ВІДГУК
ОФІЦІЙНОГО ОПОНЕНТА

на дисертаційну роботу аспіранта Грицука Валерія Юрійовича «Методи та моделі інформаційних технологій для моніторингу технічного стану транспортних засобів» на здобуття ступеня доктора філософії за спеціальністю 122 – Комп'ютерні науки

Актуальність теми дослідження

Сучасний етап розвитку автомобільного транспорту характеризується глибинною трансформацією, яку супроводжує перехід до концепції інтелектуальних транспортних засобів, що функціонують як повноцінні вузли в екосистемі Інтернету речей та які генерують значні обсяги телеметричної інформації у режимі реального часу. Умови інтенсивної експлуатації вимагають безперервного моніторингу фактичного технічного стану вузлів і агрегатів. Традиційні парадигми обслуговування – реактивна та планово-профілактична – значною мірою вичерпали свій потенціал через високі економічні втрати від непередбачуваних простоїв і неповне використання ресурсів деталей. Законною еволюційною альтернативою є проактивне прогнозне обслуговування, яке базується на аналізі потокових даних від бортових систем для точного прогнозування залишкового ресурсу механізмів.

Транспортні засоби оснащені розгалуженими мережами електронних блоків керування, об'єднаних гетерогенними протоколами в яких збір телеметрії традиційно здійснюється через стандартизований інтерфейс OBD-II. Проте існуючі діагностичні архітектури, побудовані на моделі поллінгу («запит-відповідь»), не забезпечують необхідної швидкодії (обмежуючись у середньому 9 запитами на секунду), що ускладнює достовірний моніторинг швидкоплинних газо- і термодинамічних процесів. Отже, виникає актуальне науково-прикладне завдання розробки розподілених інформаційно-

керувальних систем, здатних обробляти високошвидкісні потоки даних у режимі реального часу.

Впровадження таких розподілених інформаційних технологій має не лише локальний економічний ефект для СТО, але й виступає інфраструктурним базисом для розвитку концепції розумного міста.

Тому тема дисертаційної роботи відповідає сучасним тенденціям розвитку систем підтримки прийняття рішень, а актуальність є належно обґрунтованою.

Зв’язок роботи з науковими програмами, планами, темами

Дисертаційна робота виконана відповідно до тематичних планів наукових досліджень кафедри комп'ютерних наук та інформаційних систем Харківського національного автомобільно-дорожнього університету Міністерства освіти і науки України у межах науково-дослідної тематики відповідно до завдань держбюджетних тем МОН України: «Застосування інформаційних технологій в сфері цифровізації транспортних систем» (ДР № 0124U004241) і безпосередньо узгоджуються з пріоритетами національної транспортної стратегії та концептуальними засадами Національної стратегії Міністерства інфраструктури України «Drive Ukraine 2030» щодо впровадження інноваційних технологій обробки великих масивів інформації для розбудови інтелектуальних транспортних систем.

Наукова новизна результатів дисертації та значення для практики

Можна погодитись з автором дисертаційної роботи, що він вирішує проблему усунення протиріччя між існуючи станом інформатизації транспортних систем і потенційними можливостями інтелектуалізації транспортних процесів та систем. На прикладі оцінки технічного стану транспортних засобів за допомогою проактивного моніторингу було вперше розроблено:

Розподілену інформаційну технологію дистанційного контролю транспортних засобів, яка, на відміну від відомих реляційних систем, інтегрує нереляційні NoSQL-бази даних (MongoDB), паралельні обчислення MapReduce та потокову обробку через брокери повідомлень (Apache Kafka)., що дозволяє відмовитися від жорстко структурованих реляційних SQL-схем, що працюють за традиційною та повільною моделлю «запит-відповідь», на користь гнучкої гетерогенної архітектури (BASE), здатної обробляти швидкі нестаціонарні потоки телеметрії. Це забезпечує консистентність даних і горизонтальну масштабованість за умов нерегулярного надходження діагностичної інформації (Logged Vehicle Data, LVD).

Набули подальшого розвитку застосування архітектур нейроних мереж – багатошарового перцептрона (MLP) та рекурентних мереж довгої короткочасної пам'яті (LSTM) для виявлення прихованих дефектів у нестаціонарних потоках телеметрії OBD-II; порівняльним аналізом доведено переваги ансамблевих методів (Random Forest) над логістичною регресією за метриками F1-score та Matthews Correlation Coefficient (MCC) при класифікації технічного стану транспортних засобів.

Теоретичні положення та рекомендації дозволяють на практиці надати суб’єктам технічного обслуговування моделі та методи, засновані на використанні машинного навчання, штучних нейронних мереж, для визначення працездатності вузлів і агрегатів транспортних засобів. Загалом ці результати дають змогу приймати рішення з удосконалення технічного обслуговування власників транспорту. Результати дисертаційної роботи використано у практиці роботи організації Харківського регіону, що підтверджено актами впровадження.

Обґрунтованість наукових положень, висновків і рекомендацій та достовірність отриманих результатів

В основу дисертаційної роботи покладені основні положення, які базуються на інтелектуалізації моніторингу технічного стану транспортних засобів. Методи і моделі машинного навчання використані для прогнозування поломок для прогнозу поломок вузлів та агрегатів транспортних засобів.

Створення інтелектуальної технології моніторингу базується на програмно-апаратних методах створення розподілених комп'ютерних систем і сховищ даних.

Інструментальні методи та засоби автора не суперечать класичним, а отримані результати за своєю структурою збігаються з традиційними і стосовно до окремих випадків ведуть до оригінальних рішень. Одержані результати піддаються чіткій інтерпретації і допускають можливість їх безпосереднього використання на практиці. Все це свідчить про достатню обґрунтованість наукових положень, висновків і рекомендацій дисертаційної роботи.

Повнота викладення результатів

Результати дисертаційної роботи опубліковано у 20 наукових працях, з яких: 2 статті у наукових фахових виданнях України; 2 статті, що входять до наукометричних баз Scopus та Web of Science; 16 публікацій у матеріалах конференцій.

Зміст дисертації та відповідність встановленим вимогам

Дисертаційна робота складається з анотації двома мовами, змісту, вступу, основной частини, з чотирьох розділів, висновків, списку використаних джерел, додатку. У роботі представлено 64 рисунки і 20 таблиць. Загальний обсяг дисертації складає 248 сторінок, з них: 218 основного тексту, спи-

сок використаних джерел охоплює 27 сторінок. Список використаних джерел складає 219 найменувань.

Оцінка розділів дисертаційної роботи

У першому розділі виконано системний аналіз методів та інформаційних систем діагностики, оцінки і прогнозування технічного стану транспортних засобів. Автор розглядає еволюцію систем управління базами даних. Аналізує такі парадигми технічного обслуговування як реактивне (Run-to-Failure), планово-профілактичне та прогностичне (Predictive Maintenance). Досліджує можливості технологій Big Data для прогнозування транспортних потоків, методи Floating Car Data (FCD), алгоритми машинного навчання для керування трафіком, застосування аналітики в концепції розумного міста (Smart City, V2X). Показано потенціал OBD-II-даних як джерела формування інформаційних масивів Big Data для задач машинного навчання. За результатами аналізу сформульовано науково-прикладне завдання, наукова гіпотеза, мета та послідовність подальших досліджень. Розділ створює достатнє теоретичне обґрунтування для подальших розробок.

У другому розділі розглянуто принципи збору телеметричної інформації через інтерфейс OBD-II, детально проаналізовано стандартизовані режими діагностики та класифікацію кодів несправностей (DTC). Вирішено проблему розрідженості даних у базах LVD засобами статистичного оцінювання щільності розподілу та мінімізації втрат вибірки. Обґрунтовано методологію проектування сховища даних на базі ETL-процесів, нереляційних документо-орієнтованих СУБД (MongoDB) та концепції семантичного сховища з онтологічними зв'язками між різнорідними джерелами діагностичної інформації.

Результати розділу є змістовними, та мають самостійне наукове значення. Вони формують інформаційну основу для побудови моделі прогнос-

тичного моніторингу і інтеграції її у систему моніторингу технічного стану транспортних засобів.

У третьому розділі побудовано аналітичний конвеєр прогностичного моніторингу та розроблено математичні моделі системи. Для класифікації технічного стану транспортних засобів реалізовано й порівняно чотири підходи: багатошаровий перцептрон (MLP), рекурентні мережі LSTM для аналізу нестаціонарних часових рядів показань датчиків, алгоритм Random Forest та логістичну регресію. Оцінку ефективності розроблених моделей здійснено за розширеним набором метрик, що базується на аналізі матриці плутанини, оскільки набори даних автосервісу є глибоко незбалансованими. Зокрема, використано Accuracy (загальну точність класифікації), F1-score (гармонійне середнє між точністю та повнотою), ROC AUC (площу під кривою робочої характеристики приймача) та MCC (коефіцієнт кореляції Метьюса для комплексної оцінки якості класифікації). Головним критерієм оцінки ефективності визначено здатність алгоритму максимізувати показники F1-score та MCC при одночасній суворій мінімізації хибно-негативних прогнозів (False Negatives), що дозволяє уникнути небезпечного пропуску реальних критичних дефектів ТЗ.

Важливим результатом розділу є побудова та налаштування комплексу математичних моделей машинного навчання для класифікації технічного стану транспортного засобу. Загалом третій розділ є центральним і демонструє працездатність розробленної моделі.

У четвертому розділі проведено проектування інформаційної системи, розроблено об'єктно-орієнтовану UML-модель та деталізовано архітектуру програмного комплексу: модуль збору даних з OBD-сканера, препроцесор нормалізації та перевірки цілісності, підсистема реального часу, сховище даних та блок машинного навчання. Реалізовано відмовостійку мікросервісну архітектуру з балансувальниками навантаження та резервними вузлами. Описано механізми автоматизованого планування ETL-процесів, стиснення архівних даних та реляційну схему метабази (PostgreSQL) для збереження струк-

турованих результатів діагностики. Розроблено прикладний рівень системи для інтеграції в екосистему Smart City та периферійних обчислень (Edge Computing). У розділі також представлено методіку оцінки ефективності представлених рішень. Проведен розрахунок ефективності прогностичної системи порівняно з реактивними систами обслуговування. Розділ підтверджує можливість використання підходу в реальних системах технічного обслуговування транспортних засобів.

Зауваження по роботі

1. В оглядової частині роботи автору слід було приділити більше уваги аналізу розвитку систем технічного обслуговування у нашої країні, що так або інакше обумовлює вимоги, що висуваються до застосування методів і засобів моніторингу технічного стану транспортних засобів;
2. В дисертаційній роботі використовується термін Big Data, но відсутній аналіз які об’єми даних відносяться до Big Data;
3. В другому розділі дисертаційної роботи достатньо багато уваги приділяється опису роботи CAN, але це не є центральною темою дослідження;
4. Автор аксіоматично визначає багатошаровий перцептрон, логістичну регресія і випадковий ліс у якості моделей для класифікації технічного стану. Не зовсім зрозуміло, чому саме ці моделі були обрані для порівняння;
5. В роботі не зовсім зрозуміло який був розмір датасету для налаштування моделі прогностичного моніторингу, недостатньо детально, розкрито питання обсягу вибірок та період їх накопичення;
6. У третьому розділі дисертаційної роботи наведена частина блоку «predictions» для розробленого алгоритму. Можливо цей матеріал слід було розмістити у додатках;

7. В четвертому розділі дисертаційної роботи при розрахунку експлуатаційно-економічних коефіцієнтів автор наводить загальні витрати на ТО та вартості ремонтних робіт, но не пояснює природу формування вартості витрат;
8. Архітектуру інформаційної технології подано на достатньо концептуальному рівні, однак питання кібербезпеки, захисту даних, відмовостійкості, апаратного забезпечення, збоїв в передачі даних висвітлено стисло. Доцільно було б доповнити роботу описом захисту каналів, резервування брокерів і сховищ та режимів роботи за втрати зв'язку.;
9. При оцінці ефективності системи не приділялась увага саме вартості розгортання і підтримки системи;

Наведені зауваження мають більше уточнюючий та методичний характер, визначають перспективи подальшого розвитку і не знижують загальної позитивної оцінки дисертаційної роботи.

Висновок по роботі

Оцінюючи роботу в цілому, можна зробити висновок, що дисертація є завершеним науковим дослідженням галузі інформаційних технологій. Основні результати дисертаційної роботи досить повно надруковані у фахових виданнях, що визнані МОН України та журналах які входять до науковометричних баз Scopus, апробовано на міжнародних наукових конференціях та впроваджено на практиці.

Враховуючи обсяг дослідження, наукову та практичну значимість отриманих результатів, вважаю, що дисертаційна робота "Методи та моделі інформаційних технологій для моніторингу технічного стану транспортних засобів" відповідає вимогам Порядку присудження ступеня доктора філософії та скасування рішення спеціалізованої вченої ради, затвердженого постановою Кабінету Міністрів України від 12 січня 2022 р. №44, а також Вимо-

гам до оформлення дисертації, затвердженим наказом МОН України від 12 січня 2017 р. №40., а її автор, Грицук Валерій Юрійович, заслуговує на присудження йому наукового ступеня доктора філософії з галузі знань 12- Інформаційні технології за спеціальністю 122 Комп'ютерні науки.

Офіційний опонент,
професор кафедри інформаційних систем та технологій,
Національний технічний університет
«Харківський політехнічний інститут»,
доктор технічних наук, професор Валентина МОСКАЛЕНКО

Підпис Москаленко
ЗАСВІДЧУЮ:
ВЧЕНИЙ СЕКРЕТАР
НАЦІОНАЛЬНОГО ТЕХНІЧНОГО УНІВЕРСИТЕТУ
"ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ"
"___" _____ 20____ р.

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний технічний університет "Харківський політехнічний інститут"
код 02071180