МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДОНЕЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ВАСИЛЯ СТУСА
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ І ПРИКЛАДНИХ ТЕХНОЛОГІЙ
КАФЕДРА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

**ЮРЧАК ВЛАДИСЛАВ ОЛЕКСАНДРОВИЧ**

Допускається до захисту:
в.о. завідувача кафедри
інформаційних технологій,
д-р. техн. наук, професор
_____Наталія ВЕСЕЛОВСЬКА
«____»_____ 2025 р.

# ВИЯВЛЕННЯ ТІНЬОВОГО ФЛОТУ КРАЇНИ-АГРЕСОРА НА ОСНОВІ АНАЛІЗУ ГЕОЛОКАЦІЙНИХ ДАНИХ СУДЕН

Спеціальність 122 Комп’ютерні науки

**Кваліфікаційна (магістерська) робота**

Науковий керівник:
Роман БАБАКОВ, професор кафедри
інформаційних технологій,
д. т. н., доцент

_____

Оцінка: _____ / ___ / _____
(бали за шкалою ЄКТС / за національною шкалою)

Голова ЕК: _____

Вінниця 2025

# ЗМІСТ

# АНОТАЦІЯ

**Юрчак В. О.** Виявлення тіньового флоту країни агресора на основі аналізу геолокаційних даних суден. Спеціальність 122 Комп'ютерні науки. Освітня програма "Комп'ютерні технології обробки даних". Донецький національний університет імені Василя Стуса, Вінниця, 2025.

Кваліфікаційна робота присвячена дослідженню методів виявлення тіньового флоту на основі аналізу геолокаційних та поведінкових даних автоматичної ідентифікаційної системи (AIS) морських суден. У роботі розглянуто сучасні схеми обходу санкцій, характерні для "тіньового" флоту, включно з вимкненням AIS-трансляції, підміною ідентифікаторів суден, зміною прапорів, маніпуляціями траєкторіями та проведенням прихованих STS-операцій (ship-to-ship). Здійснено збір, очищення та обробку великомасштабних даних руху суден з використанням технологій Apache Spark, геопросторових бібліотек та власних алгоритмів детекції аномалій.

У роботі запропоновано методику автоматичного виявлення підозрілих суден шляхом аналізу часових розривів у передачі AIS-сигналів, виявлення перетину міжнародної лінії зміни дат, аналізу нетипових маршрутів, порушень швидкісних характеристик та просторових аномалій. Розроблено програмний прототип системи, здатний будувати траєкторії суден, визначати небезпечні "тіні" в даних, автоматично класифікувати поведінкові патерни та формувати аналітичні звіти для подальшого моніторингу.

Практична цінність роботи полягає у можливості застосування запропонованого підходу для посилення контролю за морськими перевезеннями, підтримки санкційної політики, виявлення незаконних операцій з переміщення нафти та інших вантажів, а також побудови систем раннього попередження щодо аномальної поведінки суден.

## ABSTRACT

**Yurchak V. O.** Detection of the shadow fleet of the aggressor country based on the analysis of geolocation data of ships. Specialty 122 Computer Science. Educational program “Computer Science”. Vasyl’ Stus Donetsk National University, Vinnytsia, 2025.

The qualification work is devoted to the study of methods for detecting the shadow fleet based on the analysis of geolocation and behavioral data of the automatic identification system (AIS) of sea vessels. The work considers modern schemes for circumventing sanctions typical of the “shadow” fleet, including disabling AIS broadcasts, replacing ship identifiers, changing flags, manipulating trajectories and conducting covert STS (ship-to-ship) operations. Large-scale vessel movement data was collected, cleaned and processed using Apache Spark technologies, geospatial libraries and proprietary anomaly detection algorithms.

The main scientific result of this research is a method for automatic detection of suspicious vessels by analyzing time gaps in AIS signal transmission, detecting crossings of the international date line, analyzing atypical routes, violations of speed characteristics and spatial anomalies. A software prototype of the system has been developed, capable of building vessel trajectories, identifying dangerous “shadows” in the data, automatically classifying behavioral patterns and generating analytical reports for further monitoring.

The practical value of the work lies in the possibility of applying the proposed approach to strengthen control over maritime transportation, supporting sanctions policy, detecting illegal operations in the movement of oil and other cargo, as well as building early warning systems for anomalous vessel behavior.

# ВСТУП

Морські перевезення відіграють ключову роль у світовій торгівлі, забезпечуючи транспортування енергоресурсів, сировини, продовольства та готової продукції між державами та континентами. З огляду на це прозорість та безпека морського судноплавства є критично важливими для стабільності глобальних економічних процесів. Проте в останнє десятиліття у світовому морському транспорті все більшого поширення набуває явище, яке отримало назву “тіньовий флот” — сукупність суден, що навмисно приховують свою діяльність з метою обходу санкцій, транспортних обмежень, податкового контролю або участі у незаконних операціях.

Особливу актуальність ця проблема набула у зв’язку з посиленням міжнародних санкцій щодо транспортування нафти та нафтопродуктів із держав, що перебувають під економічним тиском міжнародної спільноти. З метою уникнення контролю оператори тіньового флоту застосовують широкий спектр технічних та поведінкових тактик: вимкнення або фальсифікація даних AIS, зміна прапорів та назв суден, перереєстрація у юрисдикціях з низьким рівнем контролю, маніпуляції з MMSI та IMO-ідентифікаторами, а також виконання прихованих ship-to-ship (STS) операцій у віддалених регіонах світового океану. Такі дії не лише підривають ефективність санкційної політики, а й становлять серйозну загрозу для екологічної та транспортної безпеки.

З огляду на масштабність морських перевезень та величезні обсяги геолокаційних даних, які щоденно генеруються системами AIS, виявлення тіньового флоту неможливе без сучасних методів аналізу великих даних. Сьогодні аналітичні рішення у сфері морського моніторингу активно використовують технології обробки потокових даних, машинного навчання, виявлення аномалій, графових моделей та геопросторової аналітики. Проте ефективне застосування таких технологій потребує розробки чітких алгоритмів

і критеріїв, які дозволяють відрізнити нормальну поведінку суден від ризикових та потенційно незаконних патернів.

Актуальність даної роботи полягає у необхідності створення ефективних інструментів автоматизованої аналітики для ідентифікації суден тіньового флоту на основі великомасштабних геолокаційних даних. Виявлення тривалих пропусків сигналу, атипових маршрутів, неочікуваних змін швидкості та напрямку руху, а також аномальних поведінкових характеристик є важливим завданням для державних органів, аналітичних центрів, операторів портів, державних служб морської безпеки й компаній, що займаються моніторингом логістичних ланцюгів.

Метою дипломної роботи є створення методики та програмного інструменту для автоматичного виявлення суден тіньового флоту шляхом аналізу геолокаційних даних AIS, часових закономірностей передачі сигналів та просторово-поведінкових аномалій. Для досягнення поставленої мети в роботі передбачається розв’язання таких завдань:

- аналіз сучасного стану проблеми тіньового флоту та методів його виявлення;
- збір, попередня обробка та нормалізація великомасштабних AIS-даних;
- виявлення підозрілих часових розривів у передачі сигналів та геопросторових аномалій;
- дослідження структурних особливостей траєкторій суден і визначення критеріїв аномальної поведінки;
- створення програмного комплексу з візуалізацією маршрутів та аналітичними звітами.

Об’єктом дослідження є процес просторово-часового переміщення морських суден у глобальному масштабі. Предметом дослідження є моделі та методи виявлення аномалій у геолокаційних даних AIS, що дозволяють визначати судна із характеристиками, притаманними тіньовому флоту.

Практичне значення роботи полягає у можливості інтеграції запропонованих алгоритмів у сучасні системи морського моніторингу, аналітичні платформи, системи підтримки санкційної політики та державні інформаційні ресурси. Розроблені підходи можуть допомогти підвищити прозорість морських перевезень, виявляти незаконні операції на ранніх етапах, а також знизити ризики, пов’язані з міжнародними транспортними та енергетичними загрозами.

# РОЗДІЛ 1

# ТЕОРЕТИЧНІ ЗАСАДИ ДОСЛІДЖЕННЯ ГЕОЛОКАЦІЙНИХ ДАНИХ ПЕРЕМІЩЕННЯ СУДЕН

## 1.1. Сутність та значення тіньового флоту транспортування нафтопродуктів

У міжнародній практиці термін “тіньовий флот” використовується для позначення сукупності морських суден, що здійснюють діяльність поза межами прозорих механізмів глобального судноплавства та навмисно приховують свою приналежність, маршрути, вантажі або операції. Тіньовий флот формується переважно з метою обходу міжнародних санкцій, уникнення контролю над торгівлею стратегічними ресурсами, ведення незаконних перевезень та участі у фінансових схемах, пов’язаних з відмиванням коштів чи ухиленням від оподаткування. Сутність тіньового флоту полягає у забезпеченні безперервного транспортування нафти й нафтопродуктів у ситуаціях, коли легальні маршрути та канали збуту обмежені міжнародними нормативами.

Тіньовий флот виконує ключову роль у підтриманні тіньової економіки енергоресурсів, створюючи паралельну логістичну систему, що дозволяє реалізовувати нафтопродукти поза контролем міжнародних організацій, фінансових інститутів та регуляторних органів. Для власників та операторів суден участь у таких схемах є економічно вигідною: попит на нелегальні морські перевезення стимулює завищення фрахтових ставок, що більше ніж компенсує ризики експлуатації застарілих танкерів. Таким чином, тіньовий флот фактично формує альтернативний морський ринок, який за останні роки стрімко розширився та перетворився на важливу складову механізмів обходу санкцій.

Тіньовий танкерний флот перевозить по світу зростаючі обсяги санкційної нафти, забезпечує багатомільярдні доходи рф, Ірану, загрожує екологічній

безпеці у світовому океані, слугує шпигунською платформою, використовується для диверсійної діяльності в інтересах агресора.

Однак значення тіньового флоту виходить далеко за межі економічної площини. Він формує нові виклики для міжнародного морського права та глобальної системи безпеки транспортування нафтопродуктів. Судна, задіяні в тіньових операціях, часто не відповідають вимогам MARPOL та SOLAS, не проходять інспекцій, використовують несертифіковані екіпажі та не мають страхового покриття. Це створює високу ймовірність аварій із масштабними розливами нафти, особливо під час STS-перевантажень у міжнародних водах, де відсутній чіткий контроль юрисдикцій. Таким чином, існування тіньового флоту підвищує рівень ризиків для довкілля, морських екосистем та прибережних держав, які у разі аварій залишаються сам на сам з наслідками.

Підсанкційні країни шукають нові ринки збуту нафти, нарощують танкерний флот для обходу обмежень.

Тіньовий танкерний флот з загальним “дедвейтом” понад 100 млн тон (приблизно 17 % світового флоту нафтових танкерів) налічує понад тисячу переважно застарілих, погано обслуговуваних суден без належного страхування, з заплутаними структурами власності та управління, розташованими у “дружніх” юрисдикціях, під “зручними” прапорами. Такі судна вдаються до оманливих тактик у морі для приховування походження вантажів, загрожують екологічним хаосом та мільярдними збитками прибережним країнам через проходження жвавих вузьких міжнародних транспортних шляхів без лоцманських послуг. З початку повномасштабного вторгнення росії в Україну тіньові танкери вже брали участь у понад 50 інцидентах від Данських проток до Малайзії.

Поява і розвиток тіньового флоту зумовлені низкою факторів:

1. Глобальна політична напруга та санкційні режими. Країни, що перебувають під міжнародними обмеженнями, потребують непрозорих логістичних схем для експорту нафти, газу, вугілля та інших стратегічних

ресурсів. Тіньовий флот часто використовується як засіб обходу санкційних обмежень.

2. Економічна вигода. Перевезення ризикованих або заборонених вантажів є високомаржинальним бізнесом. Власники старих суден погоджуються працювати в тіньовому сегменті, адже можуть отримати значні прибутки.
3. Недосконаліть міжнародного регулювання. Розбіжності у правових нормах різних юрисдикцій, а також наявність “прапорів зручності” (flags of convenience) створюють можливості для маніпулювання даними.
4. Технічні можливості для фальсифікації. Використання застарілих або навмисно модифікованих AIS-передавачів дозволяє тіньовому флоту приховувати маршрути або передавати світові мережі оманливі сигнали.

Таким чином, тіньовий флот є продуктом поєднання політичних, економічних, технологічних та юридичних факторів, які створюють умови для масштабних нелегальних операцій у морському середовищі.

За даними CREA (Центр досліджень енергетики та чистого повітря), у листопаді доходи Росії від експорту викопного палива становили 611 млн євро на день. Найбільшим покупцем російських вуглеводнів традиційно є Китай. У листопаді він придбав російського викопного палива на 5 млрд євро, з яких 3,5 млн євро забезпечила сира нафта. Загалом саме на неї припадає 70% імпорту Китаю з Росії.


Рисунок 1.1 – Статистика купівлі країнами російської нафти

Другим найбільшим імпортером була Туреччина, яка забезпечила РФ 3 млрд євро за листопад.

На третьому місці – Індія, яка забезпечила РФ 2,1 млрд євро надходжень за місяць.

ЄС був четвертим за величиною покупцем російського викопного палива в листопаді (1,8 млрд євро) від п’ятірки найбільших покупців. Найбільшу частку закупівель російського викопного палива з боку ЄС складав трубопровідний газ (46%), а також ЗПГ (34%), які наразі поза санкціями. Окрім того, ЄС зробив виняток для російської сирої нафти, тому південною гілкою нафтопроводу "Дружба" до Угорщини, Словаччини та Чехії все ще імпортується російська нафта.

Підсумовуючи, сутність тіньового флоту транспортування нафтопродуктів полягає у функціонуванні нелегальної транспортної системи, яка забезпечує обхід регуляторних норм і санкцій та підтримує тіньову економіку енергетичного сектору. Його значення проявляється в економічній, політичній та екологічній площинах: від забезпечення фінансової стабільності країн-порушників до створення серйозних загроз морській безпеці та міжнародному правопорядку. Масштабність і стійкість цього явища визначають необхідність формування ефективних механізмів глобального моніторингу, посилення відповідальності судновласників та впровадження сучасних технологій контролю судноплавства, які здатні протидіяти подальшому розвитку тіньового флоту.

### 1.2. Методи та тактики діяльності тіньового флоту

Важливою ознакою тіньового флоту є маніпулятивна поведінка з навігаційною інформацією. Судна навмисно відключають AIS-транспондери, змінюють ідентифікатори MMSI або IMO, використовують підроблені

документи і приховують реальні маршрути. Така практика ускладнює моніторинг глобальних логістичних потоків та створює сприятливе середовище для протизаконної діяльності — від контрабанди нафти й зброї до забезпечення військових операцій держав-порушників. У разі конфліктних ситуацій це підвищує ризики військово-політичних інцидентів, оскільки правоохоронні та військово-морські структури не можуть оперативно ідентифікувати судно та його належність.

Одним із найпоширеніших методів є відключення або навмисне приглушення автоматичної ідентифікаційної системи (AIS). Попри вимоги конвенції SOLAS, судна тіньового флоту регулярно вимикають AIS-транспондери, залишаючи великі прогалини у своїх маршрутах. У деяких випадках використовуються так звані «темні зони», де наведення сигналу блокується, а судно стає тимчасово невидимим для супутникового моніторингу. Така практика дозволяє приховувати STS-операції, зміну пунктів призначення, передачу вантажу або заходи в порти, що не повинні відображатися в офіційних логістичних ланцюгах.

Іншим важливим інструментом є маніпулювання судновими ідентифікаторами — MMSI, IMO або назвами суден. Тіньові оператори змінюють ідентифікаційні номери, створюючи «дзеркальні» або дубльовані записи, або ж використовують фіктивні номери, що не відповідають реєстрам Міжнародної морської організації. Зміна прапорів також є типовою практикою, яка дозволяє переходити під юрисдикції з низькими вимогами до перевірки суден, уникати відповідальності та зменшувати ризик огляду портовими службами. Прапори зручності забезпечують мінімальний регулятивний нагляд і дозволяють швидко змінювати власника чи оператора судна через офшорні структури.

Юридичні та фінансові тактики доповнюють технічні методи маскування. Тіньовий флот використовує багаторівневі схеми володіння суднами, що включають офшорні компанії, номінальних власників і складні трастові

структури. Такі моделі ускладнюють встановлення реальної відповідальності у разі аварії або санкційних порушень. Крім того, перевізники уникають страхування в міжнародних P&I клубах і звертаються до малих, часто псевдо страхових компаній, що не забезпечують реального покриття. Це зменшує фінансові витрати, але підвищує ризики для інших учасників судноплавства та екосистем.

Тіньовий флот зосереджений переважно у регіонах із низьким рівнем контролю або складністю здійснення моніторингу, а саме:

- акваторії біля берегів Китаю та Південно-Східної Азії;
- Перська затока та територіальні води Ірану;
- території біля Західної Африки та Гвінейської затоки;
- райони Середземного та Чорного морів, де здійснюються STS-операції;
- південна частина Індійського океану.

Спільною рисою цих регіонів є велика кількість суден, складність покриття територій супутниковими системами та наявність економічної чи політичної мотивації для прихованих операцій.

Судна тіньового флоту використовують різноманітні тактики для приховування своєї діяльності, серед яких найпоширеніші:

- Вимкнення AIS-трансляції (dark activity), що призводить до тривалих розривів у сигналі.
- GPS-спуфінг — передача фальшивих координат для маскування реального маршруту.
- Маніпуляції з ідентифікаторами: зміна MMSI, IMO, назви судна або прапора.
- STS-операції (передача вантажу між суднами у відкритому морі) у відкритому морі для приховування передачі вантажу.
- Аномальні маршрути, які включають раптові зміни напрямку, зупинки в несподіваних локаціях або обхід основних транспортних коридорів.
- Перереєстрації у слабо контрольованих юрисдикціях, де управління судноплавством здійснюється з мінімальним наглядом.

Такі тактики роблять тіньовий флот одним із найскладніших об’єктів для моніторингу у глобальному масштабі.

### 1.3. Загрози та наслідки діяльності тіньового флоту

Тіньовий флот є багатовимірним явищем, що створює комплекс загроз для міжнародної безпеки, екології, економіки та глобальної транспортної інфраструктури. Його діяльність має довгострокові наслідки, які впливають як на окремі держави, так і на глобальні енергетичні ринки та міжнародний порядок. У цьому підрозділі розглянуто ключові групи загроз, зумовлених функціонуванням тіньового флоту, а також їх можливі наслідки.

Однією з основних небезпек є вплив на світові ринки енергоресурсів. Судна тіньового флоту перевозять сировину, що формально перебуває під санкціями, створюючи неконтрольовані обсяги пропозиції. Це призводить до спотворення ринкових сигналів, зниження ефективності санкційних режимів і послаблення конкурентоспроможності легальних перевізників. Значні обсяги нелегального транспортування позбавляють держави податкових надходжень, адже такі судна часто використовують прапори зручності, офшорні юрисдикції та схеми мінімальної звітності, що унеможливлює стягнення портових зборів та інших платежів.

З геополітичної точки зору тіньовий флот підриває ефективність міжнародних санкцій і механізмів стримування. Держави, що використовують такі схеми, отримують можливість продовжувати експорт стратегічних товарів, обходячи вимоги міжнародних інституцій. Це знижує результативність зовнішньої політики та ускладнює координацію дій між союзниками. Водночас тіньовий флот ставить під сумнів ефективність системи контролю прапорів і портового контролю, оскільки численні юрисдикції, через які здійснюється перереєстрація суден, часто не забезпечують належного нагляду.

Однією з основних причин існування тіньового флоту є намагання обійти міжнародні санкції, зокрема щодо експорту нафти та нафтопродуктів. За допомогою прихованих схем перевезення санкційні держави продовжують отримувати доходи від експорту енергоресурсів, що:

- знижує ефективність політичного тиску, створює нерівні умови на міжнародному енергетичному ринку,
- підриває міжнародну довіру до санкційних механізмів,
- дає змогу продовжувати фінансування військових або агресивних дій.

Фактично тіньовий флот виступає інструментом геополітичної боротьби та економічного виживання санкційних режимів.

Більшість суден тіньового флоту — це старі танкери, яким понад 20–30 років. Вони часто працюють без належного технічного огляду, міжнародного страхування та підтвердженого класу.

Це створює низку екологічних загроз:

- висока ймовірність аварій та розливів нафти, особливо під час STS-операцій у відкритому морі;
- відсутність страхового покриття, що унеможливлює фінансування ліквідації екологічних катастроф;
- порушення MARPOL-конвенцій, зокрема щодо транспортування небезпечних вантажів;
- ризики забруднення міжнародних морських шляхів, що призводить до довготривалих екологічних наслідків.

У випадку аварій держава або прибережний регіон часто залишаються без можливості ідентифікувати власника або відповідальну сторону.

Судна тіньового флоту часто навмисно вимикають AIS, що:

- ускладнює навігацію у густонаселених морських районах,
- збільшує ймовірність зіткнень,
- створює загрозу для інших суден, портів та екосистем,
- унеможливлює участь таких суден у системах попередження небезпеки.

Прихована поведінка суден негативно впливає на безпеку морського руху, що є критично важливим для глобальної транспортної системи.

Тіньовий флот тісно пов'язаний із ухиленням від оподаткування, контрабандою, фінансуванням кримінальних і терористичних організацій, нелегальною торгівлею зброєю або ресурсами.

Прибутки від таких схем часто спрямовуються на тіньові фінансові операції, а складні ланцюги судновласників, офшорних компаній та номінальних операторів ускладнюють ідентифікацію реальних бенефіціарів.

Комплексний характер загроз, створюваних тіньовим флотом, робить його одним із найбільш небезпечних явищ для глобальної морської галузі[16]. Він:

- підриває принципи міжнародного права,
- створює екологічні та техногенні ризики,
- порушує безпеку мореплавства,
- дестабілізує енергетичні ринки,
- підтримує незаконні економічні й військові операції.

Тіньовий флот є багатовимірною загрозою, що поєднує економічні, екологічні, безпекові та геополітичні ризики. Його діяльність підриває міжнародний правопорядок, знижує ефективність санкційних механізмів та створює реальні небезпеки для морських екосистем. З огляду на масштаби тіньового флоту та його зростаючу активність, міжнародні інституції потребують комплексних реформ системи моніторингу, контролю та притягнення до відповідальності учасників нелегальних морських операцій. Таким чином, ефективна боротьба з тіньовим флотом є стратегічно важливою для забезпечення міжнародної стабільності та економічної безпеки.

## 1.4. Значення аналізу геолокаційних даних для виявлення тіньового флоту

Аналіз геолокаційних даних відіграє ключову роль у виявленні, моніторингу та дослідженні діяльності тіньового флоту, який активно застосовує методи маскування для обходу санкцій та приховування руху нафтопродуктів. У сучасних умовах глобальної цифровізації саме геопросторові технології дозволяють з високою точністю відстежувати переміщення суден, визначати аномальні патерни поведінки та встановлювати зв’язки між окремими учасниками нелегальних логістичних схем. Значення таких даних зростає, оскільки традиційні інструменти контролю, засновані на документації, портовій звітності та страхових реєстрах, дедалі частіше виявляються неефективними щодо суден, які діють поза межами регулювання.

Геолокаційні дані морських суден, отримані через автоматичну ідентифікаційну систему (AIS), є одним із найцінніших джерел інформації для сучасних систем моніторингу морської діяльності. AIS-сигнали містять детальну інформацію про координати судна, швидкість, курс, час передачі, статус навігації та основні технічні характеристики. Комплексний аналіз цих даних дозволяє відстежувати поведінку суден у реальному часі та в ретроспективі, виявляти закономірності руху, ідентифікувати аномалії та встановлювати зв’язки між різними типами активності.

У контексті боротьби з тіньовим флотом аналіз геолокаційних даних набуває критично важливого значення, оскільки саме поведінкові ознаки та аномальні патерни в русі суден найбільш чітко відображають нелегітимну діяльність, приховані операції та спроби уникнути контролю[17].

Автоматична ідентифікаційна система була створена для підвищення безпеки судноплавства, однак із часом перетворилася на глобальну мережу спостереження за морським рухом. Щодня AIS генерує мільярди точок даних, що дозволяють:

- відстежувати позиції суден у режимі майже реального часу;
- аналізувати маршрути на глобальному рівні;
- визначати часово-просторові особливості руху;
- ідентифікувати взаємодії між суднами (зближення, зустрічі, STS-операції);
- формувати історичні профілі поведінки конкретного судна або групи суден.

Саме тому AIS став основою для технологій аналізу морського трафіку та виявлення підозрілих патернів руху.

Для тіньового флоту характерні тривалі паузи у передачі AIS-повідомлень, які можуть тривати від кількох годин до кількох днів. Аналіз таких пропусків дозволяє визначати спроби приховати небажану активність (STS-операції, зміну маршруту, захід у заборонену зону), виявляти переходи через міжнародну лінію зміни дат, де маніпуляції сигналами відбуваються найчастіше, ідентифікувати “темні періоди” — ключовий індикатор тіньового флоту. У сучасних аналітичних системах пропуски сигналів автоматично перетворюються на ризикові метрики, які дозволяють ранжувати судна за ступенем підозрілості.

Важливим аспектом є можливість реконструкції повної логістичної картини переміщення вантажів. Поєднання супутникових спостережень, AIS-сигналів, даних портових заходів та інформації про погодні умови дозволяє створити багатовимірні моделі поведінки флоту. За допомогою таких моделей можна визначити ймовірні маршрути транспортування санкційної нафти, виявити операції змішування вантажів, а також відстежити взаємодію між кількома судами, що здійснюють нелегальні STS-операції у відкритому морі. Це суттєво розширює можливості держав та організацій щодо ідентифікації ланцюгів постачання, які раніше залишалися прихованими.

Завдяки аналізу геолокаційних даних стає можливим автоматично виявляти характерні патерни:

- часті зміни MMSI або IMO під час руху;

- аномальні швидкісні характеристики — від нульової швидкості протягом тривалого часу до надмірно швидкого руху;
- аутлайєри у положенні — раптові “стрибки” координат;
- маніпуляції з маршрутом, пов’язані з обманом супутникового моніторингу;
- ряди повторюваних прихованих транзакцій (dark-STSs) у віддалених акваторіях.

Такі патерни дозволяють автоматизовано формувати профілі ризику суден та їхніх операторів.

Аналіз геолокаційних даних є фундаментально важливим для:

- підвищення прозорості морських перевезень;
- виявлення нелегальних торговельних ланцюгів;
- забезпечення енергетичної та економічної безпеки;
- підтримки санкційних режимів;
- зниження екологічних ризиків;
- формування глобальної системи моніторингу морських перевезень.

Особливе значення геолокаційний аналіз має для визначення рівня ризику суден. На основі траєкторного аналізу, історії відключень транспондерів, віку судна та патернів поведінки можна формувати рейтинги небезпеки, які використовуються державами, страховими компаніями та агентствами з контролю за санкціями. Це дозволяє виявляти судна, що ймовірно належать до тіньового флоту, навіть за відсутності прямої доказової бази, а також оптимізувати ресурси для морського моніторингу.

Без якісного аналізу AIS та супутникових даних виявлення тіньового флоту практично неможливе, оскільки більшість схем його діяльності маскується саме через маніпуляції геолокаційною інформацією.[18]

Отже, геолокаційні дані відіграють критичну роль у виявленні тіньового флоту, забезпечуючи можливість точного відстеження суден, аналізу аномалій, визначення зон ризику та прогнозування нелегальних операцій. Їхнє

використання суттєво підвищує ефективність міжнародних механізмів контролю та є невід’ємною складовою сучасної стратегії протидії незаконній діяльності в морському просторі.

У контексті зростання кількості тіньових танкерів та ускладнення логістичних схем значення геолокаційної аналітики лише зростатиме, стаючи одним із ключових інструментів забезпечення енергетичної та екологічної безпеки.

# РОЗДІЛ 2
# АНАЛІЗ ДОСЛІДЖУВАНИХ ДАНИХ ТА ПОШУК ОПТИМАЛЬНИХ ІНСТРУМЕНТІВ ПРОВЕДЕННЯ ДОСЛІДЖЕННЯ

## 2.1. Опис і структура використовуваних даних

Завдяки великій кількості статичної та динамічної інформації про судна, дані AIS мають величезний потенціал, який можна застосувати для вирішення різних наукових проблем, пов'язаних з морською галуззю. У цьому розділі представлено детальне тлумачення структури даних AIS та основних застосувань у прогнозуванні траєкторії руху суден, а також теми вищого рівня, що стосуються безпеки та сталого розвитку на морі.

Розроблена в 1990-х роках, система автоматичного відстеження (AIS) — це система автоматичного відстеження ближнього радіусу дії, призначена для надання інформації про ідентифікацію та позиціонування як суден, так і берегових станцій. Згідно з Конвенцією про безпеку людського життя на морі (SOLAS), судна вагою 300 валових тонн (GT) або більше, що здійснюють міжнародні рейси, вантажні судна вагою 500 валових тонн (GT) або більше, що не здійснюють міжнародні рейси, та пасажирські судна будь-якого розміру зобов'язані мати на борту систему AIS (Міжнародна морська організація, 2002).

Конкретні цілі обладнаної системи AIS включають ідентифікацію суден, допомогу у відстеженні цілей, допомогу в пошуково-рятувальних операціях (SAR), спрощення обміну інформацією та надання додаткової інформації для сприяння ситуаційній обізнаності (Міжнародна морська організація, 2015). Через морський діапазон дуже високих частот (VHF) інформація з бортових пристроїв AIS може передаватися з судна на судно, з судна на берегові станції та з судна на супутник.

AIS-дані складаються з двох основних категорій інформації: динамічної, яка генерується автоматично бортовими датчиками судна, та статичної/рейсової,

що вводиться вручну екіпажем або автоматично передається з навігаційних систем. Динамічна інформація включає координати судна (широту та довготу), курс, швидкість відносно ґрунту (SOG), курс судна (COG), істинний напрямок руху (HDG), швидкість повороту (ROT) та час останньої передачі сигналу. Ці параметри оновлюються з частотою від кількох секунд до кількох хвилин залежно від типу судна та його швидкості. Висока частота оновлення дозволяє точно відстежувати траєкторії руху та формує основу для аналізу поведінкових патернів суден.

Статичні та рейсові дані AIS мають іншу природу та змінюються значно рідше. До них належать унікальні ідентифікатори судна, такі як MMSI (Maritime Mobile Service Identity), IMO-номер, назва судна, тип і розміри корпусу. Крім того, AIS передає інформацію про призначений порт, очікуваний час прибуття (ETA), осадку (draught) та характер вантажу. Ці дані дозволяють класифікувати судно, визначати його можливості та маршрути, а також аналізувати відповідність заявленої інформації реальній поведінці. Саме тому статичні параметри є важливими для виявлення аномалій — наприклад, неправдивих MMSI, зміни ідентифікаторів або невідповідності між типом судна і його фактичним режимом роботи.

У сучасних аналітичних системах AIS-дані зберігаються у форматі таблиць або розподілених партиційованих наборів, де кожним записом є окреме повідомлення із зазначенням часу (timestamp), координат, ідентифікаторів та додаткових параметрів. Це дозволяє виконувати складні форми обробки, такі як відстеження траєкторій, кластеризація маршруту, пошук аномалій, аналіз STS-операцій або виявлення несанкціонованої діяльності. Геолокаційна складова AIS-даних робить їх незамінними для побудови просторово-часових моделей руху флоту.

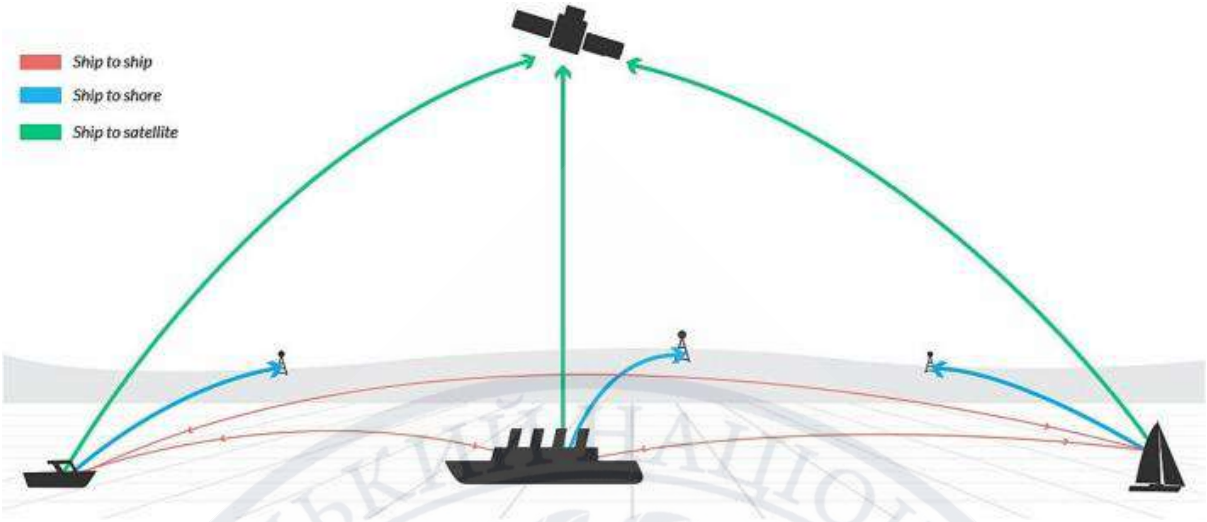


Рисунок 2.1 – Способи передачі AIS даних

Автоматична ідентифікаційна система (AIS) використовує кілька каналів і технологій для передачі інформації про місцезнаходження, параметри руху та ідентифікаційні характеристики суден. Від вибраного способу передачі залежить дальність дії, надійність отримання даних, рівень покриття та можливість моніторингу судноплавства у реальному часі. У сучасній практиці AIS-дані передаються у трьох основних режимах: наземними приймальними станціями, супутниками, а також у внутрішніх мережах суден.

Наземний AIS працює через VHF-радіочастоти (морські діапазони 161.975 MHz та 162.025 MHz).

Передавачі на суднах надсилають повідомлення, які приймаються прибережними радіостанціями, портовими системами, маяками, береговими центрами моніторингу. Цей спосіб передачі є базовим та історично першим у системі AIS.

Переваги наземного AIS: висока частота сигналів (оновлення кожні 2–10 секунд), низька затримка передачі, висока точність сигналу, підтримка двосторонньої VHF-комунікації.

Недоліки: обмежена дальність — залежить від висоти антен (приблизно 40–70 км), немає покриття у відкритому океані., слабкий прийом у районах з високою щільністю судноплавства (сигнальні колізії).

Сучасні супутники оснащені приймачами AIS (S-AIS), які збирають сигнали від суден на всій поверхні Землі.

Супутники рухаються полярними або низькими орбітами (LEO), що дозволяє отримувати AIS-дані у відкритому океані, покривати важкодоступні райони, збирати глобальні траєкторії.

Переваги S-AIS: глобальне покриття, включаючи океанічні зони, можливість отримання сигналів у регіонах без наземної інфраструктури, незамінне джерело для виявлення тіньового флоту.

Недоліки: менша частота сигналів (оновлення раз на 5–60 хв залежно від супутникової групи), колізії сигналів у зонах з високою щільністю руху, затримки, пов’язані з проходженням супутника над районом.

Внутрішні корабельні системи (Intra-ship AIS routing) – AIS-дані можуть передаватися всередині корабля на інші навігаційні та інформаційні системи, такі як:

- ECDIS (Electronic Chart Display and Information System),
- radar/ARPA,
- суднові навігаційні панелі,
- бортові сервери,
- внутрішні локальні мережі (LAN).

Це забезпечує повну інтеграцію даних AIS у суднові навігаційні процеси.

| **Назва поля** | **Опис** |
|---|---|
| MMSI | Унікальний номер, який ідентифікує судно або морську станцію, як номер телефону для радіозв'язку |
| IMO | Унікальний ідентифікаційний номер, що присвоюється судну Міжнародною морською організацією (IMO) для |

| | |
|---|---|
| | підвищення безпеки, запобігання шахрайству та забрудненню. Цей номер незмінний протягом усього життя судна, на відміну від його назви, прапора або власника. |
| Longitude / Latitude | Довгота та широта місця положення судна, з точністю до 10 метрів |
| Timestamp | Дата і час фіксування сигналу |
| Heading | Напрямок руху судна |
| Country Code | Код країни, під прапором якої ходить судно |

Таблиця 2.2 – Ключові поля геолокаційного датасету морських суден

Основна інформація AIS, що надсилається суднами, наведена в таблиці 2.2. Ці повідомлення в основному створюються трьома способами: встановлюються під час встановлення, оновлюються автоматично та вводяться вручну (Міжнародна морська організація, 2015). Наприклад, навігаційний статус має вводитися вахтовим помічником капітана (OOW) вручну, включаючи: рух за допомогою двигунів, рух під вітрилами, судно на якорі, без керування, обмежене в маневруванні, пришвартоване, обмежене осадкою, на мілині, зайняте риболовлею тощо.

Крім того, частота оновлення даних AIS може відрізнятися залежно від типу повідомлення, стану руху та типу судна. Наприклад, статична інформація та інформація, що підлягає рейсу, оновлюється кожні 6 хвилин або за запитом; для суден, оснащених судновим обладнанням класу A, динамічна інформація передається автоматично кожні 2–10 секунд під час руху та кожні 3 хвилини під час стояння на якорі та руху зі швидкістю менше 3 вузлів; для суден, оснащених судновим обладнанням класу B, динамічна інформація автономно надсилається кожні 3–30 секунд під час руху та кожні 3 хвилини під час руху зі швидкістю менше 2 вузлів.

Оскільки дані AIS стають дедалі доступнішими, зростає попит на виявлення аномальних даних AIS. Однак ручне вилучення підозрілої активності суден з масивів даних AIS є недоцільним. Згідно з нещодавніми дослідженнями виявлення аномалій суден, підходи машинного навчання (ML) набули більшої популярності в автоматичному виявленні різних аномальних треків суден.[4]

Аномалії в треках AIS стосуються поведінки, яка відхиляється від норми або є неочікуваною під час типових операцій. У контексті руху суден це може включати раптові зміни швидкості судна або незвичайні маршрути руху, які не відповідають стандартним практикам. Загальні форми аномальної поведінки суден включають відхилення від нормальних маршрутів, переривчасті траєкторії, неочікувані прибуття в порт, аномалії близького зближення та аномалії входу в зону:

- Відхилення від звичайних маршрутів: Відхилення від прямого маршруту у відкритому морі або від прямих шляхів між заздалегідь визначеними точками маршруту у складних водах може свідчити про аномалію.
- Переривчасті траєкторії: Навмисне вмикання-вимикання обладнання AIS або інші навмисні дії можуть призвести до втрати сигналу AIS, що потенційно може приховати зловмисні дії.
- Несподівані прибуття в порт: судна можуть прибувати в неочікувані порти з різних незаконних причин (наприклад, незаконний промисел). Ці аномалії можна виявити за допомогою інформації, пов'язаної з рейсом, з даних AIS, такої як порт призначення, очікуваний час прибуття та відповідні точки маршруту.
- Аномалії близького зближення: Близьке зближення між суднами, яке триває тривалий час, є рідкістю, за винятком надзвичайних ситуацій. Такі моделі руху можуть свідчити про незаконну діяльність на морі (наприклад, обмін контрабандою або наркотиками).
- Аномалії входу в зону: Вхід суден в обмежену зону (наприклад, морські охоронні зони) на значний період слід вважати аномальною діяльністю.

Для визначення підсанкційних суден додатково потрібно сформувати єдиний датасет санкцій об’єднавши списки які надають Управління контролю за іноземними активами США та Європейський союз.

У рамках своїх правоохоронних зусиль управління контролю за іноземними активами (OFAC) публікує список осіб та компаній, що належать або контролюються країнами, що підпадають під дію цих країн, або діють від їхнього імені. Він також містить список осіб, груп та організацій, таких як терористи та наркоторговці, визначених у рамках програм, що не стосуються конкретної країни. Разом такі особи та компанії називаються «особисто зареєстрованими особами» або «SDN». Їхні активи блокуються, а громадянам США, як правило, заборонено мати з ними справу.

## 2.2. Пошук оптимальних інструментів для аналізу геолокаційних даних

Аналіз геолокаційних даних – це процес аналізу даних про місцезнаходження для виявлення закономірностей, тенденцій та аналітики. Це дозволяє підприємствам, урядам та дослідникам приймати обґрунтовані рішення на основі географічних та просторових зв'язків. Від міського планування до ліквідації наслідків стихійних лих, геопросторовий аналіз трансформує галузі промисловості в усьому світі.

Пошук оптимальних інструментів для аналізу геолокаційних даних є ключовим завданням сучасної морської аналітики та моніторингу. Оскільки обсяг, різноманітність і швидкість надходження даних про рух суден стрімко зростають, виникає потреба у виборі платформ і технологій, здатних забезпечити ефективну обробку, інтеграцію та інтерпретацію просторово-часової інформації. Геолокаційні дані, зокрема AIS-повідомлення, супутникові спостереження та сигнали радіонавігаційних систем, характеризуються високою частотою оновлення, нерівномірністю покриття та потенційними помилками, що потребує

застосування інструментів із високою стійкістю до шуму та здатністю працювати з великими даними.

Одним із основних критеріїв вибору інструментів є їхня здатність обробляти розподілені та великомасштабні просторово-часові дані. У цьому контексті важливу роль відіграють платформи паралельної обробки та розподілених обчислень, такі як Apache Spark, Apache Sedona, Google BigQuery або ClickHouse. Вони дозволяють агрегувати мільйони записів AIS, виконувати складні геометричні операції, аналізувати траєкторії суден та швидко виконувати просторові фільтри, навіть у випадках багаторічних або глобальних вибірок. Перевагою таких систем є масштабованість і можливість інтеграції з іншими джерелами даних, що є критично важливим для аналізу діяльності тіньового флоту.

Другим важливим аспектом є інструменти геопросторового аналізу. Геоінформаційні системи та бібліотеки обробки геометрії забезпечують засоби для аналізу структури маршрутів, визначення зон активності, виявлення STS-операцій та моделювання просторових патернів. Ці інструменти дозволяють виконувати геометричні перетворення, створювати буферні зони, аналізувати перетини та визначати області інтересу. Вони є необхідними для задач, у яких точність геопросторових розрахунків має першорядне значення.

Третім важливим критерієм є використання інструментів машинного навчання та штучного інтелекту. Бібліотеки та фреймворки, такі як TensorFlow, PyTorch, Scikit-learn чи XGBoost, дозволяють будувати моделі класифікації та кластеризації траєкторій, виявляти аномальні патерни руху, прогнозувати поведінку суден та ідентифікувати підозрілі операції. Застосування нейронних мереж та алгоритмів послідовнісного аналізу (LSTM, GRU, трансформери) відкриває можливості для автоматизованого виявлення періодів відключення AIS або аномальних змін курсу суден. Це дозволяє створити систему раннього попередження щодо можливої діяльності тіньового флоту.

Сучасний геопросторовий аналіз спирається на потужні технології, такі як геоінформаційні системи (GIS), дистанційне зондування та GPS-відстеження в режимі реального часу. Методи візуалізації, такі як інтерактивні карти, теплові карти та просторова кластеризація, дозволяють дослідникам ефективно аналізувати географічну аналітику. Інструменти спрощують доступ до геопросторових даних, пропонуючи рішення для геокодування, маршрутизації, ізохрон та пошуку місць, що спрощує інтеграцію просторової аналітики в додатки.

Зі зростанням популярності Big Data та штучного інтелекту геопросторовий аналіз став доступнішим, ніж будь-коли. Підприємства можуть оптимізувати ланцюги поставок, міста можуть покращити транспортний потік, а екологічні організації можуть моніторити зміну клімату — все це завдяки можливостям географічної аналітики.

Аналіз геопросторових даних пропонує потужні аналітичні дані, але також пов'язаний з такими проблемами, як точність даних, проблеми конфіденційності та обчислювальні вимоги. Водночас, досягнення в галузі штучного інтелекту, обробки в режимі реального часу та хмарних рішень формують майбутнє цієї галузі.

Виклики:

- Точність і надійність даних – Помилки в сигналах GPS, застарілі карти та несумісні джерела даних можуть призвести до оманливих висновків. Забезпечення високоякісних, оновлених даних є надзвичайно важливим. Дізнайтеся більше про пошук і перевірку адрес для підвищення точності.
- Конфіденційність і безпека – Збір і зберігання даних про місцезнаходження викликає занепокоєння щодо конфіденційності. Дотримання таких правил, як GDPR, та безпечні методи обробки даних мають вирішальне значення.
- Обчислювальна складність – Обробка великих наборів даних, таких як супутникові знімки або дані датчиків Інтернету речей, вимагає значної

обчислювальної потужності та ефективних алгоритмів. Хмарні геопросторові рішення допомагають впоратися з цією складністю.

- Проблеми інтеграції – Багато галузей мають труднощі з інтеграцією геопросторової аналітики з існуючими IT-системами та системами бізнес-аналітики, що робить безперебійну сумісність ключовим завданням.
- Доступність і експертиза – Геопросторові інструменти часто вимагають технічних знань. Для подолання цієї прогалини потрібні зручніші API та автоматизовані картографічні рішення.

Пошук оптимальних інструментів для аналізу геолокаційних даних, таким чином, передбачає баланс між обчислювальною продуктивністю, точністю геопросторової обробки, можливостями машинного навчання та ефективністю інтерактивної візуалізації. Комплексний підхід — поєднання розподілених систем обробки, геоінформаційних технологій, інструментів штучного інтелекту та спеціалізованих платформ даних — створює основу для ефективного виявлення нелегальної або аномальної активності тіньового флоту. Саме вибір оптимального набору інструментів визначає якість сучасного морського моніторингу та здатність аналітичних систем протидіяти загрозам, пов’язаним з непрозорими логістичними схемами.

Існує кілька інструментів та бібліотек для роботи з “Big Data”, розроблених для підтримки геопросторових робочих навантажень (або розширення загальних механізмів великих даних для цієї мети). Нижче наведено список таких інструментів з описами, перевагами/недоліками та варіантами використання, особливо для геоданих у великому масштабі:

**Mosaic (Databricks Labs)**

Тип: Просторове розширення для Apache Spark

Опис: Mosaic — це геопросторова бібліотека, яка розширює Spark розширеними просторовими функціями, типами геометрії та оптимізаціями, особливо зосередженими на продуктивності та простоті використання для великих геопросторових наборів даних.

Переваги: оптимізовані просторові примітиви, створено спеціально для сучасних версій Spark з урахуванням продуктивності, “чистіші” API для геопросторових запитів.

Недоліки: дещо новіший/менш зрілий, ніж Apache Sedona, може мати менше ресурсів спільноти або підтримки граничних випадків, залежність від сумісності з версіями екосистеми Spark.

**Apache Sedona**

Тип: Просторове розширення на Apache Spark / Flink

Опис: просторовий обчислювальний рушій, який розширює можливості Apache Spark (та інших рушіїв) геопросторовими типами даних, просторовими операторами запитів, просторовим індексуванням, секціонуванням, просторовими об'єднаннями та просторовим SQL. Він обробляє просторові дані як першокласні дані в розподілених обчислювальних контекстах.

Переваги: використовує розподілену інфраструктуру Spark, підтримує просторове індексування (наприклад, R-дерево, квадродерево) та ефективні просторові об'єднання, просторові функції (наприклад, відстань, перетин, буфер), інтегровані в API SQL / DataFrame, працює з великомасштабними наборами даних, може обробляти геолокаційні та інші дані разом, мовні прив'язки (Scala, Java, Python), інтегрується з форматами сховища, файловими системами тощо.

Недоліки: для надзвичайно великих наборів даних (особливо растрових або надзвичайно щільних векторних даних) потрібне налаштування продуктивності, накладні витрати на розподіл та перетасування просторових даних, крива навчання правильному використанню просторового секціонування та налаштуванню індексу, не завжди оптимальний для дуже частих невеликих оновлень (потокове передавання може мати обмеження)[12].

**GeoTrellis**

Тип: Бібліотека Scala для растрової/просторової обробки

Опис: GeoTrellis — це бібліотека географічних даних, що спеціалізується на обробці великомасштабних растрових даних та геопросторовій аналітиці, побудована на Spark. Вона оптимізована для алгебри карт, мозаїчного розміщення та просторових операцій на великих растрових сітках.

Переваги: чудово підходить для растрових робочих навантажень (супутникові знімки, цифрові моделі моделювання, дистанційне зондування), підтримує мозаїчне розміщення, нашарування та паралельну обробку растрових даних, інтегрується з Spark для масштабування

Недоліки: менший акцент на операціях з векторною/точковою геометрією — хоча підтримує обидві, більша спеціалізація — може бути надмірною, якщо ваше робоче навантаження переважно пов'язане з векторною/траєкторійною обробкою, крутіша крива навчання роботі зі змішуванням растрових та векторних даних

**GeoMesa**

Тип: Розподілена просторова база даних / індекс на основі сховищ великих даних

Опис: GeoMesa — це розподілена просторово-часова база даних, побудована на основі систем великих даних, таких як Apache Accumulo, Apache HBase, Apache Cassandra та Apache Kafka. Вона підтримує просторове та часове індексування та запити у великих масштабах.

Її часто використовують як серверну частину, що підтримує швидкі геопросторові запити до великих наборів даних.

Переваги: потужний у просторово-часовому індексуванні, швидкі запити, розроблений для великомасштабного сховища + запитів (не лише пакетних обчислень), добре підходить для потокової передачі + просторового прийому даних у реальному часі та запитів, підтримує різні серверні частини сховища

Недоліки: складна інфраструктура: потрібні HBase / Accumulo / Cassandra тощо; менш підходить для виконання важкої логіки обчислень поза БД, операційні накладні витрати.

## 2.3. Аналіз даних за допомогою Apache Spark

Для виконання обчислювальних операцій і роботи з даними було обрано технології Apache Spark із розширювальною бібліотекою Apache Sedona.


Рисунок 2.3 – Ключові переваги Apache Spark + Sedona

Apache Spark є одним із найпотужніших інструментів для обробки великих даних, особливо у випадках, коли потрібно працювати з великими просторово-часовими наборами, такими як сигнали AIS, супутникові вимірювання чи великі журнали подій. Його архітектура орієнтована на розподілені обчислення, що дозволяє ефективно масштабувати обробку даних на кластері та забезпечувати високу продуктивність навіть при роботі з терабайтами або петабайтами інформації. Використання Spark є особливо актуальним у задачах, пов’язаних з аналізом руху суден, виявленням аномалій, моделюванням патернів поведінки та моніторингом діяльності тіньового флоту.

Apache Spark спочатку був розроблений для паралельної обробки величезних наборів даних на кластерах машин. Він забезпечує відмовостійку модель обчислень у пам'яті, що робить його ідеальним для ітеративної обробки даних та складних аналітичних конвеєрів. Для даних AIS, які є як безперервними, так і великого обсягу, горизонтальна масштабованість Spark дозволяє аналітикам обробляти дані, що охоплюють океани та роки історії, не перевантажуючи окрему машину.

Однак, сам Spark обробляє широту та довготу лише як числові стовпці. Він може виконувати числові фільтри, але не має здатності розуміти просторові співвідношення, такі як “в межах 5 кілометрів” або “всередині цього полігону”. Саме тут Apache Sedona фундаментально змінює рівняння[11].

Структурований підхід до даних, що реалізований у Spark SQL і DataFrame API, дозволяє виконувати обчислення високого рівня абстракції та поєднувати їх із класичними інструментами SQL-аналітики. Це спрощує обробку великих масивів даних та робить можливим застосування складних фільтрів, часових вікон, агрегацій та групування. Наприклад, визначення періодів «темної активності» судна можна здійснити за допомогою віконних функцій, а кластеризацію маршрутів — через попередню сегментацію та агрегування координат за часовими інтервалами. Spark дозволяє комбінувати ці можливості з функціями машинного навчання, що робить його універсальним інструментом для Data Science.

Apache Sedona (раніше GeoSpark) розширює рушій Spark вбудованими геопросторовими можливостями. Він вводить такі типи геометрії, як точки, лінії та полігони, а також широкий спектр просторових функцій (ST_Within, ST_Distance, ST_Intersects та багато інших). Завдяки цим функціям просторове мислення стає першокласною операцією в Spark SQL та DataFrame API.

У контексті даних AIS це означає, що можна висловлювати запити на близькість або фільтрацію на основі зон з тією ж простотою, що й стандартний SQL.

```
SELECT a.mmsi, b.mmsi
FROM ais_data a, ais_data b
WHERE ST_Distance(a.geom, b.geom) < 5000
  AND ABS(a.timestamp - b.timestamp) < 600
```

Лістинг 2.4 – приклад SQL запиту для пошуку AIS сигналів в заданому радіусі та часовому проміжку

Такі запити дозволяють знаходити випадки, коли два судна знаходилися на відстані 5 кілометрів та 10 хвилин одне від одного — поширений показник для виявлення можливих скоординованих рухів або поведінки стеження. Без Sedona реалізація цього вимагала б використання спеціального коду та порівнянь методом грубої сили, що є обчислювально непосильним у глобальному масштабі.

Однією з найважливіших переваг Sedona є її здатність інтелектуально розділяти та індексувати просторові дані. Замість порівняння позицій кожного судна з іншими, Sedona розділяє простір на комірки (використовуючи квадро-дерева, k-d дерева або R-дерева) та створює локальні просторові індекси в межах кожного розділу. Це гарантує, що порівнюються лише відповідні підмножини даних, що значно зменшує обчислювальні витрати.

Іншою важливою перевагою Apache Spark є його здатність масштабувати обчислення горизонтально. Обробка, яка потребувала б години або навіть дні на одній машині, може бути виконана на десятках вузлів у кластері протягом хвилин. Це дає змогу проводити глибокий аналіз історичних даних, моделювати багаторічні тренди, будувати карти активності флоту та проводити ретроспективні розслідування інцидентів, які могли б залишитися непоміченими у випадку ручної обробки або обмежених обчислювальних ресурсів.

У практичному сценарії, наприклад, аналізі глобальних даних AIS за рік, ця оптимізація може означати різницю між обчисленням, яке займає години, а не дні. Там, де традиційні системи, такі як PostGIS, намагалися б зберегти в пам'яті навіть місячний обсяг даних, Spark із Sedona може обробляти петабайти, розподіляючи робоче навантаження по кластеру.

Головною перевагою Apache Spark із Sedona є його єдине аналітичне середовище. Аналітика суден рідко включає лише просторові дані. Вона часто поєднує додаткові виміри, такі як тип судна, власність, списки санкцій або умови навколишнього середовища, такі як вітер чи течія. Spark забезпечує основу для безперешкодної інтеграції всіх цих джерел даних.

За допомогою Sedona можна виконувати просторові об'єднання, а за допомогою вбудованих функцій Spark застосовувати часові фільтри, агрегації та навіть моделі машинного навчання (через Spark MLlib) — все в межах одного конвеєра. Наприклад, можна спочатку ідентифікувати всі судна, які наблизилися до санкціонованого судна, а потім використовувати алгоритми кластеризації для виявлення повторюваних шаблонів близькості. Таке об'єднання робочих процесів геопросторового мислення та науки про дані робить Spark + Sedona комплексним рішенням, а не спеціалізованим нішевим інструментом.

Ще однією причиною, чому Apache Spark із Sedona чудово підходить для аналітики AIS, є його широка сумісність з екосистемою. Spark може зчитувати дані практично з будь-якого джерела даних, включаючи Parquet, ORC, CSV, GeoJSON та GeoParquet, і може працювати на хмарних системах зберігання даних, таких як Amazon S3, Google Cloud Storage або локальних кластерах MinIO. Sedona розширює цю універсальність до просторових форматів та систем координат, що спрощує інтеграцію траєкторій суден, шейп-файлів портів та морських кордонів в один аналітичний робочий процес.

У підсумку, Apache Spark є універсальним інструментом для аналізу великих масивів геолокаційних даних, що забезпечує високу продуктивність, гнучкість і можливість інтегрувати просторові, часові, статистичні та машинні методи в єдиному середовищі. Його застосування дає змогу здійснювати точний, масштабований і автоматизований аналіз поведінки флоту, забезпечуючи необхідний рівень аналітичної підтримки для виявлення аномалій, прогнозування руху суден і протидії діяльності тіньового флоту. Spark фактично стає фундаментом сучасної інфраструктури морської аналітики та одним із ключових інструментів дослідження глобальних морських процесів.

## 2.4. Алгоритм аналізу даних за допомогою Apache Spark

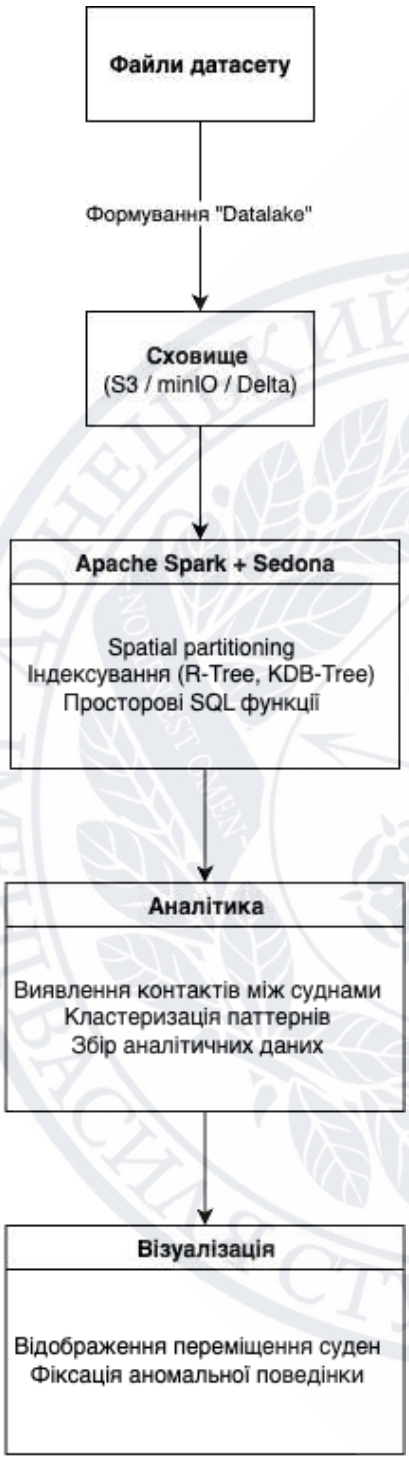


Рисунок 2.5 – Схема роботи алгоритму аналітичного інструменту

Ця архітектура ілюструє, як дані суден AIS та пов'язані з ними морські набори даних рухаються аналітичним конвеєром — від отримання та

розподіленого сховища до масштабного геопросторового аналізу в Apache Spark з Apache Sedona, і, нарешті, до систем візуалізації.

Система розроблена як модульна, хмарно-незалежна та масштабована, підтримуючи як пакетну аналітику (для історичних даних AIS), так і потокову аналітику (для моніторингу суден у реальному часі).

**Рівень джерел даних**

На вершині архітектури знаходяться постачальники даних, які генерують та надають вхідні потоки до системи. Вони можуть включати:

Дані AIS: дані місцезнаходження суден в певний момент часу (MMSI, широта, довгота, швидкість, позначка часу).

Списки санкцій: Набори даних, що ідентифікують судна та їх власників, на які накладено санкції або судна з високим рівнем ризику.

**Рівень сховища**

Рівень сховища даних є фундаментальним компонентом будь-якого аналітичного інструменту, оскільки саме на цьому етапі відбувається збереження, структуризація та оптимізація доступу до великих масивів інформації. Для систем, які обробляють геолокаційні дані — зокрема сигнали AIS, супутникові спостереження та результати аналітичної обробки — ефективність роботи аналітичного алгоритму значною мірою залежить від того, наскільки правильно організовано дані в сховищі. Саме цей рівень визначає швидкість виконання запитів, стабільність роботи, можливість масштабування та якість подальших аналітичних операцій.

На цьому етапі очищені та нормалізовані дані зберігаються в масштабованих, відмовостійких системах сховища:

Amazon S3, Google Cloud Storage, HDFS або MinIO (для локального зберігання).

Delta Lake або Apache Iceberg можна використовувати для керування версіями даних та еволюцією схеми.

Цей шар діє як озеро даних, зберігаючи як необроблені, так і оброблені розділи даних, він дозволяє Spark паралельно зчитувати дані з багатьох розділів, що є критично важливим аспектом для високопродуктивної аналітики наборів даних AIS, які можуть перевищувати кілька терабайтів.

Сховище даних виконує декілька ключових функцій. Передусім воно забезпечує надійне збереження та доступність інформації, що є критичним для аналізу великомасштабних потоків даних, характерних для морської аналітики. Геолокаційні дані мають високу частоту оновлення та великий обсяг, що потребує використання форматів з ефективною компресією та можливістю швидкого вибіркового читання — таких як Parquet або ORC. Вони дозволяють оптимізувати обсяг збереженої інформації, зменшити навантаження на файлову систему та прискорити виконання аналітичних операцій, таких як агрегації, фільтрація або побудова часових вікон.

Окрему увагу слід приділити можливості інтеграції результатів аналізу в сховище. Аналітичний інструмент зазвичай генерує нові дані — класифіковані траєкторії, зони ризику, моделі поведінки суден або оцінки ймовірності аномалій. Збереження таких результатів у сховищі дає змогу виконувати повторні аналітичні запити, будувати комбіновані моделі та створювати довгострокові дослідження тенденцій. Це робить сховище не лише архівом сирих даних, а й повноцінним компонентом циклу аналітики.

Таким чином, рівень сховища даних є ключовою частиною алгоритмічної системи для аналізу геолокаційних даних. Він забезпечує основу для ефективної обробки, збереження, масштабування та повторного використання інформації, необхідної для виявлення складних патернів руху суден і аналізу діяльності тіньового флоту. Від якості реалізації цього рівня залежить точність, швидкість і надійність усієї аналітичної інфраструктури, що робить його стратегічно важливим компонентом сучасних досліджень морського трафіку.

**Рівень обробки — Apache Spark з Apache Sedona**

Рівень обробки в архітектурі аналітичного інструменту, заснованого на геолокаційних даних, виконує центральну роль, оскільки саме на цьому етапі здійснюється трансформація сирих даних у корисну інформацію, необхідну для виявлення закономірностей, аналізу поведінки об’єктів та ідентифікації потенційних аномалій. Поєднання Apache Spark з Apache Sedona формує високопродуктивний, масштабований та гнучкий обчислювальний середовище, оптимізоване для просторово-часової аналітики, включно з аналізом руху суден та виявленням тіньового флоту.

Apache Spark забезпечує основу для розподіленої обробки великих обсягів даних. Його архітектура дозволяє виконувати паралельні операції на кластері машин, що робить можливим опрацювання багаторічних історичних наборів AIS-даних, супутникових вимірювань або величезних потоків телеметрії у режимі реального часу. Spark пропонує високорівневі API — SQL, DataFrame та Dataset — що забезпечують зручні засоби для фільтрації, агрегації, групування та побудови часових вікон. Завдяки цьому аналітик може працювати з мільйонами або мільярдами записів без втрати продуктивності та без потреби у складній ручній оптимізації.

Spark забезпечує:

- Розподілені обчислення між вузлами (горизонтальне масштабування).
- API для обробки SQL, DataFrame та потокової передачі.
- Вбудовану відмовостійкість та керування пам'яттю.

Модель виконавця Spark розподіляє роботу по кластеру, де кожен вузол обробляє частину даних AIS.

Sedona розширює можливості Spark за допомогою геопросторових типів та функцій, що дозволяє виконувати просторові об'єднання та запити на відстань безпосередньо в SQL.

Ключові функції включають:

- Просторове розбиття (KDB-дерево, QuadTree).
- Просторове індексування (R-дерево для ефективного пошуку).

- Геометричні операції, такі як ST_Within, ST_Distance, ST_Intersects.

Sedona перетворює пари широти/довготи на фактичні геометричні об'єкти (ST_Point) та використовує індекси для ефективного обчислення зв'язків.

Використання Spark у поєднанні з Sedona дозволяє реалізувати комплексну просторово-часову аналітику. Наприклад, система може фільтрувати судна, що входили у визначену зону протягом певного періоду, аналізувати взаємодію кількох об’єктів у межах заданого радіуса або визначати аномальні зміни траєкторій. Це є критичним у задачах виявлення тіньового флоту, де часто зустрічаються приховані операції — від вимкнення AIS до перевантаження нафти «борт у борт» у міжнародних водах. Просторові операції Sedona дозволяють моделювати такі події з високою точністю та виявляти патерни, які важко визначити класичними методами.

Spark і Sedona забезпечують також ефективне управління даними у часовому вимірі. Оскільки сигнали AIS надходять із високою частотою, ключовим завданням є визначення закономірностей у часових рядах — наприклад, пошук інтервалів без передачі сигналів, аналіз швидкості судна або сегментація маршруту на логічні фази. Завдяки віконним функціям Spark ці завдання виконуються швидко і масштабовано. У свою чергу, Sedona дозволяє поєднати часові патерни з просторовими характеристиками, створюючи двовимірний підхід до аналізу поведінки суден.

Цей шар обробки може обробляти мільярди просторово-часових записів, розподілених по вузлах, що неможливо в традиційних базах даних.

Таким чином, рівень обробки на основі Apache Spark та Apache Sedona забезпечує комплексний, масштабований та високопродуктивний механізм аналітики геолокаційних даних. Його можливості дозволяють здійснювати детальний просторово-часовий аналіз, виявляти аномалії у поведінці суден, ідентифікувати сигнатури тіньового флоту та формувати прогнозні моделі. Цей рівень є центральним компонентом аналітичного інструменту та значною мірою визначає ефективність роботи всієї системи.

**Рівень аналітики**

Після того, як Sedona завершить просторові об'єднання та перетворення, оброблені результати передаються до аналітичних конвеєрів, які можуть включати:

- Виявлення близькості: Виявлення суден, які часто курсують поблизу суден, на які поширюються санкції.
- Аналіз траєкторії: Кластеризація схожих шляхів руху для виявлення закономірностей.
- Виявлення аномалій: Навчання моделей машинного навчання для виявлення нестандартної поведінки, такої як відхилення суден від звичайних маршрутів.
- Оцінка ризику: Поєднання просторових характеристик (близькість, тривалість, частота) в метрики ризику для дотримання морських норм.

Spark MLlib можна використовувати тут для масштабованого навчання моделей та висновків на мільйонах записів.

Саме на цьому рівні відбувається інтерпретація просторово-часових даних, формування аналітичних моделей та виявлення закономірностей, що залишаються непомітними при класичному огляді геолокаційних сигналів. Рівень аналітики об’єднує методи статистики, машинного навчання, геопросторового моделювання та доменних експертних правил, створюючи цілісний механізм виявлення підозрілої поведінки суден.

Однією з ключових задач цього рівня є виявлення аномалій у поведінці суден. Використовуючи результати просторової обробки, система може визначати типові та нетипові маршрути, аналізувати частоту передач AIS-повідомлень, виявляти різкі зміни курсу або швидкості. Поведінкові індикатори, такі як нехарактерні зупинки у відкритому морі, відхилення від усталених торговельних шляхів чи позиційні стрибки, можуть свідчити про намір приховати реальну активність. Ці ознаки аналізуються як через класичні

статистичні методи, так і через алгоритми машинного навчання — кластеризацію траєкторій, моделі відхилення або часові нейронні мережі.

Рівень аналітики також включає побудову комплексних профілів поведінки суден. До таких профілів можуть входити дані про маршрути, історичні патерни руху, частоту вимкнення AIS, географічні зони активності та зв’язки з іншими суднами. Використовуючи ці профілі, система може оцінювати ризик належності судна до тіньового флоту та формувати індекс ризику, що базується на сукупності факторів. Це дозволяє будувати довгострокові моделі прогнозування та визначати ключові тенденції у поведінці мирового флоту.

Рівень аналітики також передбачає інтеграцію зовнішніх даних, таких як метеоінформація, інформація про порти, супутникові спостереження або економічні показники. Це дозволяє підвищити точність інтерпретації та уникнути хибнопозитивних висновків. Наприклад, зміна маршруту може пояснюватися погодними умовами або закриттям портів, а не тіньовою активністю. Інтегрованість аналітичного рівня забезпечує комплексний підхід до оцінки поведінки суден.

У підсумку рівень аналітики виступає ключовим етапом у перетворенні геолокаційних даних у реальні управлінські рішення. Саме тут формується розуміння структури морського руху, виявляються приховані операції тіньового флоту, створюються індикатори ризику та будуються моделі прогнозування поведінки суден. Ефективність роботи всієї аналітичної системи значною мірою залежить від якості цього рівня, оскільки саме аналітика встановлює причинно-наслідкові зв’язки, дозволяє виявити закономірності та створює основу для протидії незаконній діяльності в морському просторі.

**Рівень візуалізації та звітності**

Рівень візуалізації та звітності завершує цикл аналітичного процесу, перетворюючи результати обробки та аналітики на зрозумілі, доступні та інформативні форми. Його основне завдання — представити складні просторово-часові дані, аналітичні моделі та висновки у вигляді, що полегшує їх

інтерпретацію фахівцями, аналітиками, державними органами та іншими стейкголдерами. У контексті дослідження тіньового флоту цей рівень відіграє критичну роль, оскільки візуальна форма подання інформації дозволяє швидко виявляти аномалії, оцінювати масштаби порушень та приймати обґрунтовані управлінські рішення.

Результати аналітики зберігаються та візуалізуються для кінцевих користувачів, таких як морські аналітики, регулятори або групи з дотримання вимог.

Типові інструменти включають:

- Kepler.gl для 3D-геопросторової візуалізації траєкторій суден та зустрічей з ними.
- Grafana для панелей інструментів у режимі реального часу, що відображають сповіщення про близькість суден.
- Superset або Power BI для агрегованої статистики та звітності.

Візуалізації можуть відображати теплові карти щільності суден, мережі близькості або карти покадрової зйомки морського руху.

У цілому рівень візуалізації та звітності завершує роботу аналітичного інструменту, перетворюючи результати великомасштабної обробки та моделювання на доступну, структуровану та зрозумілу інформацію. Він забезпечує наочне представлення просторових і часових залежностей, формує підґрунтя для прийняття рішень та створює можливість ефективної комунікації між експертами. Візуальна аналітика стає ключовим інструментом у виявленні та дослідженні діяльності тіньового флоту, сприяючи підвищенню точності, прозорості та оперативності аналітичних процесів.

### 2.5. Візуалізація результатів дослідження

Візуалізація результатів дослідження є важливим етапом аналітичного процесу, оскільки забезпечує інтуїтивне, наочне представлення складних даних

та підвищує якість інтерпретації висновків. У контексті аналізу геолокаційних даних і виявлення діяльності тіньового флоту візуальна аналітика відіграє особливо значущу роль, адже саме графічні інструменти дозволяють зрозуміти структуру руху суден, просторово-часові взаємозв'язки, аномалії та характерні патерни поведінки, які складно визначити лише на основі числового аналізу. Візуалізація перетворює великі масиви інформації на доступні та зрозумілі моделі, які сприяють прийняттю рішень і формуванню подальших дослідницьких стратегій.

Одним із ключових завдань візуалізації є відображення траєкторій руху суден у просторі. Карти, побудовані на основі AIS-даних, демонструють маршрути пересування, їхню щільність, характер зміни курсу та швидкості. За допомогою інтерактивних геопросторових платформ можна створювати багатошарові карти, які відображають як історичні, так і поточні маршрути, виділяють зони підвищеної активності, демонструють місця STS-перевантажень і структуру морського трафіку. Візуальні карти дозволяють швидко оцінити масштаб руху, визначити нетипові маршрути та виділити зони ризику, спричинені діяльністю тіньового флоту.

Візуалізація результатів дослідження включає також кластерний аналіз та графічні методи виділення аномальних патернів. Наприклад, теплові карти (heatmaps) можуть використовуватися для відображення концентрації суден або їхньої активності у певних просторових зонах. Діаграми щільності дозволяють виявити ключові транспортні коридори та відхилення від типових маршрутів. Кластеризація траєкторій, візуалізована у вигляді груп схожих маршрутів, допомагає ідентифікувати судна, що відхиляються від усталених моделей, що є одним із індикаторів тіньової діяльності.

Окреме значення має інтерактивна візуалізація, що дозволяє користувачам самостійно досліджувати дані: масштабувати карту, переглядати інформацію про конкретні сигнали, аналізувати часові проміжки або порівнювати рух кількох суден одночасно. Такий підхід дозволяє залучити експертів різних сфер,

підвищити точність виявлення аномалій і забезпечити гнучкість аналітичного процесу. Інтерактивні панелі на основі Dash, Tableau, Power BI або власних веб-додатків сприяють ефективній комунікації результатів між аналітиками, державними органами та міжнародними структурами.

Існує кілька чудових бібліотек, які можна використовувати для візуалізації геопросторових даних на карті — залежно від того, чи потрібні статичні чи інтерактивні карти, наскільки великі дані та скільки налаштувань потрібно.

**GeoPandas**

GeoPandas розширює можливості Pandas геопросторовими типами (точки, лінії, полігони) за допомогою бібліотек, таких як Shapely та Fiona.

GeoPandas слід використовувати якщо у нас вже є векторні геопросторові дані (наприклад, шейп-файли, GeoJSON, точкові дані) і ми хочемо виконати аналіз та швидку побудову графіків у середовищі Python.

Переваги:

- Легка інтеграція з робочими процесами Pandas.
- Хороша підтримка багатьох форматів (через Fiona/GDAL).
- Просте побудування геометрії за допомогою вбудованих методів .plot().

Недоліки:

- Здебільшого для статичних графіків (хоча можна інтегруватися з інтерактивними серверами).
- Для дуже великих наборів даних (наприклад, мільйонів точок) продуктивність може бути обмеженням.

**Folium**

Обгортка Python навколо бібліотеки JavaScript Leaflet.js. Вона легко створює інтерактивні карти з Python.

Підходить якщо потрібні інтерактивні веб-карти, які можна вбудовувати в блокноти, ділитися у вигляді HTML або відображати в браузері.

Переваги:

- Інтерактивність (масштабування, спливаючі вікна) одразу після встановлення.
- Добре підходить для представлення даних аудиторії незнайомій з особливостями технічної галузі.
- Легко накладати шари даних (маркери, теплові карти, полігони).

Недоліки:

- Для величезних наборів даних можуть виникати труднощі (обмеження браузера/інтерфейсу).
- Стилізація та розширена просторова аналітика можуть вимагати більше роботи.

**Plotly**

Загальна інтерактивна бібліотека візуалізації; підтримує карти (наприклад, scatter_geo, choropleth, інтеграції Mapbox).

Переваги:

- Дуже добре підходить для інформаційних панелей та веб-додатків.
- Легко підключається до веб-фреймворків.
- Інтерактивні підказки, гарний дизайн.

Недоліки:

- Деякі просторові/географічні операції можуть бути не такими багатими, як спеціальні GIS-бібліотеки.
- Якщо у вас є важкий геопросторовий аналіз (просторові з'єднання тощо), вам можуть знадобитися додаткові бібліотеки.

**Cartopy**

Бібліотека, побудована на Matplotlib, для розширеного картографічного креслення (картографічні проекції, берегові лінії тощо).

Переваги:

- Відмінний контроль над картографічними проєкціями та стилем.
- Чудово підходить для дослідницьких/академічних карт.

Недоліки:

- Менш пристосована для інтерактивних веб-карт.
- Крива навчання для деталей просторової проекції.

Візуалізація геолокаційних даних є одним з ключових елементів аналізу морських траєкторій та виявлення аномалій, притаманних тіньовому флоту.

У рамках кваліфікаційної роботи для побудови інтерактивних карт і відображення маршрутів суден було обрано бібліотеку Folium, яка поєднує гнучкість Python та можливості картографічного рушія Leaflet.js.

Folium є однією з найбільш зручних та гнучких Python-бібліотек для візуалізації геопросторових даних, зокрема маршрутів і точкових спостережень у сфері морської аналітики. Заснований на JavaScript-бібліотеці Leaflet, він дозволяє поєднувати потужність Python-екосистеми з інтерактивними можливостями веб-карт, що робить його ефективним інструментом для дослідження геолокаційних даних, включно з AIS-сигналами, траєкторіями суден та зонами їхньої активності. Folium забезпечує гнучкість у побудові карт і багатошарових моделей, не вимагаючи спеціальних навичок веб-розробки, що суттєво спрощує інтеграцію в аналітичні процеси.

Folium також добре підходить для роботи з результатами алгоритмічної обробки даних. Після кластеризації, сегментації траєкторій або виявлення аномалій результати можуть бути миттєво візуалізовані у вигляді ліній, точок або теплових карт. Це особливо корисно під час аналізу зон високої активності

суден, виявлення STS-операцій або пошуку патернів «темної активності». Folium підтримує можливість використання колірних шкал, що дозволяє інтуїтивно відображати характерні ознаки групування суден, інтенсивність трафіку або рівень ризику.

Значною перевагою Folium є проста інтеграція з середовищами наукових досліджень, такими як Jupyter Notebook або Google Colab, де мапи можуть відображатися безпосередньо у форматі HTML. Це дозволяє будувати інтерактивні звіти, що поєднують код, аналітичний текст і візуалізацію, або експортувати карти у форматі HTML для подальшого використання в презентаціях та веб-застосунках. Бібліотека також підтримує збереження карти у файлі, що робить її придатною для публікацій, автоматизованих звітів та інтеграції у більші програмні комплекси.

Обґрунтування вибору цієї технології подано нижче:

**1. Інтерактивність і наочність**

Folium забезпечує повністю інтерактивні карти, що дозволяє:

- збільшувати та переміщувати карту в реальному часі;
- переглядати спливаючі підказки та попапи з інформацією про точку;
- виділяти аномальні сегменти маршруту різними кольорами;
- працювати з великими наборами точок без втрати зручності сприйняття.

Це критично важливо для аналізу траєкторій суден, де потрібно швидко і наочно відстежувати зміни напрямку, швидкості, «темні» періоди та інші патерни.

**2. Інтеграція з Python та обробкою даних**

Folium нативно працює в екосистемі Python, що дозволяє:

- легко інтегрувати його з pandas, PySpark, NumPy;
- будувати карти на основі попередньо оброблених геолокаційних даних;
- автоматизувати візуалізацію великих наборів даних;

- створювати сотні карт у циклах без ручного втручання.

Оскільки у роботі використовується Apache Spark для обробки AIS-даних, можливість прямої інтеграції з Python є суттєвою перевагою.

**3. Підтримка геопросторових можливостей**

Folium дозволяє легко візуалізувати:

- лінійні маршрути (PolyLine);
- точки (Marker, CircleMarker);
- області та полігони;
- heatmap-и;
- кластеризацію точок.

Це дає змогу відображати:

- траєкторії суден у вигляді ліній;
- STS-операції (кластеризовані точки зближення);
- аномальні ділянки маршруту;
- місця зупинок та стоянок;
- зміни маршруту в часі.

Folium генерує карту у форматі HTML-файлу, який можна відкривати у браузері без серверної частини, передавати аналітикам та іншим дослідникам. Генерація статичних HTML-файлів дає незалежність від зовнішніх сервісів і дозволяє зберігати повну інтерактивність[19].

Folium дозволяє:

- підсвічувати сегменти маршруту різними кольорами (наприклад, червоним при темних періодах > 12 годин),
- додавати підписані попапи з координатами, часом та додатковою інформацією,
- вмикати або вимикати обгортання карти на міжнародній лінії зміни дат,
- створювати складні візуальні компоненти (легенди, шари, контури зон ризику).

Це важливо для візуального представлення аномалій у показниках руху тіньового флоту.

Folium добре підходить для візуалізації сотень тисяч точок, довгих суднових маршрутів, багаторівневих лінійних сегментів. Його використання оптимальне у поєднанні зі Spark, який забезпечує попередню агрегацію даних на кластері.

Загалом, Folium є універсальним та доступним інструментом для візуалізації геолокаційних даних, що поєднує у собі інтерактивність, простоту користування та широкі можливості інтеграції з Python-екосистемою. У дослідженнях морського трафіку та діяльності тіньового флоту він дозволяє наочно представляти складні аналітичні результати, забезпечувати інтуїтивне розуміння просторових залежностей та сприяти підвищенню точності та обґрунтованості аналітичних висновків.

# РОЗДІЛ 3
# РЕЗУЛЬТАТИ ПРОВЕДЕНОГО АНАЛІТИЧНОГО ДОСЛІДЖЕННЯ

## 3.1. Пошук та візуалізація переміщення підсанкційних суден

Як було згадано в попередніх розділах, для реалізації системи виявлення тіньового флоту було розроблено комплексне програмне рішення, яке поєднує високопродуктивні інструменти обробки великих даних та сучасні технології геопросторового аналізу. Система складається з таких ключових компонентів:

- MinIO — об'єктне сховище, у якому зберігаються історичні дані AIS у форматах CSV/Parquet;
- Apache Spark — обчислювальний двигун для розподіленої обробки великих обсягів даних;
- Apache Sedona — надбудова над Spark для виконання геопросторових операцій;
- Folium — інструмент для візуалізації геопросторових траєкторій на інтерактивних HTML-картах.

Використання саме цієї архітектури дозволило забезпечити масштабованість, точність геолокаційних обчислень та інтерактивність кінцевого результату.

AIS-дані зберігаються в MinIO — високопродуктивному об'єктному сховищі, сумісному з Amazon S3 API.

Таке рішення обрано з кількох причин: можливість зберігання великих обсягів неструктурованих файлів (мільйони сигналів); підтримка S3-протоколу, що дозволяє Spark напряму читати дані; висока швидкість читання завдяки розподіленому зберіганню; можливість масштабування у хмарних або on-premise середовищах.

Для дослідження був сформований датасет AIS сигналів за період з 01.08.2025 по 10.10.2025. Дані зберігаються у форматі CSV (формат файлу, що

зберігає табличні дані у текстовому вигляді, де значення в кожному рядку розділені комами), початковий об’єм даних сягав понад 110 Гігабайт тому в цілях оптимізації датасет був стиснутий за допомогою утиліти gzip, яка використовує алгоритм стиснення Deflate для зменшення розміру файлу, що дозволило зменшити розмір до 11.2 Гігабайт.

Дані структуровано за паттерном:

s3a://datalake/vessel_signals/{year}/{month}/{day}/

Така схема дозволяє виконувати часові фільтри на стороні файлової системи, зменшуючи обсяг завантажених даних.

Кількість наявних сигналів для одного дня – від 10 до 12 мільйонів, тому їх загальна кількість у використаному датасеті сягає більш ніж 750 мільйонів AIS сигналів.

Також для виявлення суден, на які були накладені санкції було сформовано датасет із інформацією про назву судна, дату накладення санкцій, їх IMO (унікальний ідентифікатор судна) та MMSI (унікальний ідентифікаційний номер, який використовується для ідентифікації морських суден, станцій та інших мобільних морських об'єктів у системі морського зв'язку).

Після підключення до MinIO дані завантажуються в Spark DataFrame. Основними задачами Spark є попередня очистка даних (кастинг колонок, нормалізація timestamp), фільтрація за MMSI або IMO, сортування сигналів у часовому порядку (для відображення шляху руху), попередня агрегація даних для побудови траєкторій.

```
signals_df = (
    spark.read
    .option("header", "true")

.csv("s3a://datalake/vessel_signals/*/*/*.csv.gz")
    .withColumn("mmsi", col("mmsi").cast("long"))
```

```
    .withColumn("latitude",
col("latitude").cast(DoubleType()))
    .withColumn("longitude",
col("longitude").cast(DoubleType()))
    .withColumn("timestamp",
to_timestamp(from_unixtime(col("timestamp"))))
)
```

Лістинг 3.1 – Приклад завантаження даних в Spark DataFrame

Spark виконує розподілене обчислення, що дозволяє обробляти десятки мільйонів записів за декілька хвилин. Sedona надає змогу виконувати геолокаційні обчислення на рівні кластеру Spark, що критично важливо при аналізі великих об’ємів сигналів AIS.

Після обробки у Spark дані конвертуються в Pandas DataFrame для передачі у Folium.

```
pdf = vessel_signals.select("timestamp", "latitude",
"longitude").toPandas()
pdf["timestamp"] = pd.to_datetime(pdf["timestamp"],
utc=True)
```

Лістинг 3.2 – Приклад формування Pandas DataFrame

Цей формат оптимальний для поелементного проходження маршруту.

Після фільтрації сигналів і отримання шляхів лише суден під санкціями, для кожного з них формується інтерактивна карта для візуального аналізу та визначення подальшої можливості дослідження кожного із знайдених суден.

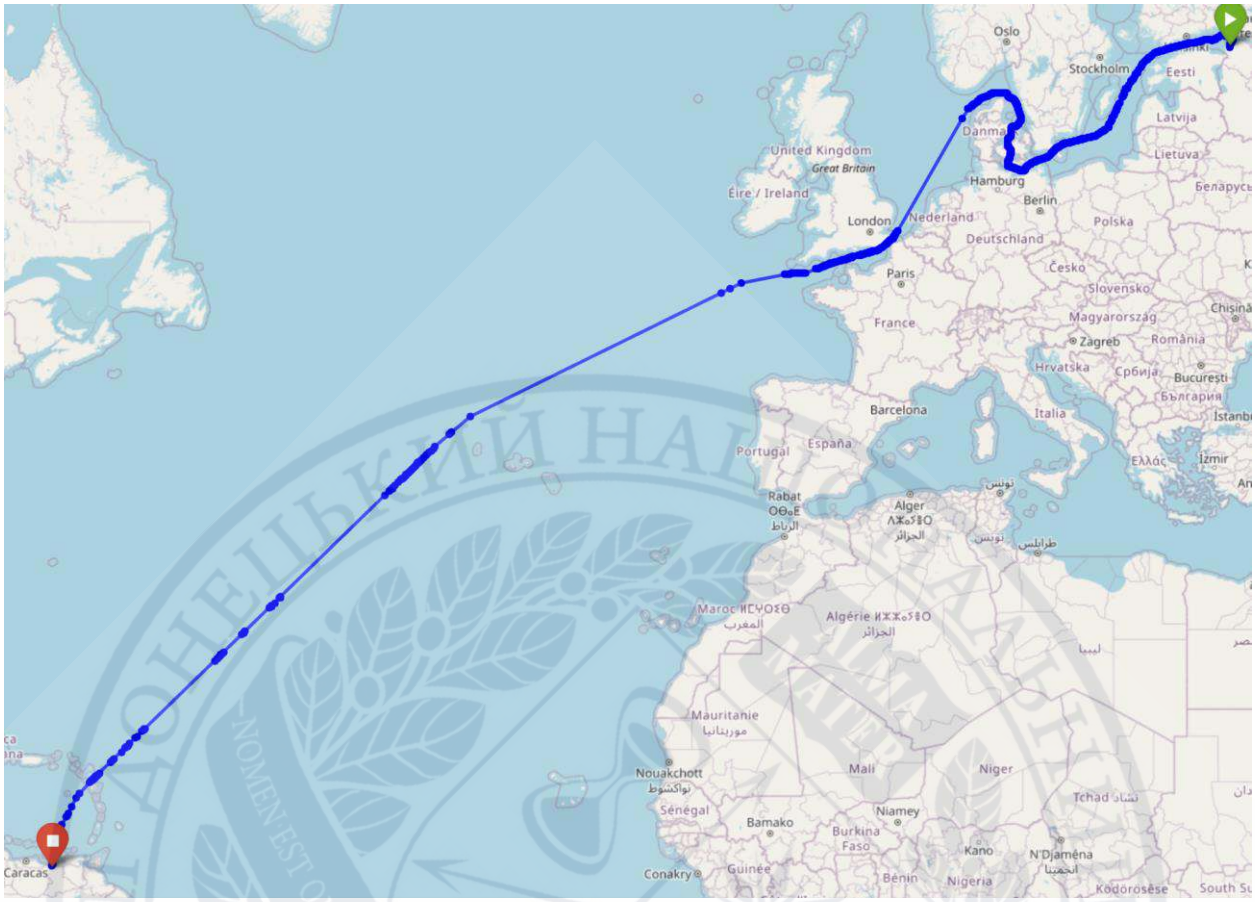


Рисунок 3.1 – Візуалізація шляху судна з MMSI – 461000249

На рисунку 3.1 продемонстрований шлях судна під назвою “Lavender” (IMO: 9339337, MMSI: 461000249) – нафтовий танкер зареєстрований під прапором Оману та перебуває під санкціями з 25.06.2024.

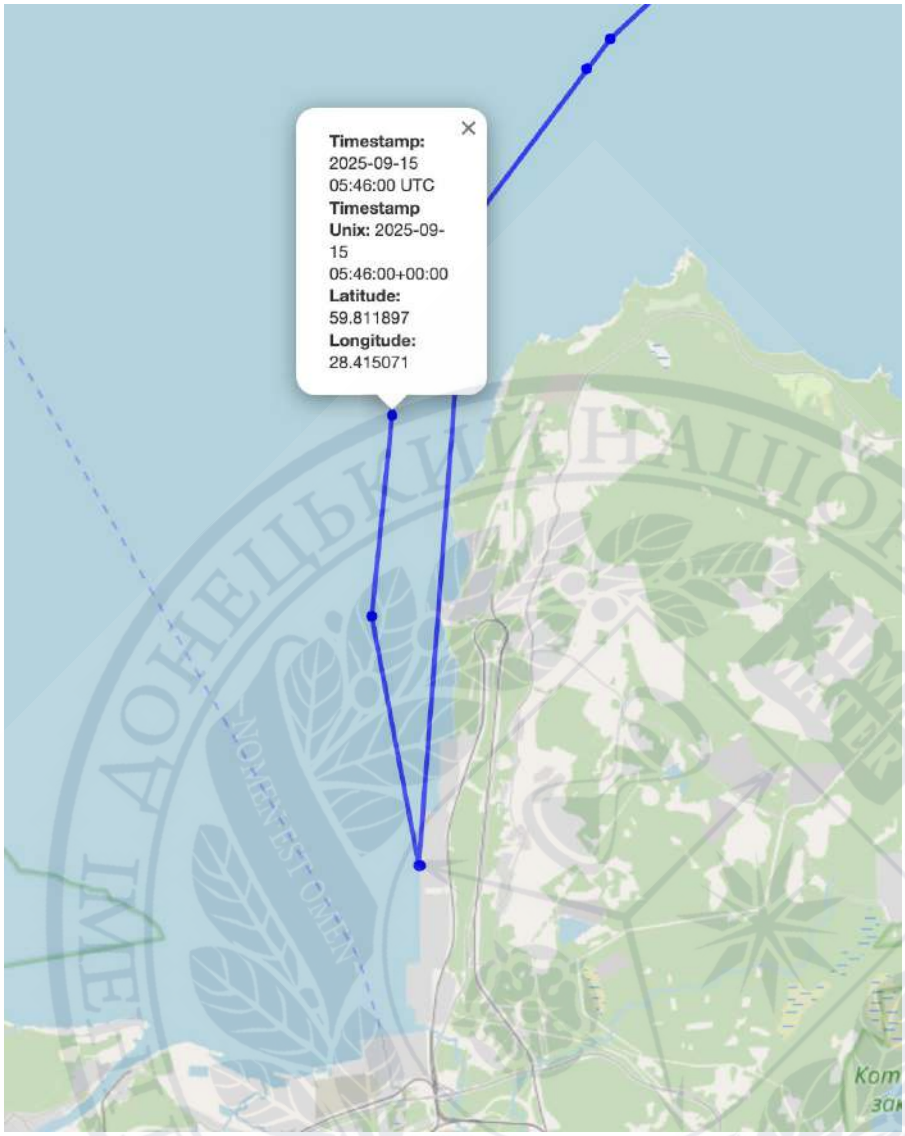


Рисунок 3.2 – Початок шляху судна з MMSI – 461000249

При детальному перегляді початкових сигналів побудованого шляху було визначено що 15.09.2025 судно перебувало біля російського порту Усть-Луга, який є одним з великих нафтових терміналів.

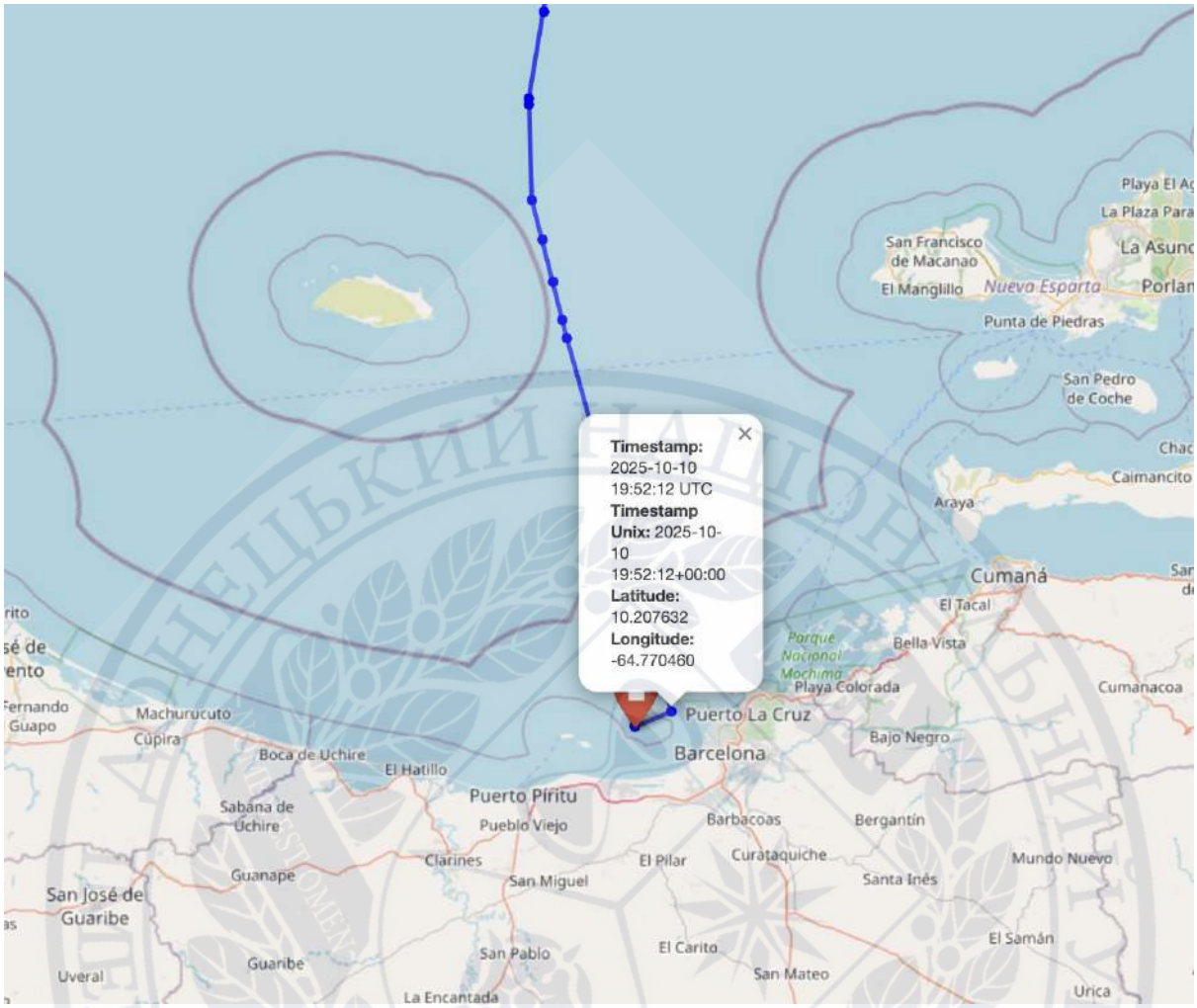


Рисунок 3.3 – Кінець шляху судна з MMSI – 461000249

Кінцева точка шляху датована 10.10.2025 і знаходиться біля берегів міста Барселона у Венесуелі. У цьому місті знаходиться морський порт в якому могло бути проведено нелегальне відвантаження російських нафтопродуктів.

Факт того що судно зареєстроване під прапором Оману свідчить про використання однієї із тактик обходу санкційної політики – реєстрація у слабо контрольованих юрисдикціях. Судна реєструють у країнах з низьким наглядом, які рідко проводять інспекції, швидко змінюють статуси суден, дозволяють не подавати обов'язкові документи.

### 3.2. Дослідження фактів вимкнення AIS-трансляції

Вимкнення AIS-передавача — один із найбільш характерних і критичних методів, який використовує тіньовий флот для приховування своєї діяльності. Автоматична ідентифікаційна система (AIS) зобов’язує судна постійно транслювати інформацію про:

- свої координати (широту, довготу),
- курс та швидкість,
- ідентифікатори (MMSI, IMO, назву),
- статус навігації,
- пункт призначення та ETA.

Коли судно навмисно припиняє трансляцію цих даних, воно стає “невидимим” для систем моніторингу. Таке явище отримало назву Dark Activity — “темна активність”.

Відповідно до:

- Конвенції SOLAS (Chapter V, Regulation 19),
- вимог Міжнародної морської організації (IMO),
- правил морської безпеки,

AIS має працювати постійно, за винятком:

- ситуацій, коли це може створити загрозу безпеці,
- випадків пошкодження обладнання.

Таким чином, навмисне вимкнення AIS заради приховування регулярної діяльності є порушенням міжнародних норм.

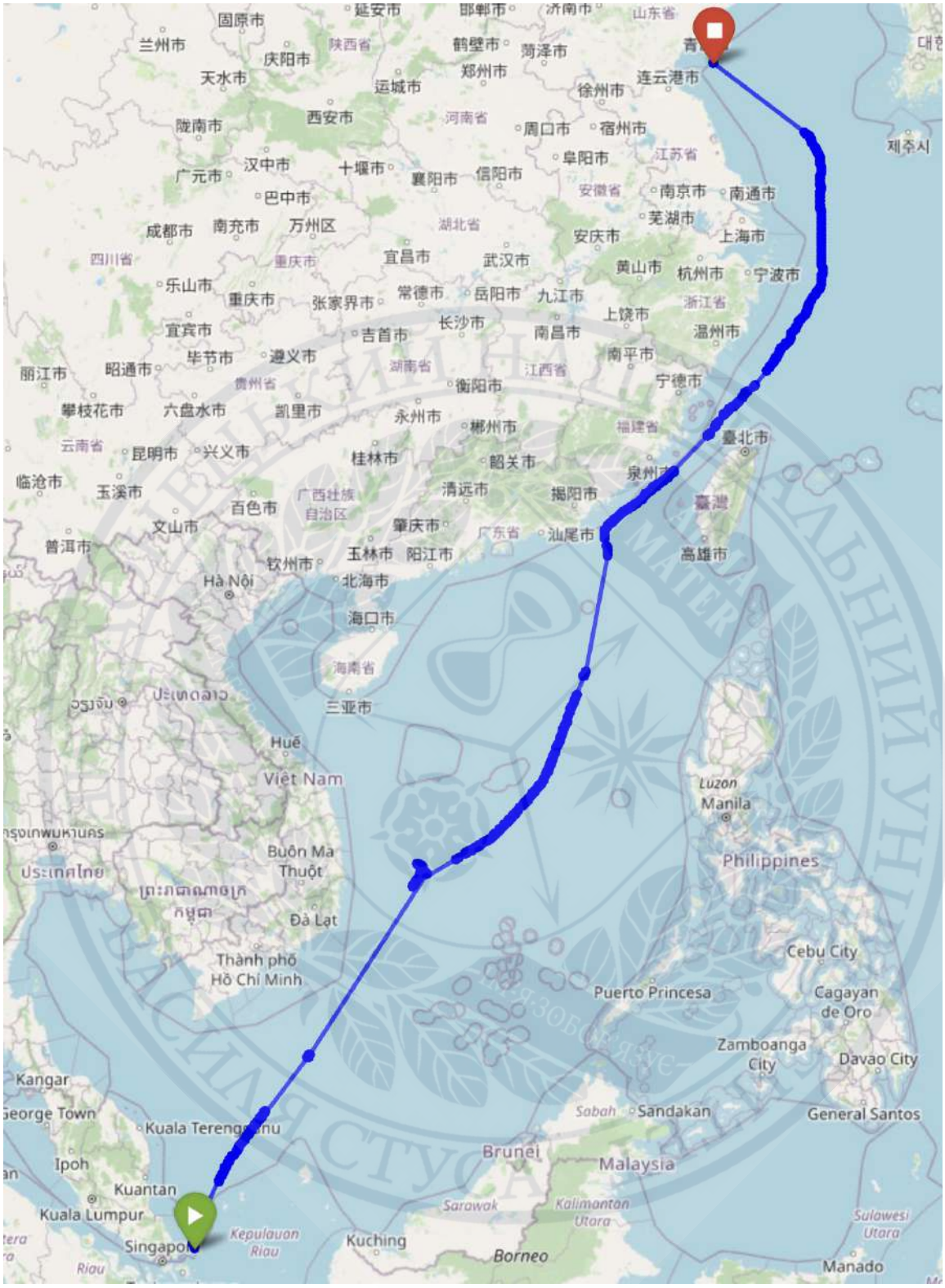


Рисунок 3.4 – Візуалізація шляху судна з MMSI – 613003728

На рисунку 3.4 продемонстрований шлях судна під назвою “Bodhi” (IMO: 9144782, MMSI: 613003728) – нафтовий танкер зареєстрований під прапором

Камеруну який перебуває під санкціями з 25.06.2024. З відкритих джерел було знайдено інформацію про те що це судно було неодноразово помічено в російському порту Усть-Луга, тому є підстави вважати що воно використовується для транспортування російських нафтопродуктів і є частиною так званого “тіньового” флоту.

Після візуального аналізу стає помітно що в побудованому шляху є декілька суттєвих проміжків часу за які AIS дані відсутні. В деяких випадках це можна вважати нормою, наприклад якщо судно перебуває у віддаленні від радіолокаційних засобів які могли б зареєструвати його сигнал, але велику підозру викликає відсутність сигналів наприкінці шляху, що є аномальним через досить стабільну кількість сигналів напередодні, а також враховуючи що судно перебуває в тому ж регіоні що і раніше.

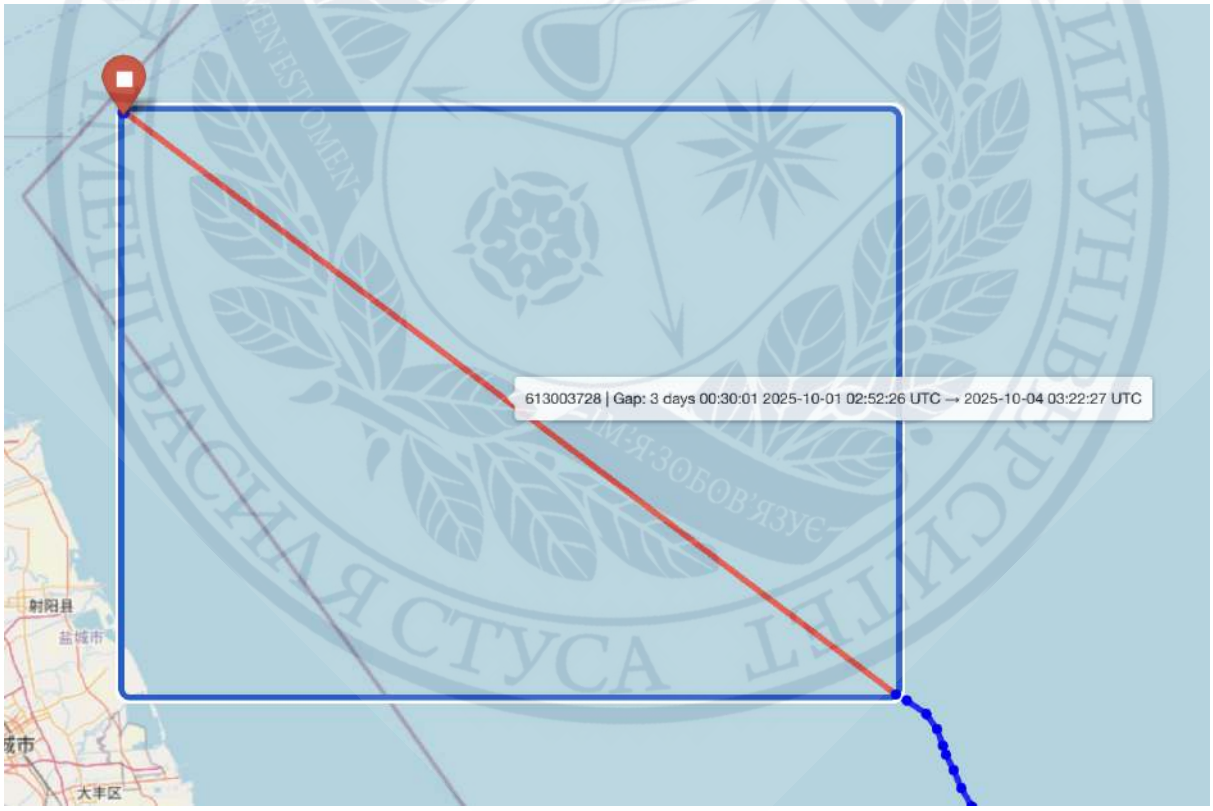


Рисунок 3.5 – Візуалізація аномальної відсутності сигналів судна з MMSI – 613003728

Після проведення додаткового аналізу було визначено що після наближення до берегів Китаю судно не передавало свій сигнал на протязі більш ніж трьох днів, що може свідчити про відключення AIS передавача в цілях приховання факту входження судна в один із портів для відвантаження нафтопродуктів.

Це може свідчити про приховування походження нафти – судно “зникає”, заходить у санкційний порт, відвантажує нафту і з’являється за сотні кілометрів.

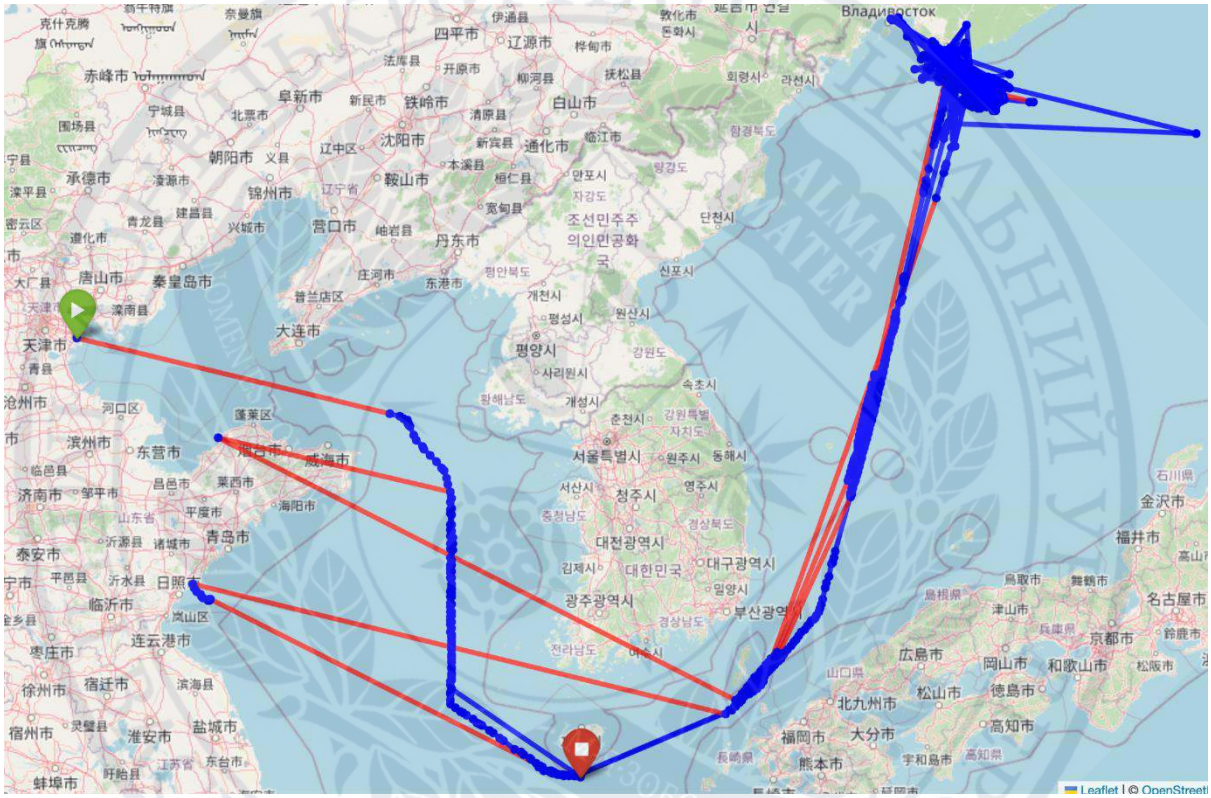


Рисунок 3.5 – Візуалізація аномалій сигналів судна з MMSI – 461000238

Наступний приклад аномальної поведінки – судно під назвою “AMULET” (IMO: 9413547, MMSI: 461000238) – нафтовий танкер зареєстрований під прапором Оману який перебуває під санкціями з 17.12.2024. Перевірка відкритих джерел вказує на те що судно змінювало свою реєстрацію на назву 10 разів

протягом 2024 та 2025 року, що є ще одним із ключових інструментів тіньового флоту для обходу санкцій.

Однією з найбільш поширених та ефективних технік, яку використовує тіньовий флот, є часта зміна формальних власників, операторів, менеджерів та технічних керуючих судном. Така практика створює складну мережу юридичних структур, які приховують реальних бенефіціарів та унеможливлюють просте відстеження санкційних зв’язків.

У поєднанні з іншими методами, такими як зміна прапора та вимкнення AIS, вона значно ускладнює моніторинг судноплавства.

Судно може мати кілька типів пов’язаних сторін:

- власник судна (Registered Owner)
- оператор або фрахтувальник (Operator / Charterer)
- менеджер судна (Ship Manager / ISM Manager)
- компанія, що займається технічним обслуговуванням (Technical Manager)
- компанія, що комплектує екіпаж (Crewing Manager)

У тіньовому флоті всі ці ролі можуть змінюватися надзвичайно часто — іноді кожні кілька місяців або навіть кілька тижнів. Це створює юридичну “туманність”, у якій складно визначити остаточного бенефіціара, важко підтвердити, що судно належить до санкційної мережі, неможливо встановити відповідальність у разі аварії або порушення.

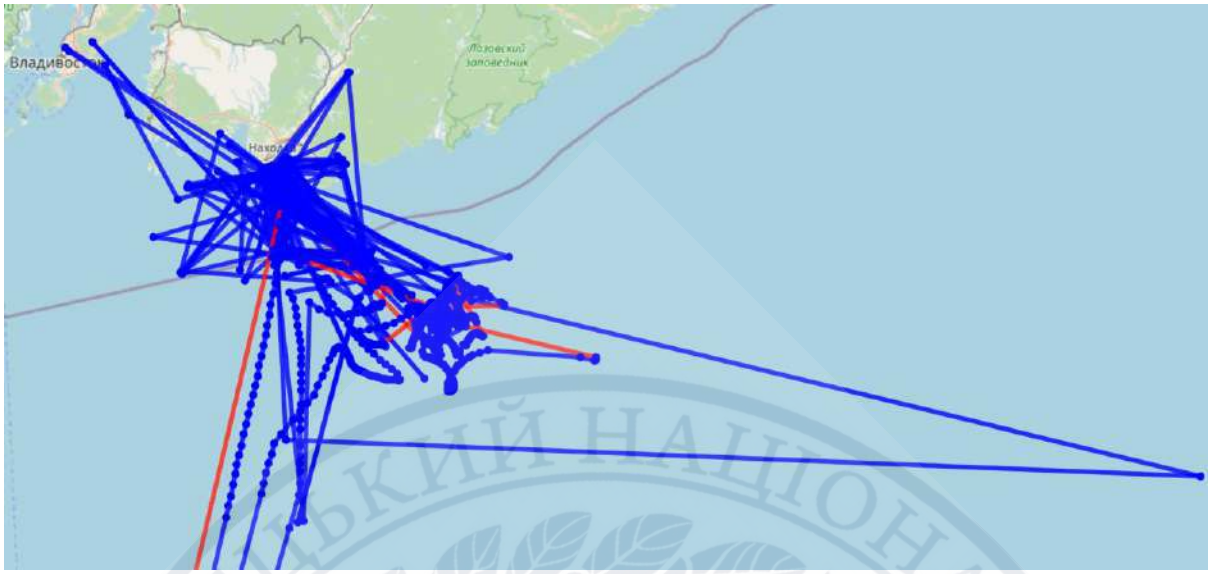

Рисунок 3.6 – Активність судна “AMULET” біля порту міста Находка

Попередній огляд маршрутів показав неодноразову активність судна поблизу нафтового порту в російському місті Находка. AIS сигнали знайдені біля даного порту мають багато аномалій в координатах сигналів, що не є характерним для більшості сигналів судна в інших локаціях. Це може свідчити про GPS-спуфінг —навмисне впровадження фальшивих геолокаційних даних у навігаційні системи судна або маніпуляція AIS-повідомленнями так, щоб вони містили вигадані координати. На відміну від вимкнення AIS, при GPS-спуфінгу судно не зникає з радарів, а продовжує передавати дані, які вводять аналітичні системи в оману.

Тіньовий флот використовує цей метод там, де не можна дозволити собі “темні періоди”, але потрібно приховати реальну активність — наприклад, STS-операції, заходи в санкційні порти або нелегальні зупинки.

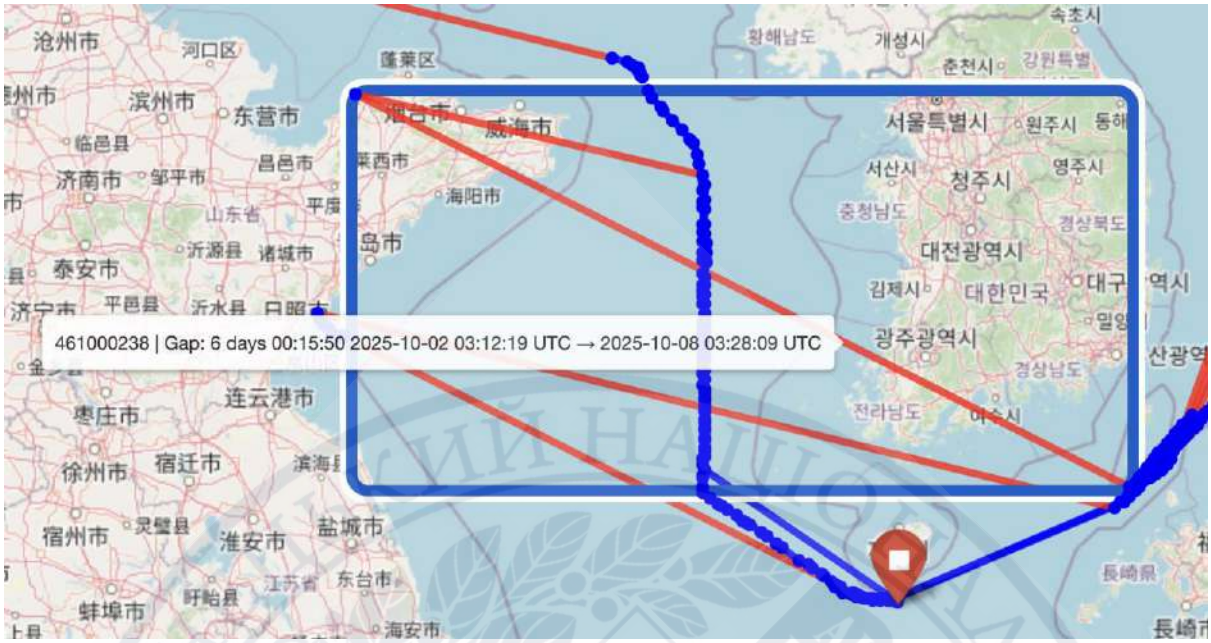


Рисунок 3.7 – Візуалізація аномальної відсутності сигналів судна з MMSI – 461000238

Серед маршрутів декількох маршрутів судна було виявлено найбільш підозрілий – на шляху, який це судно вже раніше проходило, помітна аномальна відсутність сигналів зображена на рисунку 3.7. Різниця між сигналами сягнула 6 днів, а якщо врахувати що після єдиного знайденого сигналу після такого проміжку з’являється ще один проміжок тривалістю 1 день – це свідчить про навмисне відключення AIS передавача. Ці дії могли бути вчинені в цілях приховання факту входження судна в один із портів для відвантаження нафтопродуктів або ж з ціллю перевантаження вантажу з одного судна на інше без заходу до порту. Зазвичай цей процес є легальним і широко використовується в міжнародній торгівлі, наприклад, для перевантаження нафти, LNG або хімічної продукції у районах з шару води, недостатньої для глибоководних танкерів.

Однак подібні операції стали ключовим інструментом тіньового флоту, який використовує їх для прихованого транспортування санкційних нафтопродуктів, змішування вантажів та маскування їх походження.

В межах дослідження не вдалось підтвердити потенційне проведення подібної операції через занадто високу активність суден в даному регіоні.

### 3.3. Дослідження потенційної передачі вантажу у відкритому морі (Ship-to-Ship)

STS-операція (*Ship-to-Ship*) — це процес перевантаження вантажу з одного судна на інше без заходу до порту. Санкційна нафта перевантажується на інший танкер, який має “чисту” історію, зареєстрований під іншим прапором, має іншу юридичну компанію.

Після цього вантаж легше продавати легальним трейдерам.


Рисунок 3.8 – Аномальні маршрути руху судна з MMSI – 273257120

Серед знайдених під час дослідження суден було помічено судно “MELITO CARRIER” (IMO: 8920581, MMSI: 273257120) зареєстрований під прапором Оману який перебуває під санкціями з 24.10.2025. На рисунку 3.8 можна побачити нестандартні шляхи руху судна, а також випадки тривалої відсутності сигналів, що може свідчити про незаконні дії у відкритому морі.

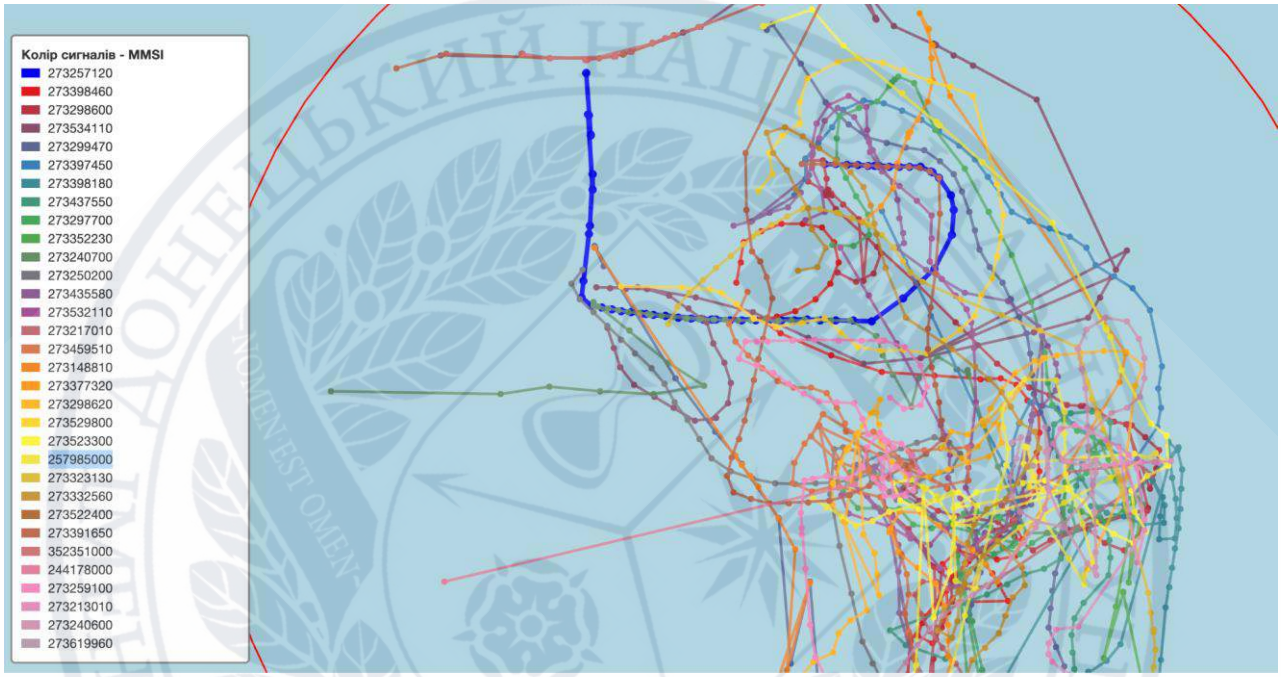


Рисунок 3.9 – Візуалізація суден в радіусі 75 кілометрів від судна з MMSI – 273257120

Для подальшого пошуку взаємодії у відкритому морі було знайдено і візуалізовано сигнали всіх суден для були в радіусі 75 кілометрів від судна “MELITO CARRIER” за аналогічний період часу.

Для зручності пошуку кожному судну був присвоєний унікальний колір сигналів та відображено відповідність кольору до їх MMSI.

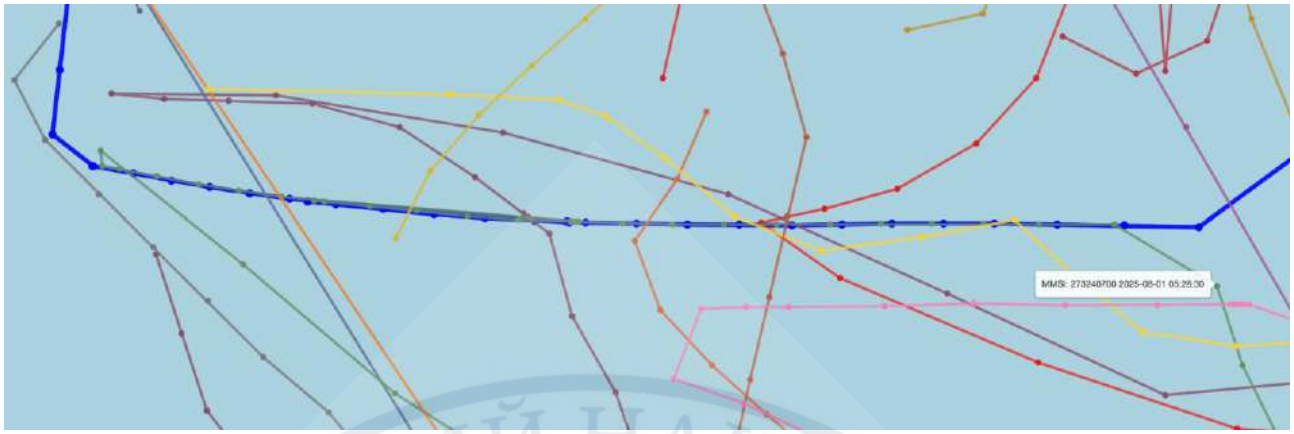

Рисунок 3.10 – Візуалізація наближення судна з MMSI – 273240700

Серед усіх суден які перебували в даному регіоні в даний період часу найбільшу підозру викликає судно під назвою “NIVENSKOYE” з MMSI: 273240700.

Протягом шести годин обидва кораблі рухались в одному напрямку та в безпосередній близькості один від одного, що вказує на класичні ознаки Ship-to-Ship операції – зближення суден на менш ніж 500 метрів, перебування в близькості більш ніж 6 годин, рух після зближення відрізняється за логікою (судно яке тривалий час перебувало переважно в одному регіоні після виконання операції одразу прямує в порт).

# ВИСНОВКИ

Основним результатом роботи стало комплексне дослідження проблеми виявлення тіньового флоту на основі аналізу геолокаційних даних суден, зокрема даних автоматичної ідентифікаційної системи (AIS). Враховуючи масштаб проблеми, глобальний вплив на безпеку морських перевезень та важливість дотримання міжнародних санкцій, проведене дослідження дозволило отримати низку важливих практичних і теоретичних результатів.

Спочатку було проаналізовано сутність явища тіньового флоту, визначено основні форми та методи його діяльності, такі як вимкнення AIS-трансляції, GPS-спуфінг, проведення STS-операцій у відкритому морі, маніпуляції з MMSI та IMO, зміна менеджерів і власників суден, використання слабко регульованих прапорів, застосування офшорних корпоративних структур тощо. Доведено, що ці методи формують комплексну систему ухилення від регуляторних обмежень і значно ускладнюють контроль за переміщенням вантажів, особливо енергетичних ресурсів.

Також було описано ризики, які несе діяльність тіньового флоту для міжнародної безпеки та економіки: зростання кількості аварій через використання технічно зношених танкерів, підвищення рівня екологічних загроз, розширення нелегальних логістичних ланцюгів, підтримка порушення санкційних режимів та фінансування небажаних акторів. Обґрунтовано, що моніторинг тіньового флоту є важливим елементом сучасної системи глобальної морської безпеки.

Далі у роботі проведено огляд технологій та методів аналізу AIS-даних, визначено їх роль у виявленні прихованої активності суден. Показано, що геолокаційні дані дозволяють ідентифікувати часові та просторові аномалії, такі як “темні періоди”, фальсифіковані траєкторії, нелогічні зміни курсу, STS-зближення та нетипові стоянки. Встановлено, що саме аналіз геопросторових

патернів є одним із найефективніших методів розпізнавання підозрілої поведінки суден.

Наступних етапом було виконано практичну реалізацію підсистеми аналізу та візуалізації даних. Для обробки великих масивів інформації було використано Apache Spark, що забезпечило масштабованість та можливість ефективного опрацювання десятків мільйонів записів. Для виконання геопросторових обчислень застосовано Apache Sedona, що дало змогу виконувати операції просторового фільтрування, побудови буферів, визначення відстаней та взаємного розташування об’єктів. Сховище MinIO забезпечило швидкий доступ до CSV та Parquet-даних у хмарному середовищі. Для візуалізації траєкторій та аномалій був використаний інструмент folium, який дозволив створювати інформативні інтерактивні карти з накладанням даних, часовими маркерами, маршрутами та аномальними відрізками.

У роботі реалізовано практичні засоби виявлення ключових ознак тіньової поведінки, таких як:

- виявлення часових розривів між сигналами (більш ніж 12 годин);
- ідентифікація GPS-спуфінгу через нереалістичні швидкості та різкі переміщення;
- пошук STS-зближень за геометричними критеріями;
- визначення нелогічних маршрутів та стоянок у відкритому морі;
- аналіз довготривалих “темних періодів” поблизу санкційних зон.

Отримані результати підтверджують ефективність використання геолокаційних даних для виявлення тіньового флоту та демонструють можливість автоматизації цього процесу засобами великих даних і геопросторової аналітики.

Загалом, виконана робота показала, що поєднання сучасних технологій обробки Big Data, просторових алгоритмів та інтерактивної візуалізації створює потужний інструментарій для моніторингу морських перевезень і детекції нелегальної активності. Розроблений підхід може бути розширений,

інтегрований у системи державних органів, аналітичних центрів або комерційних платформ та слугувати основою для подальших досліджень у галузі санкційного комплаєнсу та морської безпеки.

# СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Тіньовий флот РФ [Електронний ресурс] // Deutsche Welle – Режим доступу: https://www.dw.com/uk/tinovij-flot-rf-ci-vdastsa-zahodu-znekroviti-rezim-putina/a-70925560

2. Тіньовий флот російської нафти: статус і план дезактивації [Електронний ресурс] // Економічна правда – Режим доступу: https://epravda.com.ua/experts/shcho-take-tinoviy-flot-rosiji-i-yak-rosiya-eksportuye-naytu-popri-sankciji-800367

3. Тіньовий флот Росії: чи існують кораблі-привиди, чи це лише зручний фейк для прикриття поставок нафти [Електронний ресурс] // Еспресо – Режим доступу: https://espreso.tv/poyasnuemo-tinoviy-flot-rosii-chi-isnuyut-korabli-privid i-chi-tse-lishe-zruchniy-feyk-dlya-prikrittya-postavok-nafti

4. Ying Yang, Yang Liu, Guorong Li, Zekun Zhang, Yanbin Liu. Harnessing the power of Machine learning for AIS Data-Driven maritime Research: A comprehensive review, 2024 – С. 2-6.

5. Holden Karau, Rachel Warren. High Performance Spark: Best Practices for Scaling and Optimizing Apache Spark, 2017 – С. 56-58.Holden Karau, Rachel Warren. *High Performance Spark*, 2017 – С. 56–58.

6. Arsalan Javed, Ubaid Rahman. *Maritime Traffic Anomaly Detection Using AIS Data*, 2022 – С. 13–21.

7. Veselov G., Kharitonov D. *AIS Data Processing and Trajectory Analysis*, 2021 – С. 47–55.

8. International Maritime Organization (IMO) [Електронний ресурс] – Режим доступу: https://www.imo.org

9. EMSA — *Dark Fleet and Maritime Safety Report*, 2023 [Електронний ресурс] – Режим доступу: https://emsa.europa.eu

10. UN Panel on Maritime Sanctions — *Ship-to-Ship Transfers Report*, 2023 [Електронний ресурс] – Режим доступу: https://main.un.org/securitycouncil/en/sanctions/1718/panel_experts/reports

11. NOAA Marine Navigation Services [Електронний ресурс] — Режим доступу: https://nauticalcharts.noaa.gov

12. MarineTraffic — AIS Technical Information [Електронний ресурс] – Режим доступу: https://www.marinetraffic.com

13. Global Fishing Watch — AIS Dataset Specifications [Електронний ресурс] – Режим доступу: https://globalfishingwatch.org

14. SkyTruth — Bringing Russia's Shadow Fleet Into the Light Using Satellite Data, 2023 [Електронний ресурс] – Режим доступу: https://skytruth.org/section/bringing-russias-shadow-fleet-into-the-light-using-satellite-data

15. Robert Johnson, Apache Sedona Essentials: A Practical Guide to Spatial Data Processing, 2025 – С.141-146.

16. Paweł Tokaj, Jia Yu, Mo Sarwat, Cloud Native Geospatial Analytics with Apache Sedona, 2025 – С. 16-25.

17. Pallotta G., Vespe M. Vessel Pattern Knowledge Discovery from AIS Data // *Information Fusion*, 2016.

18. Roy, J. Anomaly Detection in the Maritime Domain. In Proceedings of SPIE: Optics and Photonics in Global Homeland Security IV, Orlando, FL, USA, 16 March 2018; Halvorson, C.S., Lehrfeld, D., Saito, T., Eds.; Volume 6945.

19. Lloyd's List Intelligence. Dark Fleet Analysis Report, 2023 [Електронний ресурс] – Режим доступу: https://lloydslist.maritimeintelligence.informa.com

20. Apache Spark Documentation [Електронний ресурс] – Режим доступу: https://spark.apache.org/docs/latest

21. Apache Sedona Documentation [Електронний ресурс] – Режим доступу: https://sedona.apache.org

22. MinIO Documentation [Електронний ресурс] – Режим доступу: https://min.io/docs

23. Folium Documentation [Електронний ресурс] – Режим доступу: https://python-visualization.github.io/folium

24. Viktor Majer Shenberger, Big Data: A Revolution That Will Transform How We Live, Work and Think, 2025 – С.23-24.

25. Що таке "тіньовий флот" і як Росія використовує його проти Заходу [Електронний ресурс] // BBC – Режим доступу: https://www.bbc.com/ukrainian/articles/c4gw87480wgo

26. Тіньовий флот [Електронний ресурс] // ГУР МО України – Режим доступу: https://war-sanctions.gur.gov.ua/transport/shadow-fleet

27. NYT: близько 17% усіх діючих нафтових танкерів у світі належать до "тіньового флоту" РФ [Електронний ресурс] // LB.ua – Режим доступу: https://lb.ua/world/2025/09/22/697688_nyt_blizko_17_usih_diyuchih_naftovih.html

28. Python Folium: Create Web Maps From Your Data [Електронний ресурс] – Режим доступу: https://realpython.com/python-folium-web-maps-from-data/

29. Interactive map with Python and Folium [Електронний ресурс] – Режим доступу: https://python-graph-gallery.com/288-map-background-with-folium/

30. Building a Local Data Lake from scratch with MinIO, Iceberg, Spark, StarRocks, Mage, and Docker [Електронний ресурс] – Режим доступу: https://blog.dataengineerthings.org/building-a-local-data-lake-from-scratch-with-minio-iceberg-spark-starrocks-mage-and-docker-c12436e6ff9d

31. MinIO Object Storage Docs – https://min.io/docs

32. Amazon S3 Storage Guide [Електронний ресурс] – Режим доступу: https://aws.amazon.com/s3

33. Kepler.gl Documentation [Електронний ресурс] – Режим доступу: https://kepler.gl

34. Deck.gl Documentation [Електронний ресурс] – Режим доступу: https://deck.gl

35. Folium Documentation [Електронний ресурс] – Режим доступу: https://python-visualization.github.io/folium

36. GeoPandas Documentation [Електронний ресурс] – Режим доступу: https://geopandas.org

37. Brown, A. Detecting AIS Spoofing with Spatial Analysis // Maritime Cybersecurity Journal. – 2022.

38. Richards, P. AIS Data Integrity and Maritime Risk Assessment // Maritime Risk Review. – 2021.

39. The Maritime Executive. Shadow Fleet Investigations [Електронний ресурс]. – Режим доступу: https://www.maritime-executive.com

40. Financial Times. Russia's Oil Sanctions Evasion Tactics [Електронний ресурс]. – Режим доступу: https://www.ft.com/content/89d9bc79-dba6-4ec2-8bc3-612b68ebf9c9

41. Al Jazeera. Ghost Tankers and Oil Laundering Routes [Електронний ресурс]. – Режим доступу: https://www.aljazeera.com/economy/2025/11/27/russias-shadow-vessels-using-false-flags-to-skirt-sanctions-report-says

42. UNODC. Maritime Crime Patterns and Detection. – 2022.

43. International Energy Agency. Oil Market Report. – 2023.

44. PostGIS. Reference Manual [Електронний ресурс]. – Режим доступу: https://postgis.net

45. QGIS Documentation [Електронний ресурс]. – Режим доступу: https://qgis.org.

46. Mapbox GL JS Documentation [Електронний ресурс]. – Режим доступу: https://docs.mapbox.com/mapbox-gl-js

47. Columbia SIPA. Sanctions and Maritime Supply Chains. – 2022.

48. RAND Corporation. Illicit Maritime Logistics Networks. – 2020.

49. NATO Maritime Command. Maritime Risk Intelligence Brief, 2023 [Електронний ресурс]. – Режим доступу: https://ccdcoe.org/news/2025/nato-ccdcoe-publishes-new-policy-brief-on-cyber-threats-to-maritime-port-infrastructure

50. Chan, K., Wong, P. Spatial-Temporal Modelling of Vessel Behavior // Navigation Analytics Journal. – 2021.

51. Singh, R. Optimization of Big Data Processing in Maritime Systems // Big Data Engineering. – 2020.

52. Zhang, W. Trajectory Reconstruction and Noise Filtering in AIS Data // Signal Processing in Navigation. – 2019.

53. The Economist. Ghost Tankers and Global Trade [Електронний ресурс] – Режим доступу:

https://www.economist.com/technology-quarterly/2014/03/06/ghost-ships