

ВІДГУК
офіційного опонента
на дисертаційну роботу Барбарук Ліни Вікторівни
«Моделі та метод обробки великих даних в інформаційно-аналітичних системах
моніторингу водних об'єктів»,
що представлена
на здобуття наукового ступеня кандидата технічних наук
за спеціальністю 05.13.06 – інформаційні технології

1. Актуальність теми.

Поширення інформаційних технологій, зокрема інтернету речей у галузі сучасної гідроекології дозволяє оцінювати і прогнозувати стан водних об'єктів з точки зору сталого розвитку і збереження базових компонент екосистеми. В Україні останнім часом приділяється величезне значення проблемі вдосконалення екологічного моніторингу стану навколишнього природного середовища та, особливо, моніторингу водних об'єктів. Моніторинг водних ресурсів є важливим інструментом ефективного контролю стану якості води природного середовища, своєчасного попередження шкідливого впливу забруднювачів, а також широкого та швидкого інформування про стан і тенденції зміни стану водних об'єктів усієї гідроекосистеми. Водночас, існуюча в країні система моніторингу ще не повністю відповідає міжнародним вимогам. Основні засади системного підходу та комплексного аналізу екологічного стану водних екосистем дають змогу удосконалити чинну систему моніторингу.

Актуальність розроблення моделей та методу обробки великих даних в системах моніторингу водних об'єктів обумовлена тим, що вони дозволяють суттєво підвищити ефективність обробки та аналізу даних, які отримані в реальному масштабі часу, в інформаційно-аналітичних системах моніторингу водних об'єктів і при виборі набору базових компонент для прийняття рішень щодо стану якості води.

У зв'язку з цим напрям наукових досліджень, обраний Л.В. Барбарук є актуальним.

2. Зв'язок роботи з науковими програмами, планами, темами.

Дисертаційна робота Л.В. Барбарук виконувалась у відповідності з науково-технічними планами кафедри комп'ютерних наук та інженерії Східноукраїнського національного університету імені Володимира Даля, в рамках державних програм і планів НДР, а також міжнародних проектів: НДР «Дослідження стратегій та механізмів прийняття рішень для інтегрованого управління водними ресурсами» (№ ДР 0116U005784); НДР «Проектування системи моніторингу та контролю водних об'єктів на основі технології інтернет речей», (№ ДР 0120U100421); проект Європейського Союзу ERASMUS+ ALIOT 573818-EPP-1-2016-1-UK-EPPKA2-CBHE-JP «Internet of Things: Emerging Curriculum for Industry and Human Applications».

3. Ступінь обґрунтованості наукових положень, висновків і рекомендацій.

Обґрунтованість наукових положень дисертаційної роботи забезпечена коректним використанням теорії нечітких множин, нечіткої логіки для розробки моделі для аналітичної обробки великих даних та нечіткої евристичної кластеризації, узагальненням праць вітчизняних і зарубіжних авторів і підтверджена використанням статистичних та аналітичних даних, оброблених з застосуванням методів теоретико-множинного опису, системного аналізу та методів інтелектуального аналізу.

Обґрунтованість наукових висновків і рекомендацій забезпечена логічною аналітичною структурою роботи, перевіркою отриманих результатів і рекомендацій на практиці.

Достовірність наукових положень, висновків і рекомендацій, отриманих у дисертації, підтверджена результатами теоретичних та експериментальних досліджень, коректним застосуванням математичного апарату теорії нечітких множин, сучасних методів комп'ютерного моделювання, а також впровадженням інформаційної технології обробки великих даних в інформаційно-аналітичних системах моніторингу водних об'єктів.

4. Новизна наукових положень, висновків і рекомендацій.

Наукова новизна результатів дисертаційного дослідження полягає в розвитку сучасних інформаційних технологій ефективною обробки та аналізу великих даних, що отримані від систем моніторингу водних об'єктів екосистеми.

В ході досліджень сформульовано нові наукові положення і висновки, отримано низку нових результатів, що мають суттєве значення.

1. Вперше розроблено модель для аналітичної обробки великих даних системи моніторингу водних об'єктів на основі системи нечіткого виведення та застосуванні нечітких правил для оцінки якості води згідно методики віднесення масиву поверхневих вод до одного з класів екологічного та хімічного станів масиву поверхневих вод.

2. Удосконалено метод нечіткої кластеризації, за рахунок введення процедури автоматичного маркування елементів, що знаходяться на межі нечітких кластерів, і відповідно, самих кластерів. Це дозволило застосовувати автоматичну розмітку при обробці великих даних і покращити конвергенцію алгоритму кластеризації.

3. Удосконалено технологію та засоби візуалізації великих даних моніторингу водних об'єктів, шляхом впровадження технології спрощення полігональних ланцюгів для вибору рівня деталізації вихідних зображень часових рядів даних про якість води, що дозволило зменшити час на пошук невідповідностей і ретроспективний аналіз великих даних.

4. Набула подальшого розвитку інформаційна технологія обробки великих

даних в інформаційно-аналітичних системах моніторингу водних об'єктів шляхом її адаптації до завдань контролю та управління рибогосподарських підприємств.

5. Практична значимість і шляхи використання отриманих результатів.

Основні наукові результати, які отримані автором у ході дисертаційних досліджень представляються досить важливими, мають фундаментальний характер і спрямовані на подальший розвиток і вдосконалення інформаційних технологій в області моніторингу водних об'єктів. Запропоновані автором моделі і метод обробки великих даних пройшли практичне випробування і впроваджені в Управлінні Державного агентства рибного господарства у Луганській області та у навчальний процес Східноукраїнського національного університету імені Володимира Даля (наявні відповідні акти впровадження).

Результати дослідження можуть бути рекомендовані для впровадження на підприємствах з інтенсивним контролем характеристик якості води, зокрема таких що використовують процеси водопідготовки, кондиціонування води, та в спеціалізованих аналітичних системах обліку вод, системах управління динамічними промисловими процесами керування якістю вод, тощо.

Крім того, слід відзначити, що теоретичні основи роботи стали основою для створення програмного забезпечення інформаційної технології для обробки великих даних в інформаційно-аналітичних системах моніторингу водних об'єктів, яке дало змогу підвищити точність інтегральної оцінки якості води, зменшити час на прийняття рішень щодо якості вод водойм рибогосподарського призначення та знизити витрати на виробництво.

6. Зміст, завершеність, стиль викладення, публікації, апробація.

Дисертаційна робота складається зі вступу, 4 розділів, висновків, списку використаних джерел зі 177 найменувань на та 3 додатків. Оформлення дисертації виконано відповідно до діючих вимог.

У вступі показано актуальність проблеми, сформульовано мету й задачі дослідження, описані об'єкт, предмет і методи дослідження, дана характеристика наукової новизни й практичного значення результатів, а також зазначені особистий внесок здобувача та дані щодо реалізації, апробації та публікації результатів.

Перший розділ містить аналіз моделей та методів обробки великих даних. В ньому розглянуто потенціал інноваційних технологій моніторингу для кращого управління водними ресурсами, проаналізовано особливості великих даних в системах моніторингу водних об'єктів, виконано аналіз нечітких моделей, особливостей кластеризації великих даних, розглянуто парадигму MapReduce. Виконано огляд проблем візуалізації великих даних та підходів до їх вирішення. Сформульоване основне науково-прикладне завдання роботи, яке

полягає в розробці моделей та методу інформаційної технології для інформаційно-аналітичних систем моніторингу водних об'єктів.

Другий розділ присвячено дослідженню та розробленню нечітких моделей для оцінювання якості води на прикладі поверхневих вод та вод рибогосподарського призначення. Нечіткий логічний формалізм використаний для опису якості води шляхом розробки індексу якості води на основі нечітких міркувань. Розроблено модель для аналітичної обробки великих даних системи моніторингу водних об'єктів на основі формалізації її атрибутів та інтерпретації невизначеності оцінки якості води у вигляді лінгвістичних змінних, яка дозволяє надати інтегровану характеристику стану водних об'єктів, для подальшого прийняття рішення.

Третій розділ присвячений розробленню методу обробки великих даних для водойм рибогосподарського призначення. Розглянуто проблему інтерпретації результатів кластеризації на випадок застосування евристичних алгоритмів кластеризації для інтуїтивістської нечіткої обробки даних. Запропоновано удосконалений метод кластеризації швидких нечітких С-середніх для оброблення великих даних моніторингу водних об'єктів. Наведено приклад використання методу для аналізу якості вод водойм рибогосподарського призначення.

У четвертому розділі представлені етапи розробки інформаційної технології для обробки великих даних, отриманих від станцій моніторингу водних об'єктів та результати реалізації цієї технології у вигляді системи підтримки прийняття рішень. Описана архітектура сховища даних, визначені критерії якості видобутку даних, представлені структура бази даних та основні інструменти для аналітичної обробки даних у вигляді гібридного інструмента інтелектуального аналізу даних та управління непрямыми знаннями баз даних спеціалізованої аналітичної системи водних об'єктів. Наведені результати удосконалення та адаптації методів візуалізації великих даних (у вигляді хордових діаграм і методів спрощення полігональних ланцюгів) на випадок моніторингу стану водних об'єктів в реальному масштабі часу.

Додатки містять список публікацій здобувача, довідкову інформацію щодо показників якості вод рибогосподарських підприємств, а також акти впровадження результатів дисертаційного дослідження.

В цілому робота характеризується достатньо високим теоретичним рівнем, базується на узагальненні значної кількості теоретичних і інформаційних джерел. Дисертація має досить логічну структуру, її мова відповідає рівню науково-технічних видань, термінологія використовується коректно, стиль викладу зручний для аналізу. Автореферат дисертації повністю відображає зміст дисертаційного дослідження.

Результати дисертації достатньо повно висвітлені в статтях і доповідях. Зокрема за темою дисертації з викладенням її основних результатів

опубліковано 18 наукових праць, з яких 8 статей в фахових наукових виданнях України, включених до міжнародних наукометричних баз, 1 стаття у закордонному виданні, що внесено до міжнародної наукометричної бази Scopus; 1 стаття у фаховому виданні іншої держави, що входить до Європейського Союзу, 8 публікацій у збірниках матеріалів і праць міжнародних та всеукраїнських конференцій. Їх аналіз дає підставу зробити висновок, що результати дисертаційних досліджень у достатній мірі висвітлені в періодичних наукових виданнях, які мають відповідний статус.

У тексті дисертаційної роботи та у публікаціях, що містять результати наукових досліджень Л.В. Барбарук, порушень принципів академічної доброчесності (плагіату та самоплагіату, фабрикації та фальсифікації) не виявлено

Дисертація та її автореферат повністю відповідають вимогам пунктів 9, 11, 12 «Порядку присудження наукових ступенів», затверджених постановою Кабінету Міністрів України від 24 липня 2013 р. № 567 (зі змінами від 30.12.2015 та 27.07.2016 р.), що висувуються до кандидатських дисертацій.

7. Зауваження по дисертаційній роботі.

1. В першому розділі роботи при огляді проблем моніторингу водних об'єктів доцільно було б розглянути існуючі спеціалізовані інформаційні системи, що використовуються для розв'язання задач подібного типу.

2. В другому розділі більш уваги слід було приділити аргументації вибору саме трієподібних функцій належності при описі параметрів якості води як лінгвістичних змінних. Також не зрозуміло яким чином формуються правила в системі нечіткого логічного виведення та задаються вагові коефіцієнти факторів, що впливають на оцінку якості води.

3. В третьому розділі доцільно було б більш детально описати як обираються параметри удосконаленого методу нечіткої кластеризації, а саме: вагові коефіцієнти в формулі розрахунку відстані (3.16); значення експоненційної ваги для розрахунків центрів кластерів та матриці належності, а також кількість кластерів. Ці параметри досить суттєво впливають на результат кластеризації.

4. В третьому розділі при оцінці ефективності запропонованого удосконаленого методу нечіткої кластеризації не наведено жодних даних щодо кількісних оцінок обчислювальної складності методу та часових параметрів його збіжності (в залежності від обсягу даних) в порівнянні з відомими методами. Це знижує можливості оцінки ефективності використання запропонованого методу при обробці великих даних.

Вказані вище зауваження не мають принципового характеру та не впливають на загальну позитивну оцінку дисертаційної роботи.

8. Загальні висновки

Дисертаційна робота Барбарук Л.В. «Моделі та метод обробки великих даних в інформаційно-аналітичних системах моніторингу водних об'єктів» є завершеним науковим дослідженням, що виконана на високому науковому рівні, в якій отримано нові науково обґрунтовані та апробовані результати.

Усі теоретичні положення дисертації доведено до відповідних інженерних рішень із застосуванням запропонованої інформаційної технології обробки великих даних в інформаційно-аналітичних системах моніторингу водних об'єктів.

Робота відповідає паспорту спеціальності 05.13.06 – інформаційні технології. За актуальністю розглянутих задач, обсягом досліджень, науковим рівнем і практичною цінністю отриманих результатів дисертаційна робота повністю відповідає вимогам пп. 9, 11-13 «Порядку присудження наукових ступенів і присвоєння вченого звання старшого наукового співробітника», затвердженого постановою Кабінету Міністрів України від 24 липня 2013р. № 567, до кандидатських дисертацій.

Здобувач Барбарук Ліна Вікторівна заслуговує присудження наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології.

Офіційний опонент,
завідувач кафедри інформаційних управляючих систем
Харківського національного університету радіоелектроніки,
доктор технічних наук, професор

К.Е. Петров

Підпис засвідчую,
Учений секретар
Харківського національного університету радіоелектроніки,
кандидат технічних наук, доцент



І.В. Магдаліна