаються наукові прогалини, адже бракує робіт, які комплексно аналізують *data poisoning* саме у візуально-інерціальних системах. Також невивченими залишаються сценарії комбінованих атак на IMU та камери, крім того відсутні моделі, що описують рівень контролю зловмисника під час впливу на реальні об'єкти міському середовища (зокрема білборди як інструмент маніпуляції). Розрив спостерігається і між описом загальних загроз ШІ та реальними технічними ризиками для служб самокерованих таксі. Це підтверджує необхідність спеціалізованого дослідження, спрямованого на оцінку загроз *data poisoning* у візуально-інерціальних системах і визначення механізмів протидії таким атакам.

Постановка завдання. Мета статті полягає в дослідженні загроз *data poisoning* у візуально-інерціальних системах самокерованих таксі, з акцентом на ключових векторах атак, можливостях зловмисника контролювати поведінку транспортного засобу і наслідках спотворення візуальних та інерціальних даних. Досягнення цієї мети передбачає виконання трьох завдань:

1. Дослідити основні вектори атак *data poisoning*, що впливають на камери та IMU самокерованих таксі, і визначити потенційний рівень контролю зловмисника над поведінкою системи.

2. Проаналізувати механізми помилкових рішень сенсорного ф'южну під час спотворення вхідних даних і встановити, які типи викривлень найчастіше призводять до втрати локалізації чи збоїв SLAM-алгоритмів.

3. Виявити практичні технічні й організаційні рекомендації для підвищення стійкості візуально-інерціальних систем до атак *data poisoning* у сервісах автономного таксі.

Виклад основного матеріалу. У сучасних дослідженнях підкреслюється, що загрози *data poisoning* охоплюють маніпуляції як із навчальними наборами, так і з даними під час експлуатації, що робить такі атаки особливо небезпечними для систем із високою залежністю від потокових сенсорних сигналів. Вразливість автономних транспортних сервісів зростає в умовах розподіленої обробки інформації, коли камера, IMU, глобальна супутникова навігаційна система (Global Navigation Satellite System, GNSS) та комунікації «транспортний засіб – усе» (V2X – Vehicle-to-Everything) формують взаємопов'язаний контур ф'южну [1, с. 9–11; 2, с. 210–212]. V2X є технологією обміну даними, яка дозволяє автомобілю напряму спілкуватися з іншими транспортними засобами, дорожньою інфраструктурою, пішоходами та мережевими сервісами для підвищення безпеки і точності навігації.

Цілеспрямовані підміни візуальних ознак (наприклад, зміщення контрасту, наклейки, патерни) можуть спричиняти системне спотворення вихідних ознак, на основі яких працюють

моделі розпізнавання та навігації. Транспортні системи вразливі до комбінованих атак, де одночасно спотворюються дані камер, GNSS або локальних сенсорів, що забезпечує зловмиснику як локальний, так і високорівневий контроль над поведінкою автомобіля.

Data poisoning застосовується для поступової зміни поведінки системи через накопичення помилок у даних або навмисне зрушення статистики сенсорних потоків [15; 16]. Для сервісу самокерованих таксі це створює як ризик непомітних дрібних інцидентів, так і можливість спрямованих атак, призначених для відхилення маршруту. У таблиці 1 узагальнено ключові вектори атак.

Сенсорний ф'южн у самокерованому таксі синтезує дані камер та IMU (інерціальний блок, що вимірює прискорення та кутові швидкості) для отримання точного стану руху. Камери забезпечують корекцію дрейфу, тоді як IMU підтримує стійкість оцінки між візуальними кадрами. Навіть невеликі спотворення у візуальних даних можуть призвести до некоректних детекцій, помилок у виборі смуги та формуванні карти середовища [5, с. 100–103]. Порушення узгодженості між камерами та IMU спричиняє збій у моделі руху, оскільки навігаційний контур може втратити стабільність, а система перейти на аварійні траєкторії або різкі маневри.

У джерелах з аналізу транспортних систем наголошується, що отруєння даних впливає як на класифікацію, так і на проміжні етапи: агрегування інформації, оцінку ризику, структуризацію карти та прогнозування руху [8, с. 10–13]. Окреме помилкове спостереження може спричиняти каскадні відмови, оскільки моделей аномалій не вистачає для ідентифікації атакувальних патернів у потокових даних. Систематизація впливів атак на ф'южн і навігацію наведена на рисунку 1.

Окремі дослідження описують приклади, коли втручання в навчальні або операційні дані призводило до критичних помилок у рішеннях бортових систем. Фіксувалися сценарії, за яких зміна невеликої частини навчальної вибірки спричиняла систематичні збої у детекції об'єктів і виявленні перешкод, що підтверджено в дослідженнях із тестування автономних транспортних моделей у реальних умовах [2; 4]. Додатково відзначено випадки, коли спрямовані атаки на дані сенсорів або журнали подій впливали на працездатність модулів керування та створювали передумови для масштабних відмов, про що докладно повідомляється в оглядах сучасних атак на бізнес-системи та критичну транспортну інфраструктуру [15; 16]. У сукупності ці приклади демонструють, що *data poisoning* не є теоретичною загрозою, а реальним механізмом впливу на навігацію, розпізнавання та безпеку автономного руху.

Підвищення стійкості візуально-інерціальних систем до атак *data poisoning* вимагає комплексної стратегії, яка має поєднувати технічні механізми захисту, організаційні процедури та формалізований контроль за якістю даних. Особливий наголос треба зробити назабезпеченні керованості джерелами даних Це може включати документування їх походження, ведення журналів змін і періодичний аудит навчальних вибірок, які використовуються для оновлення моделей навігації та розпізнавання.

Суттєву роль відіграє технічний моніторинг узгодженості між основними сенсорами візуально-інерціальної системи: камерами, інерціальним вимірювальним блоком, глобальною супутниковою навігаційною системою та комунікаціями V2X. Своєчасне виявлення відхилень у цих потоках дає змогу ідентифікувати як випадкові апаратні збої, так і цілеспрямовані атакувальні дії. Для цього необхідні локальні фільтри

Таблиця 1

Основні вектори атак data poisoning для візуально-інерціальних систем

Група векторів атак	Канали реалізації у сервісі автономних таксі	Можливий рівень контролю над авто
Отруєння навчальних наборів	Ін'єкція модифікованих фреймів, спотворення маркування, зміна розподілу ознак	Від локальних помилок до системної деградації моделі
Отруєння даних камер під час експлуатації	Адверсаріальні білборди, патерни на знаках, зміни кольорів і контрасту	Від дезорієнтації до примусової зупинки чи навмисного маневру
Отруєння IMU та допоміжних сенсорів	Спотворення інерціальних сигналів, маніпуляції магнітним полем, вібраційний вплив	Дрейф положення, втрата локалізації, некоректні траєкторні рішення
Комбіновані багатоканальні атаки	Одночасна підміна відео-, IMU-, GNSS- та V2X-даних	Найвищий: можливість сценарного керування рухом

Джерело: узагальнено авторами на основі даних [1, с. 9–11; 2, с. 210–212; 5, с. 100–103; 11; 13; 15; 16].

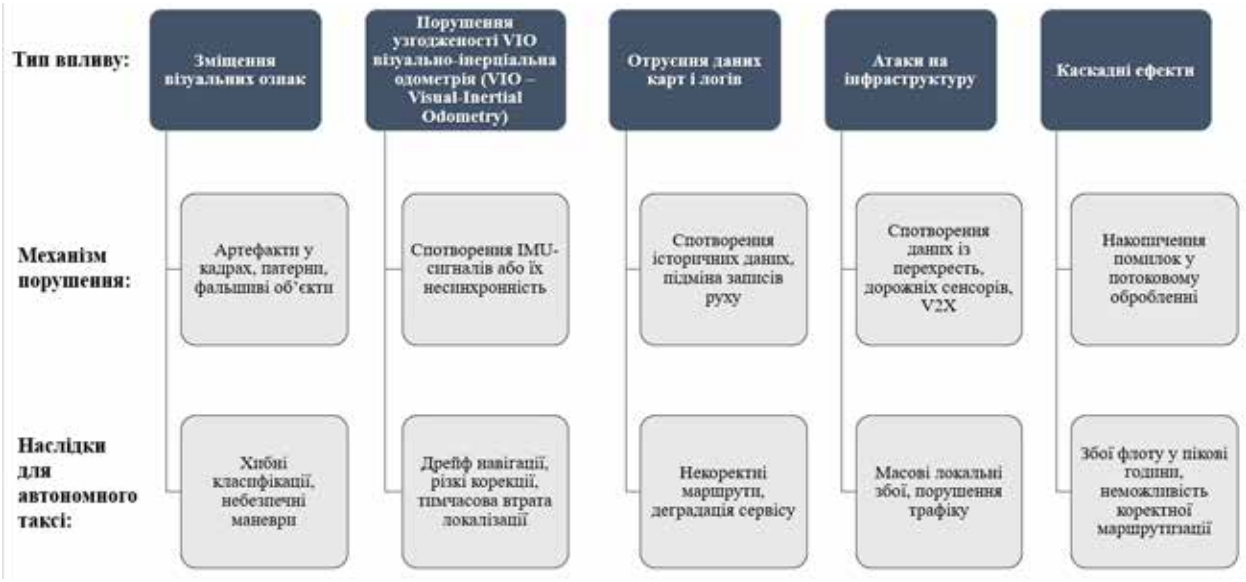

Рис. 1. Типові сценарії впливу атак на ф'южн і навігацію

Джерело: узагальнено авторами на основі даних [2; 5, с. 100–103; 8, с. 10–13; 10; 15; 16].

аномалій, перевірки узгодженості між каналами та крос-перевірка даних між сенсорами.

Технічний рівень захисту має включати процедури стійкого навчання моделей, що зменшують чутливість навігаційних і розпізнавальних модулів до отруєних зразків. До таких підходів належать очищення даних перед донавчанням, фільтрація підозрілих вибірок, а також застосування спеціалізованих методів, розроблених для зниження впливу цілеспрямованих маніпуляцій. Додаткову ефективність забезпечує тестування системи на заздалегідь змодельованих складних сценаріях, де перевіряється поведінка моделі в умовах атакувальних патернів, які у звичайних умовах не виявляються.

У сфері організаційного управління доцільно впроваджувати формалізовані процедури реагування на інциденти, навчання персоналу, який працює з даними, регулярну оцінку ризиків і створення окремих ролей, відповідальних за безпеку алгоритмів. Це дозволяє забезпечити безперервність контролю, підвищити якість рішень і зробити систему менш вразливою до людських помилок. Узагальнення ключових напрямів і сформованих заходів подано в таблиці 2.

Підвищення стійкості візуально-інерціальних систем до атак *data poisoning* потребує одночасного впровадження технічних, процедурних та організаційних заходів. Зокрема, ефективний захист неможливий без налагодженого управління життєвим циклом даних, яке забезпечує контроль їх походження та зменшує ризик непомітної ін'єкції шкідливих зразків. Додатковий рівень безпеки створюють перевірки узгодженості між ключовими сенсорами (камерами, інерціальним вимірювальним блоком, глобальною супутниковою навігаційною системою та V2X-комунікаціями), що дозволяє своєчасно виявляти спотворення або підміни в потоках даних.

Застосування стійких методів навчання й очищення даних підсилює здатність моделей протистояти цілеспрямованим маніпуляціям, а сценарне тестування дає змогу виявляти приховані вразливості в складних ситуаціях, які не покриваються стандартними перевірками. Завершальним елементом є організаційні механізми (формалізовані процедури реагування, навчання персоналу та створення спеціалізованих ролей із безпеки ШІ), що забезпечують безперервність контролю та знижують ризики людського фактору. Сукупність цих напрямів формує багаторівневу систему захисту, здатну значно зменшити ефективність атак *data poisoning* у сервісах самокерованих таксі.

Висновки. Стрімке впровадження сервісів самокерованих таксі формує новий рівень залежності транспортних систем від якості сенсорних даних і стійкості моделей, що поєднують візуальні та інерціальні сигнали. Будь-яке навмисне спотворення цих даних здатне змінити поведінку автономного транспортного засобу значно сильніше, ніж традиційні атаки на інформаційні системи. Саме тому аналіз загроз *data poisoning* у візуально-інерціальних комплексах є критично важливим для розвитку безпечних міських транспортних сервісів.

У межах дослідження встановлено, що маніпуляції даними, спрямовані на камери та інерці-

Таблиця 2

Заходи для підвищення стійкості візуально-інерціальних систем до атак *data poisoning*

Напрям захисту	Конкретні заходи для сервісу самокерованих таксі	Очікуваний ефект та можливі обмеження
Управління життєвим циклом даних	Введення політик походження даних, журналювання змін і контроль доступу до навчальних вибірок, регулярний аудит датасетів, періодична перевірка оновлень моделей	Зменшення ризику непомітного отруєння даних; підвищення прозорості процесів. Обмеження: потреба в додаткових ресурсах і зрілості інфраструктури
Захист сенсорної інтеграції (камери + IMU + GNSS + V2X)	Перевірки узгодженості між камерами та інерціальним вимірювальним блоком, крос-перевірка з GPS, крос-перевірка з V2X-даними, локальні фільтри аномалій у потоках сенсорних даних	Зниження ймовірності раптової втрати локалізації, вчасне виявлення маніпуляцій. Обмеження: підвищення обчислювального навантаження
Стійкі методи навчання та фільтрації даних	Використання robust training підходів, очищення даних перед донавчанням (data sanitization), перевірка нових вибірок перед інтеграцією в моделі навігації і розпізнавання	Зменшення чутливості моделей до отруєних зразків. Обмеження: можливе уповільнення циклу оновлень моделей
Сценарний аналіз ризиків і тестування	Тестування на спеціально змодельованих атакувальних сценаріях, моделювання наслідків для групи транспортних засобів у реальних дорожніх умовах, оцінка меж працездатності моделей у складних середовищах	Виявлення слабких місць, недоступних у стандартних тестах; підвищення рівня готовності до інцидентів. Обмеження: значні витрати часу та потреба у фаховій команді
Організаційні процедури й управління безпекою ШІ	Формалізація процедур реагування на інциденти, навчання персоналу, який працює з даними та моделями, впровадження спеціалізованих ролей або підрозділів, відповідальних за безпеку алгоритмів	Підвищення культури безпеки, зменшення ризику людських помилок. Обмеження: потреба в підтримці керівництва та виділенні ресурсів

Джерело: власна розробка авторів.

альний вимірювальний блок, можуть надавати зловмиснику широкий діапазон можливостей від локального викривлення навігації до керованої зміни траєкторії руху. Такі впливи виникають як під час навмисного спотворення візуальних елементів міського середовища, так і під час модифікації інерціальних сигналів, що підтверджує високий рівень чутливості класичних SLAM- та VIO-рішень до атакувальних втручань. Дослідження механізмів сенсорної інтеграції показало, що помилкові рішення формується переважно через порушення узгодженості між каналами, адже зрушення візуальних ознак, дрейф IMU, несинхронність сигналів і накопичення відхилень у потоках даних здатні провокувати втрату локалізації, аварійні маневри або спонтанні зупинки. Отримані результати дають можливість застосовувати розроблені підходи в транспортних платформах, що працюють із потоковими даними та використовують комбіновані сенсорні модулі. Зокрема, запропоновані заходи можуть бути інтегровані в системи маршрутизації, модулі локалізації і блоки розпізнавання об'єктів, забезпечуючи стійкість сервісів самокерованих таксі до зловмисних маніпуляцій. Їх практичне використання дозволяє зменшити ризики неконтрольованої поведінки, запобігти накопичуваним несправностям і підвищити рівень безпеки користувачів.

Наукова новизна роботи полягає в структуризації загроз *data poisoning* саме для візуально-інерціальних систем самокерованих таксі й у формуванні комплексної моделі їх впливу на поведінку транспортного засобу з урахуванням реальних умов міського середовища. Практична цінність полягає в можливості безпосереднього застосування запропонованих підходів під час проєктування інфраструктури безпеки, створення механізмів контролю якості сенсорних даних і підвищення стійкості навігаційних моделей у реальних сервісах автономного транспорту.

Перспективи подальших досліджень пов'язані з необхідністю моделювання складних комбінованих атак, що одночасно впливають на камери, IMU, GNSS та V2X-комунікації, а також із розробленням нових підходів до виявлення маніпуляцій у реальному часі. Доцільним є створення адаптивних моделей, здатних самостійно оцінювати надійність даних, і формування симуляційних полігонів, де можна відтворювати реалістичні міські сценарії та тестувати реакції автономного транспорту на атакувальні патерни.

Список літератури:

1. Неретін О., Харченко В. Забезпечення кібербезпеки систем штучного інтелекту: аналіз вразливостей, атак і контрзаходів. *Вісник Національного університету «Львівська політехніка». Інформаційні системи та мережі*. 2022. № 12. С. 7–22. DOI: https://doi.org/10.23939/sisn2022.12.007 (дата звернення: 28.11.2025).

2. Wang F., Wang X., Ban X. Data poisoning attacks in intelligent transportation systems: A survey. *Transportation Research Part C: Emerging Technologies*. 2024. Vol. 165. Article 104750. DOI: https://doi.org/10.1016/j.trc.2024.104750 (дата звернення: 28.11.2025).

3. Омельченко О., Шелестов А. Оцінка стійкості систем розпізнавання образів до впливу зловмисних втручань. *Проблеми керування та інформатики*. 2025. Т. 70. № 4. С. 96–108. DOI: https://doi.org/10.34229/1028-0979-2025-4-6 (дата звернення: 28.11.2025).

4. Wang S., Li Q., Cui Z., Hou J., Huang C. Bandit-based data poisoning attack against federated learning for autonomous driving models. *Expert Systems with Applications*. 2023. Vol. 227. Article 120295. DOI: https://doi.org/10.1016/j.eswa.2023.120295 (дата звернення: 28.11.2025).

5. Марценюк Є., Партика А., Крет Т. Дослідження вразливостей штучного інтелекту та побудова комплексної моделі безпеки організації. *Сучасний захист інформації*. 2025. № 1 (61). С. 206–218. DOI: https://doi.org/10.31673/2409-7292.2025.018929 (дата звернення: 28.11.2025).

6. Гайдур Г., Гахов С., Скибун О. Оцінка стану кібербезпеки критичної інфраструктури з використанням ШІ. *Сучасний захист інформації*. 2025. № 2 (62). С. 31–41. DOI: https://doi.org/10.31673/2409-7292.2025.020831 (дата звернення: 28.11.2025).

7. Ящик О., Симонов В., Іваненко Р. Забезпечення кібербезпеки в еру штучного інтелекту: аналіз технологічних підходів та стратегій для захисту інформації. *Бізнес Інформ*. 2024. № 1. С. 81–86. DOI: https://doi.org/10.32983/2222-4459-2024-1-81-86 (дата звернення: 28.11.2025).

8. Скіцько О., Складанний П., Ширшов Р., Гуменюк М., Ворохоб М. Загрози та ризики використання штучного інтелекту. *Кібербезпека: освіта, наука, техніка*. 2023. Т. 2. № 22. С. 6–18. DOI: https://doi.org/10.28925/2663-4023.2023.22.618 (дата звернення: 28.11.2025).

9. Zavrazhnyi K., Kulyk A. Methodological principles of assessing the impact of artificial intelligence on the information security of management systems of enterprises. *Kyiv Economic Scientific Journal*. 2024. No. 7. P. 71–78. DOI: https://doi.org/10.32782/2786-765X/2024-7-10 (дата звернення: 28.11.2025).

10. Wang F., Wang X., Ban X. Data poisoning attacks in intelligent transportation systems: a survey. *arXiv*. 2024. DOI: https://doi.org/10.48550/arXiv.2407.15855 (дата звернення: 28.11.2025).

11. Data poisoning in deep learning: a survey / Y. Jiang et al. *arXiv*. 2025. DOI: https://doi.org/10.48550/arXiv.2503.22759 (дата звернення: 28.11.2025).

12. Grosse K., Alahi A. A qualitative AI security risk assessment of autonomous vehicles. *Transportation Research Part C: Emerging Technologies*. 2024. Vol. 169. Article 104797. DOI: https://doi.org/10.1016/j.trc.2024.104797 (дата звернення: 28.11.2025).

13. Tahir B., Tariq M. Vulnerability assessment and federated intrusion detection of Air Taxi-enabled smart cities. *Sustainable Energy Technologies and Assessments*. 2022. Vol. 53. Article 102686. DOI: https://doi.org/10.1016/j.seta.2022.102686 (дата звернення: 28.11.2025).

14. Lightweight and mobile artificial intelligence and immersive technologies in aviation / G. Wild et al. *Visual Computing for Industry, Biomedicine, and Art*. 2025. Vol. 8. Article 21. DOI: https://doi.org/10.1186/s42492-025-00203-z (дата звернення: 28.11.2025).

15. Obadiaru A. Data poisoning attacks: a new attack vector within AI. *Cobalt*. URL: https://www.cobalt.io/blog/data-poisoning-attacks-a-new-attack-vector-within-ai (дата звернення: 28.11.2025).

16. Robb B. Data poisoning attacks: how hackers target AI-Driven business systems. *BlackFog*. URL: https://www.blackfog.com/data-poisoning-attacks-hackers-target-ai-systems/ (дата звернення: 28.11.2025).

Skop A.S., Vostrikov S.O. ANALYSIS OF DATA POISONING THREATS IN VISUAL-INERTIAL SYSTEMS

The rapid deployment of self-driving taxi services increases the reliance of transport infrastructure on the accuracy of sensor data, which underpins visual-inertial navigation. Any distortions of this data, artificially introduced during collection, transmission or retraining of models, can affect the system's decisions, provoke deviations from the route or cause loss of localization. Data poisoning attacks are becoming an increasingly realistic tool of influence due to the ability to manipulate elements of the urban environment, inertial signals or training samples without having physical access to the vehicle.

The purpose of the study is to conduct a comprehensive analysis of data poisoning threats in visual-inertial systems, identify key attack vectors for self-driving taxi services, and assess the level of control over the vehicle's behavior that an attacker can obtain. The paper uses the analysis of sensory integration disruption

processes, modeling of possible attack scenarios on cameras and inertial measurement units, and comparison of approaches to detecting distorted data.

The results obtained demonstrate that visual-inertial systems remain vulnerable to even minor distortions in data streams. Modified objects in the urban environment can introduce spurious biases in motion estimation, and inertial signal errors can cause cumulative drift, which is barely noticeable in the initial stages. It has been established that combined attacks lead to the most long-lasting and difficult-to-detect violations, since they create the illusion of internal consistency between sensory channels. Based on these observations, a set of organizational and technical measures is proposed: data origin control, sensor synchronization verification, sample cleaning before pre-training, and scenario testing of models under challenging conditions.

The study confirms that data poisoning threats pose a real danger to self-driving taxi services, as they allow the manipulation of a vehicle's behavior without physical access to its equipment. The conclusions drawn in the work form the basis for the development of multi-level protection mechanisms that can be integrated into data monitoring systems, navigation modules, and real-time anomaly detection tools. Prospects for further research include modeling multi-component attacks, adaptive assessment of the reliability of sensor streams, and creating simulation environments for testing the behavior of autonomous vehicles in complex urban scenarios.

Keywords: *sensor fusion, data threats, autonomous taxi, intentional distortions, model stability, visual-inertial systems.*

Дата першого надходження статті до видання: 29.01.2026
Дата прийняття статті до друку після рецензування: 23.02.2026
Дата публікації (оприлюднення) статті: 08.04.2026