

18.05.2020  
116-05.65

## ВІДГУК

офіційного опонента

доктора технічних наук, професора Шостака Ігоря Володимировича

на дисертацію Захожая Олега Ігоровича

«Моделі, методи та інформаційна технологія гібридного розпізнавання образів для консолідованої обробки даних в складних системах», поданої на здобуття наукового ступеня доктора технічних наук за спеціальністю 05.13.06 – Інформаційні технології

### **Актуальність теми дисертаційної роботи.**

Постійне ускладнення сучасних об'єктів інформатизації призводить до появи все більшої кількості складних систем, для яких неможлива повна формалізація їхньої поведінки. В цьому аспекті апарат розпізнавання образів, який базується на використанні апріорної інформації та інтелектуальних алгоритмах обробки даних, дозволяє отримати переваги для прийняття достовірних рішень.

Однак, підчас використання методів та інформаційних технологій розпізнавання образів, постійно виникає проблема збільшення часової складності процесів обробки даних, а також залежність достовірності результату від якості вхідних даних. Особливо критичного характеру ці процеси набувають у випадку наявності неоднорідних даних, що значно збільшує кількість операцій співставлення ознак при прийнятті рішення.

В цьому аспекті, тема дисертаційної роботи Захожая О.І. яка направлена на розробку моделей, методів та інформаційної технології для обробки неоднорідних даних в складних системах є, безсумнівно, актуальною.

### **Оцінка змісту і загального рівня роботи, міри вирішення наукової проблеми.**

Наукова проблема, яка вирішується автором дисертаційної роботи, полягає у підвищенні достовірності та зниження часової складності обробки

даних та прийняття рішення в складних системах за умови наявності різноманітних перешкод і викривлень. В дисертаційній роботі представлено вирішення даної проблеми через розробку моделей, методів, а на їхній основі нової інформаційної технології гібридного розпізнавання образів, яка базується на отриманні, селекції, ранжуванні та співставленні інформативних ознак різної природи виникнення.

### **Зв'язок роботи з науковими програмами, планами, темами.**

Дисертаційна робота виконувалась відповідно до пріоритетних напрямів науково-дослідних робіт Луганського національного університету імені Тараса Шевченка, в рамках науково-дослідних робіт: «Розробка інформаційної технології моделювання та оцінювання фінансово-економічних ризиків із врахуванням невизначеностей різної природи» (№ ДР 0113U000650); «Розробка методології системного аналізу моделювання та оцінювання фінансових ризиків» (№ ДР 0115U000356); «GreenCo – Green Computing and Communication» (530270-TEMPUS-1-2012-1-UK-TEMPUS-JPCR); «Розробка і дослідження методів синтезу індивідуальних стратегій управління в адаптивних навчальних системах на основі World wide web-технологій» (№ ДР 0109U007908); «Розробка та дослідження методів синтезу адаптивних комп'ютерних мультимедіа систем» (№ ДР 0198002639).

### **Наукова новизна отриманих результатів:**

*вперше:*

– побудовано ситуаційно-подійну модель гібридного розпізнавання образів, засновану на поданні характеристик складної системи у вигляді сукупності образів, ознаки яких мають різну природу виникнення, множини характеристик зовнішніх умов – як проява ситуації, статичної складової ситуації – у вигляді множини стаціонарних характеристик інформативності, динамічної складової – у вигляді нестаціонарних характеристик інформативності, та сукупності класів, як результату виконання послідовності дій по розпізнаванню;

– розроблено метод консолідації неоднорідних даних при гібридному розпізнаванні, який оснований на теорії грубих множин та визначенні стаціонарних і нестаціонарних критеріїв інформативності, в залежності від наявних перешкод і викривлень, та дозволяє здійснити селекцію образів складної системи з метою зменшення кількості співставлень даних для отримання рішення;

– розроблено метод та інформаційну модель визначення послідовності обробки даних, які базуються на узагальненні моделі емоційних процесів людини, представленої Аткинсоном-Шифріном, для складних систем з неоднорідними даними та різнорідним проявом перешкод і викривлень, що дозволяє в динаміці системи зменшити кількість даних, які підлягають співставленню та отриманню достовірного результату;

– розроблено метод та інформаційну технологію гібридного розпізнавання образів, які основані на роздільному аналізі інформативних ознак та пошуку груп образів з ідентичною класифікацією, що дозволяє забезпечити апріорно завданий рівень достовірності результату консолідованої обробки неоднорідних даних в складних системах з різнорідним проявом перешкод і викривлень;

– розроблено метод екстенсiонально-інтенсiонального аналізу ознак, який оснований на прийнятті швидкого рішення по узагальненим характеристикам складної системи та використання, в разі необхідності, деталізованого аналізу даних для уточнення результату, що дозволяє зменшити кількість співставлень даних при отриманні апріорно визначеного рівня достовірності класифікації;

– розроблена інформаційна модель підтримки прийняття рішення, яка у випадку отримання багатоальтернативного результату класифікації стану складної системи, за рахунок використання інформації про групи образів з ідентичною класифікацією, дозволяє представити для експертного висновку отримані результати класифікації (альтернативи), ранжовані у порядку

зменшення їхньої достовірності, а також інформацію про сукупності образів, на основі яких ці класифікації були отримані.

*отримала подальший розвиток* теорія комбінованих систем розпізнавання образів, в якій опис об'єкту розпізнавання здійснюється ознаками різної природи виникнення, а не різного характеру інформації, що дозволяє здійснювати достовірну класифікацію за умови широкої зміни різнорідних перешкод і викривлень.

*удосконалено* матричну інформаційну модель складної системи, в якій неоднорідні дані компонуються не в одному, а в різних шарах, що забезпечує більшу гнучкість у виборі альтернатив консолідації даних та методів їхньої обробки.

**Обґрунтованість та достовірність наукових положень, висновків і рекомендацій.**

Розроблені в роботі нові наукові положення пройшли перевірку експериментальними чисельними дослідженнями, експериментальними дослідженнями в умовах лабораторії та на трьох реальних інформаційних системах обробки даних різної структурної організації та призначення.

**Значимість отриманих результатів для науки і практичного використання** полягає в придатності застосування розробленого підходу, методів, інформаційної технології гібридного розпізнавання образів для прийняття достовірних та оперативних рішень в широких межах змін умов отримання вхідних даних та рівня перешкод і викривлень.

Результати досліджень були використані: у коксохімічному виробництві для вдосконалення автоматизованої системи контролю просторового розподілу температури коксової печі; в системі ультразвукового вимірювання лінійної відстані для засобів автоматизації; в системі перевірки унікальності текстових даних.

Результати наукових досліджень також використані в освітньому процесі при викладанні низки дисциплін спеціальностей галузі знань 12 «Інформаційні технології».

Основні результати роботи можуть бути використані для побудови різноманітних систем автоматизованої переробки інформації і управління, де об'єкт інформатизації представляє собою складну, багатопараметричну систему з неоднорідними даними, для якої неможлива побудова повної математичної моделі поведінки, або така модель значно переускладнена та не може використовуватися для прийняття достовірного та оперативного рішення.

### **Повнота викладення результатів досліджень в опублікованих працях.**

Основні результати дисертаційного дослідження опубліковані у 42 наукових працях, в тому числі: 10 праць у фахових виданнях, рекомендованих МОН України для опублікування результатів наукових робіт; 16 статей в фахових виданнях, рекомендованих МОН України та таких, що входять до наукометричних баз даних Scopus та ін.; 1 стаття в іноземному виданні, що входить до наукометричних баз даних Web of Science та ін; 9 тез доповідей на міжнародних та всеукраїнських конференціях; отримано 6 патентів України.

Аналіз змісту наведених вище робіт надає змогу стверджувати, що всі основні положення дисертації повною мірою представлені у цих публікаціях.

### **Оцінка змісту дисертаційної роботи.**

Дисертаційна робота Захожая О.І. складається зі вступу, п'яти розділів, висновків, списку літературних джерел з 236 найменувань та додатків. Загальний обсяг складає 317 сторінок (з них 259 сторінок основного тексту). Також робота містить 49 рисунків і 7 таблиць.

У вступі наведено загальну характеристику роботи, яка підкреслює її актуальність, визначено наукову проблему, предмет і об'єкт дослідження, висунуто наукову гіпотезу, що підлягає перевірці, сформульовано мету, задачі дослідження та отримані наукові результати.

У першому розділі проведений аналіз методів і засобів обробки даних в складних системах. В результаті цього аналізу визначається основна

проблема, пов'язана з певним рівнем невизначеності та стохастичності впливу вхідних даних на визначення стану складної системи. Таким чином, для обробки даних у складних системах обґрунтовується доцільність використання апостеріорної інформації як про саму систему, так і про зовнішнє для неї середовище.

З урахуванням того, що в складних системах фактично відсутній формалізований опис закономірностей впливу окремих параметрів на поведінку системи, обґрунтована доцільність підходу, коли для прийняття рішень використовується ідентифікація не окремих параметрів, а стану складної системи.

Далі в розділі проводиться аналіз існуючої матричної інформаційної моделі та методів комбінованого розпізнавання образів, які використовуються для обробки неоднорідних даних в складних системах. В результаті аналізу визначається, що існуюча модель не враховує ступень інформативності ознак складної системи, який значно впливає на достовірність результату розпізнавання. Крім цього, існуючі методи комбінованого розпізнавання, які направлені на порівняння усіх поєднань ознак складної системи для отримання результату, мають достатньо високу часову складність, яка істотно збільшується зі збільшенням кількості даних, що підлягають обробці.

На основі проведеного аналізу визначено науково-технічну проблему, яка полягає у необхідності забезпечення одночасного підвищення достовірності розпізнавання та зниження часової складності аналізу даних, оскільки ці показники є конкуруючими.

На основі проведених досліджень, автором сформульовано наукову гіпотезу, що підлягає перевірці в ході дослідження.

У другому розділі обґрунтовано новий принцип консолідації даних, який полягає у формуванні для складної системи не одного, а сукупності образів, ознаки яких мають різну природу виникнення. В цьому випадку, для різних зовнішніх умов, ознаки мають різну інформативність для отримання

достовірного результату розпізнавання. Таким чином, відповідно до поточних зовнішніх умов, прийняття рішення можна здійснювати на основі не повного порівняння ознак, а тільки тієї сукупності, що мають достатню інформативність. Таким чином, у будь-який момент часу, прийняття рішень здійснюється на основі співставлення меншої кількості але найбільш інформативних даних. Такий підхід в роботі названий гібридним розпізнаванням.

Застосування нового принципу консолідації призвело до необхідності удосконалення існуючої матричної інформаційної моделі складної системи, де вводяться сукупності образів різної природи виникнення та відповідні до них характеристики інформативності.

Надалі, для забезпечення прийняття рішень на основі обмеженої кількості але найбільш інформативних даних, була розроблена ситуаційно-подійна модель, яка враховує зовнішні фактори як ознаку поточної ситуації, та дозволяє здійснювати селекцію ознак складної системи на основі визначених характеристик інформативності.

Для ранжування ознак образів та визначення черги обробки даних в роботі запропонована інформаційна модель, яка є адаптацією моделі розумової активності людини відносно технічних систем.

В третьому розділі, на основі раніше представлених моделей, розробляються методи гібридного розпізнавання образів та прийняття остаточного рішення у випадку багатоальтернативного результату класифікації.

В основі розробленого методу гібридного розпізнавання покладено принцип роздільного аналізу ознак, та визначення груп ідентичних класифікацій, так як апіорно відомо, що за кожним образом з наявної сукупності, у випадку абсолютно точної класифікації, об'єкт розпізнавання може біти віднесений до одного й того ж самого класу.

Розроблений екстенсіонально-інтенсіональний метод аналізу ознак базується на теорії Д.О. Поспелова, щодо визначення узагальнених і

деталізованих ознак складних систем. Згідно методу, пріоритет у обробці мають узагальнені ознаки, і тільки якщо в процесі співставлення не отримується заданий рівень достовірності, здійснюється уточнення результату через порівняння деталізованих ознак.

Розроблені інформаційна модель і метод прийняття рішень у випадку багатоальтернативного результату класифікації базується на ранжуванні отриманих груп ідентичних класифікації за кількісними показниками. Таким чином, результати розпізнавання представляються для остаточного прийняття рішення у порядку зменшення їхньої достовірності.

Слід зазначити, що усі розроблені методи покладено в основу отриманих патентів України.

В четвертому розділі, на основі представлених моделей і методів, розробляється інформаційна технологія гібридного розпізнавання образів, в основі якої покладено побудову багатопоточної черги обробки даних, де кожен потік містить ознаки відповідного образу складної системи, що ранжовані у порядку зменшення їхньої інформативності.

В процесі співставлення ознак, визначається ступінь подібності стану складної системи до одного з наперед визначених класів. Кількісне визначення ступеня подібності здійснюється на основі обчислення міри Хеммінга, яка для спрощення програмної реалізації, виражена в роботі через показники чисельності груп ідентичних класифікацій.

Також в розділі приділена увага вимогам до апаратного забезпечення реалізації інформаційної технології гібридного розпізнавання образів у випадку виконання операцій в режимі реального часу. Зроблене обґрунтування використання багатопоточної обробки.

П'ятий розділ присвячений розробці прикладних інформаційних технологій гібридного розпізнавання образів та перевірки достовірності та ефективності отриманих наукових положень.

Для перевірки універсальності розробленого підходу, для апробації було вибрано три різні за структурою та функціоналом інформаційні



системи: аналізу просторового розподілу температури коксового пирога для коксохімічного виробництва, ультразвукового вимірювання відстані для засобів автоматики та перевірки текстових даних на унікальність для систем антиплагіату. В усіх прикладних застосуваннях був отриманий позитивний ефект від запровадження моделей і методів гібридного розпізнавання.

У висновках наведені основні результати дисертаційного дослідження, вказані сфери можливого їх застосування, а також зроблено аналіз виконання усіх поставлених задач та факти підтвердження висунутої наукової гіпотези.

У додатках наведені таблиці та графіки результатів математичного планування експерименту, та результати апробації прикладних інформаційних технологій на підприємстві, а також акти про впровадження результатів наукової роботи та перелік публікацій автора за тематикою дисертації.

#### **Зауваження до дисертації.**

По дисертаційній роботі можна зробити наступні зауваження:

1. Автор, на мій погляд, неприпустимо звужив клас складних систем, бо неможливість побудови адекватної математичної моделі внаслідок браку апріорної інформації є наслідком лише деяких особливостей такого явища як складність. Складність по своїй природі багатоаспектна, тому, наприклад, таку ознаку складності як змішаність природи (соціо-економо-технічні, природньо-штучні системи, тощо), необхідно враховувати при розробці методичних засобів, а на їх основі – відповідних інформаційних технологій.

2. В рамках сучасної парадигми вирізняють концептуальну (вона прописана у відповідних стандартах, зокрема, у ДСТУ), базові та прикладні інформаційні технології. Автор пише лише про прикладні ІТ, хоча ним розроблена, по суті, одна базова та три прикладних технології на її основі.

3. На стор. 98 (3-й абзац) автор пише про різну природу завод і перешкод, та їхній вплив на функціонування складної системи. При цьому як тут, так і далі по ходу викладення матеріалу, розглядаються тільки зовнішні

збурення, про внутрішні ж (параметричні), згадується лише при описі прикладної ІТ у коксохімічному виробництві (дефекти стінок коксової печі).

4. Розглядаючи функціонування систем реального часу (РЧ) автор пише лише про «жорсткий» та «м'який» режими РЧ, але не згадує про «стійкий (firm)» режим, у якому функціонує доволі велика кількість складних систем РЧ, зокрема, телекомунікаційних.

5. При формулюванні основних положень роботи автор робить акцент на зниженні часової складності методів класифікації. Викликає сумнів, що часова складність, при наявній потужності обчислювальних ресурсів, є проблемою при створенні стаціонарних систем управління великомасштабними об'єктами, функціонування яких характеризується високою інерційністю та порівняно низьким темпом протікання процесів.

6. При описі прикладної ІТ визначення унікальності текстів автор не вказав, для яких саме з шести відомих різновидів плагіату, зокрема плагіату на рівні ідеї, може бути застосована дана ІТ.

7. Матеріали оглядово-критичного характеру, стосовно «неоперцептрона» та визначення емоційних станів ОПП, доречно було б подати не в другому, а у першому розділі рукопису.

8. Представлені на рис. 3.4, 3.5 методи гібридного розпізнавання краще було б представити не у формі алгоритму, а у формі покрокової процедури, бо це більше відповідає загально прийнятому визначенню поняття «метод». При цьому метод консолідації різнорідних даних не описано ані у вигляді алгоритму, ані у вигляді покрокового набору процедур.

9. На графіках лінійних розподілів температури коксових печей (Додаток Б, Г) не вказано одиниць виміру по осі абсцис.

10. У рукопису дисертації мають місце русизми, в ряді випадків не правильно вказано відмінки.

Вказані недоліки не впливають на загальну позитивну оцінку виконаної роботи.

**Загальні висновки.**

Дисертаційна робота Захожая О.І. «Моделі, методи та інформаційна технологія гібридного розпізнавання образів для консолідованої обробки неоднорідних даних в складних системах» є завершеним науковим дослідженням, виконаним на високому науковому рівні, в ході цього дослідження отримано нові обґрунтовані наукові результати.

Робота відповідає паспорту спеціальності 05.13.06 – інформаційні технології. За актуальністю вирішеної проблеми, обсягом досліджень, науковим рівнем і практичною цінністю отриманих результатів дисертаційна робота повністю відповідає пп. 9, 10, 12-14 «Порядку присудження наукових ступенів», затвердженого Постановою Кабінету Міністрів України від 24 липня 2003 р. №567 (зі змінами та доповненнями), до докторських дисертацій.

Здобувач, Захожай Олег Ігорович, заслуговує на присудження наукового ступеня доктора технічних наук за спеціальністю 05.13.06 – інформаційні технології.

Офіційний опонент,  
доктор технічних наук, професор,  
професор кафедри інженерії програмного  
забезпечення Національного аерокосмічного  
університету імені М.Є. Жуковського «ХАІ»

  
І.В. Шостак

Підпис д.т.н., проф. Шостака І.В. засвідчую:

Помічник ректора із забезпечення якості освіти,  
кандидат технічних наук, професор



  
Ю.А. Воробйов