

УДК 681.518:004
КП
№ держреєстрації 0109U001380
Інв. №

Міністерство освіти і науки України
Сумський державний університет
(СумДУ)

40007, м. Суми, вул. Римського-Корсакова, 2
тел. (0542) 33-54-79 факс (0542) 33-54-79

ЗАТВЕРДЖУЮ

Проректор
з наукової роботи
д.ф-м.н., професор
А.М. Черноус
2009-12-24 . . .

ЗВІТ
ПРО НАУКОВО-ДОСЛІДНУ РОБОТУ
« РОЗРОБЛЕННЯ НАУКОВО-МЕТОДИЧНИХ ОСНОВ ТА
ІНФОРМАЦІЙНИХ ЗАСОБІВ ПРОЕКТУВАННЯ ЗДАТНИХ
САМОНАВЧАТИСЯ АДАПТИВНИХ СИСТЕМ КЕРУВАННЯ
ТЕХНОЛОГІЧНИМИ ПРОЦЕСАМИ»

(проміжний звіт)

Начальник НДЧ
канд. техн. наук, доцент
Керівник НДР
доктор техн. наук, професор

В.А. Осіпов
А.С. Довбиш

2009

Рукопис завершено 21 грудня 2009 р.

Результати даної роботи розглянуто
науковою радою СумДУ, протокол від 2009-11-26 № 5

СПИСОК АВТОРІВ

Керівник НДР, доктор техн. наук, професор	А.С.Довбиш (вступ, розділ 2)
Молодший наук. співробітник, канд. техн. наук	М.В.Козинець (підрозділ 1.4)
Молодший наук. співробітник	І.В. Шелехов (підрозділ 3.3)
Молодший наук. співробітник	С.О. Петров (підрозділ 2.2)
Молодший наук. співробітник	С.С. Мартиненко (підрозділ 2.3)
Молодший наук. співробітник	В.А. Тронь (підрозділ 3.4)
Інженер другої категорії	Н.М.Бровко (підрозділ 1.3)
Провідний інженер	О.М. Склокін (підрозділ 3.4)
Провідний інженер	О.Д. Пачола (підрозділ 1.1)
Провідний інженер	Ж.Ю. Шматко (підрозділ 2.3)
Провідний інженер	О.В. Пустовгар (підрозділ 1.2)
Провідний інженер	І.В. Возна (підрозділ 2.4)
Аспірант	О.О. Дзюба (підрозділи 3.1, 3.3)
Аспірант	С.М. Котенко (підрозділ 3.2)
Аспірант	О.А. Якушев (підрозділ 2.1)
Студент	Р.С. Волков (підрозділ 3.1)
Студент	Н.І. Андрієнко (підрозділ 2.5)
Студент	С.П. Занченко (підрозділ 2.5)

РЕФЕРАТ

Звіт: 106 с., 19 рисунків, 1 таблиця, 69 джерел.

Мета роботи – розроблення методологічних та теоретичних засад створення адаптивних систем керування складним технологічним процесом.

Об’єктом дослідження є слабо формалізований технологічний процес.

Предметом дослідження є моделювання, оцінка функціональної ефективності, алгоритми оптимізації та адаптації СППР, що навчається.

Даний звіт є проміжним.

На основі аналітичного огляду та аналізу тенденції розвитку інтелектуальних автоматизованих систем керування сформовано задачу аналізу та синтезу системи підтримки прийняття рішень (СППР), що навчається в режимі факторного кластер-аналізу. Розроблено науково-методологічні основи інформаційно-екстремального методу аналізу і синтезу СППР, що навчається в режимі факторного кластер-аналізу за умов апріорної невизначеності, інформаційних і ресурсних обмежень. Запропонований метод дозволяє побудувати безпомилкові за навчальною вибіркою вирішальні правила шляхом максимізації інформаційної спроможності СППР. Розроблено категорійні моделі та алгоритми оптимізації фенотопних і генотипних параметрів функціонування і досліджено вплив параметрів навчання на функціональну ефективність СППР.

АДАПТИВНА СИСТЕМА КЕРУВАННЯ, ОПТИМІЗАЦІЯ,
ФУНКЦІОНАЛЬНА ЕФЕКТИВНІСТЬ, ІНФОРМАЦІЙНИЙ КРИТЕРІЙ,
СИСТЕМА ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ, ТЕХНОЛОГІЧНИЙ
ПРОЦЕС, ІНФОРМАЦІЙНО-ЕКСТРЕМАЛЬНА ІНТЕЛЕКТУАЛЬНА
ТЕХНОЛОГІЯ, ПРОГНОЗУВАННЯ

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ	6
ВСТУП	7
РОЗДІЛ 1	
АНАЛІЗ ПРОБЛЕМИ ТА ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ	9
1.1 Системологічні питання проектування інтелектуальних СППР для автоматизації технологічних процесів	9
1.2 Аналіз та тенденція розвитку інтелектуальних СППР для автоматизації виробничих процесів.....	15
1.3. Аналітичний огляд методів автоматичної класифікації.....	24
1.4 Основні висновки та постановка задачі дослідження	32
РОЗДІЛ 2	
НАУКОВО-МЕТОДОЛОГІЧНІ ОСНОВИ АНАЛІЗУ І СИНТЕЗУ АСК, ЩО ФУНКЦІОНУЮТЬ В РЕЖИМІ ФАКТОРНОГО КЛАСТЕР-АНАЛІЗУ	36
2.1 Основні принципи, поняття та визначення інформаційно-екстремального факторного кластер-аналізу	36
2.2 Постановка задачі та математична модель інформаційно-екстремального факторного кластер-аналізу	41
2.3 Забезпечення статистичних стійкості й однорідності навчальної вибірки в процесі функціонування СППР в режимі ФКА.....	46
2.4 Критерії оптимізації процесу навчання СППР	50
2.5 Визначення мінімального обсягу репрезентативної навчальної вибірки.....	57
РОЗДІЛ 3	
ОПТИМІЗАЦІЯ ПАРАМЕТРІВ ФУНКЦІОНУВАННЯ АДАПТИВНОЇ АСКТП, ЩО НАВЧАЄТЬСЯ В РЕЖИМІ ФКА.....	62
3.1 Оптимізація фенотипних параметрів функціонування СППР, що навчається	62

3.2 Оптимізація контрольних допусків на ознаки розпізнавання	66
3.3 Максимізація міжцентрових відстаней для заданого алфавіту класів розпізнавання	76
3.4 Прогностичне навчання АСК виробництва фосфорної кислоти.....	89
ВИСНОВКИ.....	98
СПИСОК ЛІТЕРАТУРИ	100

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

- АСК – автоматизована система керування;
- АСКТП – автоматизована система керування технологічним процесом;
- БІР – блок індикації та реєстрації даних;
- БД – база даних;
- БЗ – база знань;
- БУ – блок керування;
- БКОІ – блок кінцевого оброблення інформації;
- БОЕ – блок оцінки ефективності;
- БП – блок прогнозування;
- БПОІ – блок первинного оброблення інформації;
- БПР – блок прийняття рішень;
- БСК – блок статистичної корекції;
- ЕПС – екстремальна порядкова статистика;
- ІЕІТ – інформаційно-екстремальна інтелектуальна технологія;
- КТЗ – комплекс технічних засобів;
- КФЕ – критерій функціональної ефективності;
- ОПР – особа, що приймає рішення;
- ППП – пристрій постійної пам'яті;
- СКД – система контрольних допусків;
- СМД – складні мінеральні добрива;
- СППР – система підтримки прийняття рішень.

ВСТУП

Підвищення функціональної ефективності керування виробничими процесами органічно пов'язано із розробкою та впровадженням інтелектуальних інформаційних технологій, що базуються на ідеях і методах машинного навчання, яке є важливим розділом теорії розпізнавання образів. Застосування у виробництві здатних навчатися (самонавчатися) автоматизованих систем керування (АСК), дозволяє здійснити перехід від застарілих суб'єктивних методів ручного керування до інтелектуальних інформаційних технологій керування розподіленими у просторі та часі слабо формалізованими технологічними процесами. При цьому виникає в теоретичному плані необхідність переходу від строгих традиційних математичних моделей до кібернетичних (евристичних) методів моделювання керованих процесів за умов апіорної невизначеності, нечітких даних і ресурсних обмежень, а в методологічному плані – перехід від методів операторного оброблення даних, що мають місце в лінійній теорії керування, до методів автоматичної класифікації на основі самонавчання й автоматичної класифікації.

Дослідження механізму прийняття рішень при керуванні технологічними процесами знайшло у науково-технічній літературі відносно широке висвітлення завдяки ідеям і науковим здобуткам, перед усім, представників українських наукових шкіл О. Г. Івахненка, В. І. Костюка, О. А. Павлова, В. І. Скуріхіна, В. Н. Васильєва, І. В. Кузьміна, Л. А. Пономаренка, М. Д. Кошового, Л. С. Ямпольського, А. П. Ладанюка та інших. При цьому науково-методологічні питання підвищення функціональної ефективності АСК, що навчаються, все ще залишаються недостатньо дослідженими. Так, основними недоліками відомих методів навчання систем керування, які роблять неефективним їх застосування на практиці, є ігнорування перетину класів розпізнавання у просторі ознак розпізнавання і

необхідність застосування навчальних вибірок великих обсягів. Указані вище методологічні та теоретичні ускладнення вирішення проблеми побудови високоефективних АСК, що навчаються, обумовлюють актуальність розроблення математичних моделей класифікаційного керування, методів оцінки функціональної ефективності та оптимізації просторово-часових параметрів функціонування, а так само відповідних засобів інформаційної технології проектування високоефективних АСК та широке їх впровадження в різні галузі соціально-економічної сфери України.

Науково-дослідна робота присвячена подальшому розвитку методів кластер-аналізу у рамках прогресивної інформаційно-екстремальної інтелектуальної технології (ІЕІ-технологія), яка ґрунтується на максимізації інформаційної спроможності адаптивних АСК шляхом введення на етапі навчання додаткових інформаційних обмежень. Наукове дослідження спрямовано на одержання нових результатів і удосконалення аналізу та синтезу адаптивних АСК, складовою частиною якої є інтелектуальна СППР, що навчається в режимі факторного кластер-аналізу, та розроблення засобів інформаційних технологій синтезу таких систем.

У першому розділі наведено аналітичний огляд методів аналізу та синтезу інтелектуальних СППР для керування складними технологічними процесами і постановку задачі синтезу СППР, що навчаються в режимі факторного кластер-аналізу.

Другий розділ присвячено висвітленню теоретичних і методологічних положень інформаційно-екстремального методу аналізу і синтезу СППР, що навчаються в режимі факторного кластер-аналізу.

У третьому розділі наведено результати оптимізації параметрів функціонування СППР, що навчаються в режимі факторного кластер-аналізу.

Одержанні при виконанні науково-дослідної роботи результати опубліковано у працях [1-19].

РОЗДІЛ 1

АНАЛІЗ ПРОБЛЕМИ ТА ПОСТАНОВКА ЗАДАЧІ
ДОСЛІДЖЕННЯ

1.1. Системологічні питання проектування інтелектуальних СППР для автоматизації технологічних процесів

Подальший розвиток інформаційного суспільства характеризується інтенсивним розвитком усіх галузей науки та техніки з якісно новим інтелектуальним технологічним забезпеченням, яке ґрунтується на комп'ютеризованих засобах праці й інформаційних технологіях і потребує принципово нового інформаційного середовища – сукупності ЕОМ, систем комп'ютерних комунікацій, баз даних, баз знань і програмних комплексів. При цьому важливого значення для автоматизації керованих процесів набуває аналіз і синтез адаптивних АСК складними технологічними процесами на основі самонавчання, розпізнавання образів та класифікаційного прогнозування за умов апріорної невизначеності, обумовленої впливом неконтрольованих вхідних факторів. Оскільки СППР є обов'язковою складовою адаптивної АСК, то важливого науково-методологічного значення набуває розроблення методів аналізу та синтезу здатних навчатися (самонавчатися) СППР та інформаційних засобів їх проектування. Інтелектуальні СППР виконуючи основні функції АСК слабо формалізованими процесами, допомагають особі, що приймає рішення, (ОПР) в аналізі ситуації, формулюванні задач, виробленні, контролі й оцінюванні варіантів рішень для досягнення поставленої мети.

СППР або DSS (Decision Support Systems) виконують більшість функцій АСКТП та інших автоматизованих АСК слабо формалізованими процесами. Оскільки в процесі керування ОПР вибирає варіант розв'язання задачі з множини можливих альтернатив із врахуванням принципів оптимальності та обмежувальних умов і несе відповідальність за своє рішення, то СППР допомагає ОПР в аналізі ситуації, формулюванні задач,

виробленні, контролі й оцінюванні варіантів рішень, що забезпечують досягнення поставленої мети. Якщо при аналізі ситуація вважається задовільною, то приймати рішення не потрібно. Якщо ж стан керованого процесу незадовільний, формується мета на основі певних критеріїв. З її урахуванням визначаються шляхи і засоби досягнення мети. Оскільки при виборі альтернатив аналізуються результати їх реалізації та прогнозуються наслідки, тому процес прийняття рішень має динамічний характер. На заключному етапі реалізується прийняте рішення й у взаємодії ОПР із СППР починається цикл ухвалення наступного рішення.

Архітектуру інтелектуальної СППР, побудованої за класифікаційним принципом керування, представлено на рис. 1.1.

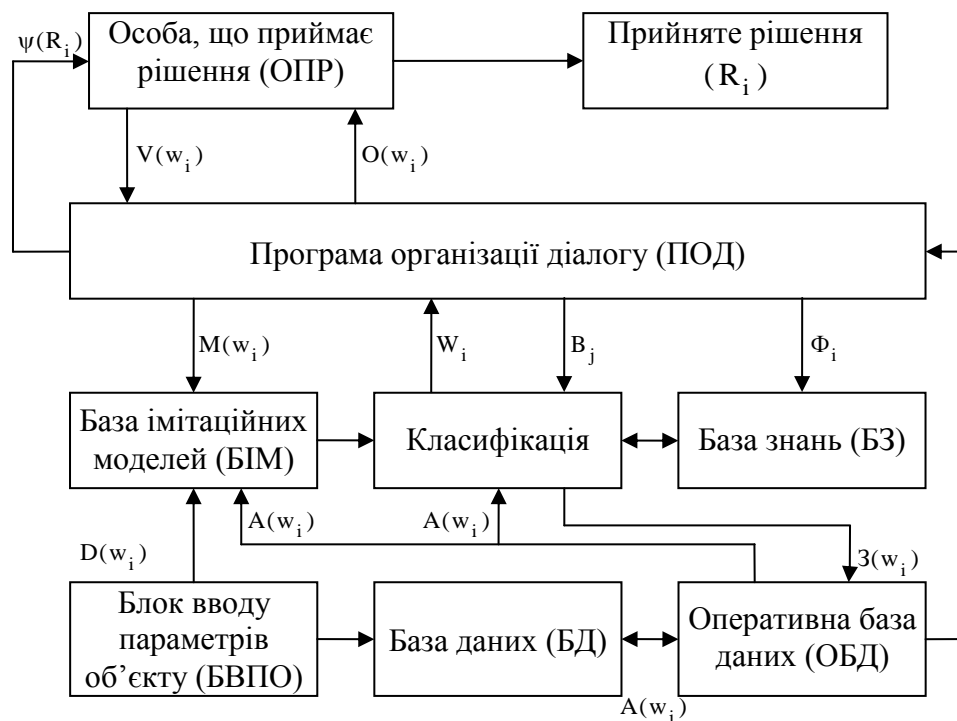


Рисунок 1.1 – Архітектура класифікаційної СППР

У класифікаційній СППР ОПР, якій надається право остаточного вибору та прийняття рішення R_i із допустимої множини $\psi(R_i)$, працює з системою через інтерфейс ПОД (програма організації діалогу) за допомогою

запитань $V(w_i)$ до користувача та відповідей $O(w_i)$. Вся множина завдань W_i , що розв'язує СППР, розбивається на три групи: завдання аналізу, синтезу й оптимізації. За бажанням ОПР вибирається одна з груп і вказується, який критерій ефективності функціонування об'єкта необхідно визначити. Відповідно до обраного завдання W_i ПОД вибирає необхідні моделі M_i у БІМ та робить класифікацію. На початку роботи СППР утвориться оперативна база даних (ОБД) (копія БД), у яку надходить запит із класифікатора для вибору вихідної інформації в базі імітаційних моделей (БІМ). Оперативна база даних (ОБД) може проглядатися й редагуватися ОПР за допомогою ПОД. Вихідна інформація надходить із ОБД у БІМ і класифікатор для вирішення обраного завдання W_i . Результат роботи класифікатора виводиться ПОД на екран дисплея у вигляді значень, таблиць, графіків залежностей показників функціонування від факторів, що впливають (B_j). Впливаючі фактори позначені на блок-схемі СППР інформаційним потоком $A(w_i)$. Оброблені результати роботи класифікатора після відповідної інтерпретації й визначення за розділами знань (Φ_i) записуються у базу знань за відповідним напрямом досліджень, які надалі будуть узагальнюватися і використовуватися при вирішенні завдань.

Інтелектуальні СППР, що оптимізують параметри функціонування, передбачають реалізацію багатоциклічної ітераційної процедури оптимізації структурованих просторово-часових параметрів функціонування АСК шляхом пошуку глобального максимуму усередненого за алфавітом класів розпізнавання значення критерію функціональної ефективності (КФЕ) навчання.

Якщо в експертних системах передбачається використання знань людини-експерта, то в інтелектуальних СППР перекладення функції експерта на ЕОМ здійснюється шляхом навчання (самонавчання) системи при забезпеченні технологічної гнучкості, яка характеризується властивостями переналагодження, багатомеклатурністю та оновленням продукції.

Для формулювання загальносистемних принципів побудови АСК, що навчається, необхідно ввести деякі поняття, що характеризують її структуру та функції. До основних загальносистемних принципів проектування АСК, що навчається, із урахуванням принципів, наведених у праці [20], відносяться:

- комплексність – реалізується побудовою системи, що охоплює всі галузі діяльності АСК, що навчається, фази (функції) керування, рівні організаційної структури, фази життєвого циклу, етапи оборту виробничих ресурсів та періоди планування;

- декомпозиція системи на елементи: підсистеми (цільові, функціональні, лінійного керування, забезпечення керування), спеціальні функції керування, задачі керування, модулі забезпечення окремих операцій технологічної структури;

- ієрархічність – здійснюється формуванням багаторівневої функціонально-лінійної структури АСК, що навчається відповідно до поділу на елементи і прикріплення повноважень (функцій) прийняття управлінських рішень відповідному рівню (органу) керування;

- цілеспрямованість – здійснюється формуванням цільових підсистем АСК, що навчається відповідно до основних галузей і цілей діяльності;

- замкненість циклу керування – здійснюється як у системі в цілому, так і в окремих її елементах повного циклу керування (прогнозування, планування, організація, координація, регулювання, активізація, стимулювання, урахування, контроль, аналіз);

- редукція простору ознак розпізнавання та замкнених роздільних гіперповерхонь класів розпізнавання, для яких будь-яким способом визначено геометричний центр і які за термінологією праці [21] далі будемо називати контейнерами класів розпізнавання;

- принцип максимізації інформації, який реалізується шляхом введення додаткових обмежень, що зменшують різноманітність об'єктів;

- принцип несумісності для складних систем, який утверджує, що складність системи і точність, із якою можна її можна проаналізувати математичними методами, у першому наближенні обернено пропорційні [22];

- принцип апріорної недостатності обґрунтування гіпотез (принцип Бернуллі-Лапласа), за яким виправдано прийняття рівноймовірних гіпотез, тобто прийняття рішень в СППР, що навчається, здійснюється за найгіршими у статистичному розумінні умовами її функціонування;

- рандомізація (приведення до випадковості) вхідних даних, що дозволяє разом з детермінованими характеристиками функціонального стану АСК, що навчається вводити у розгляд випадкові реалізації образів розпізнавання для обчислення інформаційної спроможності системи;

- принцип обмеженої багатоваріантності рішень, що приймаються, на якому базуються ітераційні процедури оптимізації просторово-часових параметрів функціонування СППР, що навчається у процесі максимізації інформаційного КФЕ.

До загальних принципів проектування СППР відносяться:

- принцип неперервного розвитку – полягає у тому, що система повинна будуватися так, що б поява нових засобів і сам розвиток системи не призводили до повної її перебудови;

- принцип модульності – полягає у тому, що система або підсистема повинна бути побудована з окремих автономних модулів, модифікація яких не призводить до поломки всієї системи, і кожний з яких може впроваджуватися і функціонувати окремо;

- відомі принципи системного підходу, які дозволяють сформулювати критерії досягнення мети, а на її основі сформулювати принципи побудови системи у цілому.

СППР складається з обслуговуючих та функціональних підсистем. Якщо функціональні підсистеми вирішують задачі керування, то обслуговуючі підсистеми призначаються для забезпечення ефективної

роботи функціональних підсистем. До обслуговуючих підсистем відносяться: підсистеми технічного забезпечення – комплекс технічних засобів (КТЗ), що включає засоби комунікації, зв'язку та інше; інформаційне забезпечення – довідники, класифікатори, кодифікатори, система документообігу, інформаційні масиви даних і методи технології оброблення інформації; програмне забезпечення; організаційне забезпечення – сукупність наказів, директив, інструкцій, вказівок, методичне забезпечення, лінгвістичне забезпечення – комплекс мовних засобів, юридичного та метрологічного забезпечення.

Керування розподіленим технологічним процесом полягає у розв'язанні таких задач:

- ідентифікація ситуації, у якій перебуває інтелектуальна СППР в поточний момент часу;
- прийняття рішень про відповідність поточного функціонального стану СППР раніше прийнятим плановим рішенням;
- вироблення оперативних змін раніше прийнятих планових рішень – регулювання;
- перспективне та нормативне прогнозування зміни функціонального стану інтелектуальної СППР;
- корекція моделі об'єкта, що використовується для виконання функцій керування, прогнозування, планування;
- планування та визначення погоджених дій підсистем інтелектуальної СППР на період часу, що планується.

Залежно від вхідних даних і уявлень про побудову та функціонування системи задачі синтезу поділяються на три класи [23]:

- інформаційний синтез, який передбачає оптимізацію (в інформаційному розумінні) функціональної ефективності системи;
- структурний синтез, який спрямований на визначення оптимального складу та конфігурації системи при заданих алгоритмах її функціонування;

– узагальнений синтез структури та алгоритмів функціонування, пов'язаний з розподілом функцій за елементами системи та визначенням їх оптимального складу.

Для інтелектуальної СППР інформаційний синтез набуває важливого науково-практичного значення, що викликано необхідністю створення високоефективних алгоритмів прийняття рішень у процесі керування слабо формалізованими процесами.

1.2 Аналіз та тенденція розвитку інтелектуальних СППР для автоматизації виробничих процесів

Сучасний стан автоматизації керованих складних технологічних процесів характеризується переходом від традиційних методів детермінованого керування до інтелектуальних технологій, здатних моделювати когнітивні процеси особи, що приймає керуючі рішення. При цьому важливого значення набуває проблема підвищення функціональної ефективності АСК розподіленими в просторі та часі складними технологічними процесами у виробництві за умов апріорної невизначеності та особливо відсутності через технічні та інші причини поточного вхідного контролю сировини і матеріал. Як приклад розглянемо технологічний процес виробництва складних мінеральних добрив у ВАТ “Сумихімпром”. На рис. 1.2 показано мнемо-схему технологічного процесу нейтралізації фосфорної та сірчаної кислот у процесі виготовлення пульпи, яка подається безпосередньо в барабанний гранулятор, де, крім того, здійснюється охолодження та сушка готового виробу.

Основна схема процесу нейтралізації кислот така. Розбавлені фосфорна та сірчана кислоти поступають відповідно із ємностей 1-5 і 1-1 в ємності 3-1 і 3-2, де вони змішуються. Із ємності 3-1 суміш поступає послідовно у сатуратори 4-1 і 4-2, а із ємності 3-2 – у сатуратори 4-6 і 4-5. У цих сатураторах відбувається перша стадія нейтралізації кислот (амонізація).

Утворена в процесі нейтралізації пульпа поступає в ємність 1-3, де здійснюється її лабораторний хімічний аналіз. Потім пульпа із позиції 1-3 поступає послідовно у сатуратори 4-4 і 4-5, де відбувається друга стадія нейтралізації з метою доведення азоту до заданого рівня концентрації в пульпі.

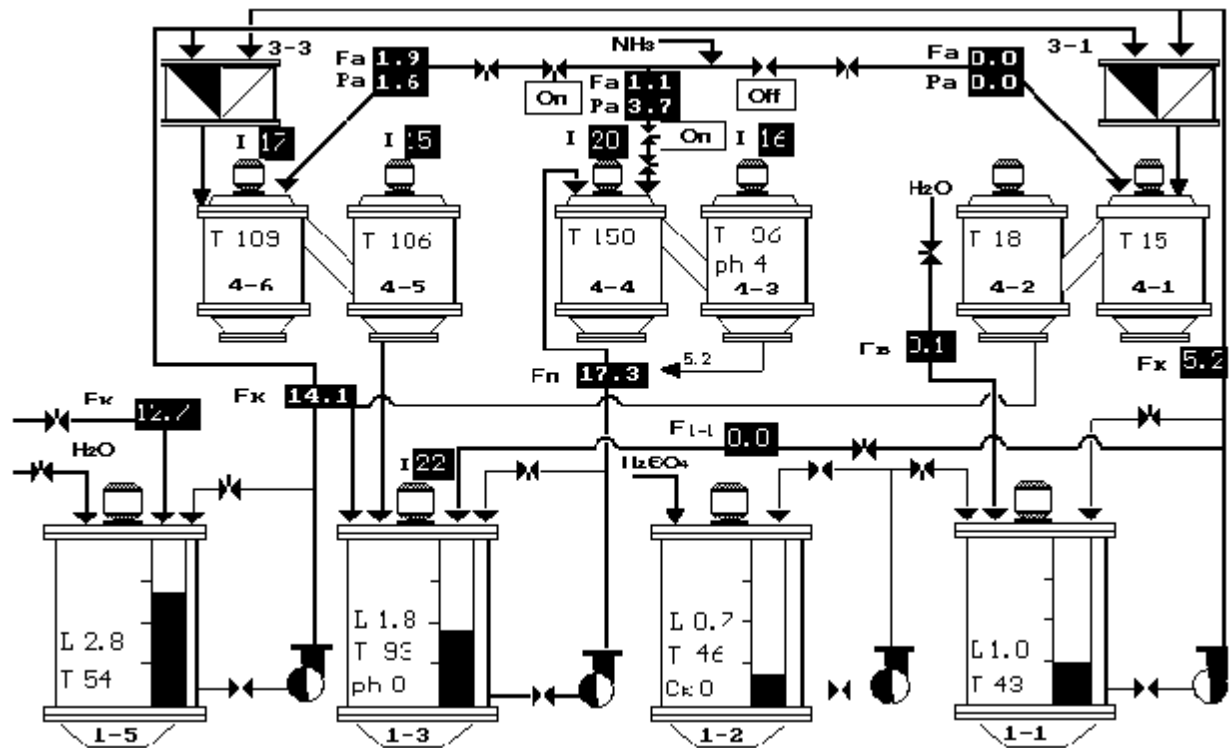


Рисунок 1.2– Мнемосхема технологічного процесу нейтралізації кислот

Після завершення другої стадії нейтралізації пульпа поступає в ємність 5-2, де так само здійснюється її лабораторний хімічний аналіз. Після корегування в пульпі змісту фосфору та калію, вона поступає в барабанный гранулятор, де набуває кінцевих показників якості СМД.

Незважаючи на те, що при виробництві СМД НКР використовується для локального керування окремими технологічними операціями 15 ПІД-регуляторів і частотних регуляторів, аналіз якості готового виробу показав достатньо високий відсоток відхилення змісту основних хімічних

компонентів та характеристик грануляції від існуючих стандартів. Основна причина такого незадовільного стану полягає у відсутності вхідного поточного контролю основних компонентів сировини – фосфорної та сірчаної кислот, що обумовлено як технічними причинами складності динамічного контролю характеристик кислот, які безперервно поступають в цех СМД із відповідних виробництв, так і суттєвим розкидом поточних вхідних характеристик цих кислот через відсутність вхідного контролю сировини.

Для прикладу проаналізуємо зміст P_2O_5 у фосфорній кислоті за результатами хіманалізу, взятими на виході її виробництва (п'ята ступінь екстрактора) протягом січня-лютого 2008 року (рис. 1.3).

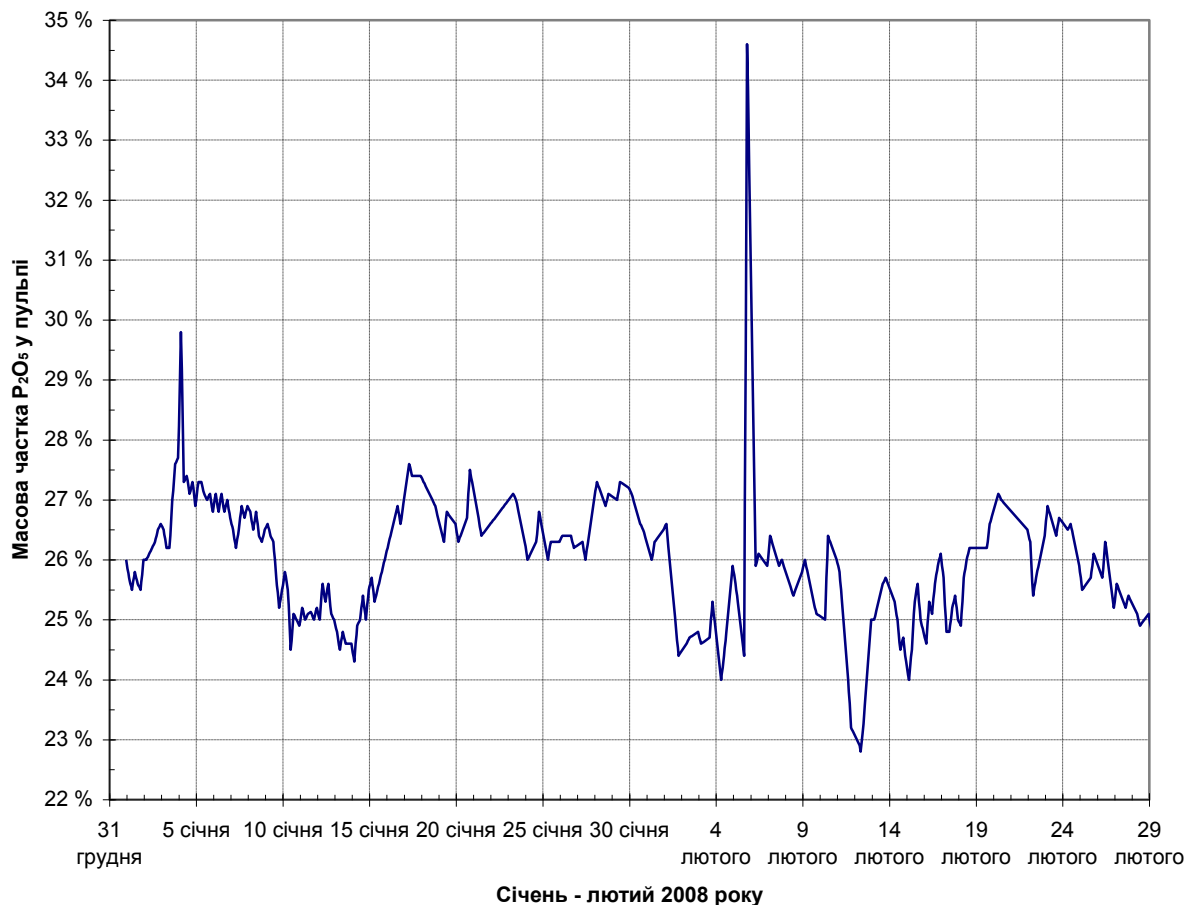


Рисунок 1.3 – Графік змісту P_2O_5 у пульпі на виході екстрактора за період січень-лютий 2008 р.

Аналіз рис.1.3 показує, що зміст $P_2 O_5$ є випадковою величиною, при цьому в періоди з 04.01.08 до 14.01.08 і з 12.02.08 до 21.02.08 процес виробництва фосфорної кислоти був взагалі нестационарним і, крім того, суттєво імплікативним (наприклад, випадки 04.01.08, 06.02.08, 11.2.08 і 12.02.08).

Таким чином, використання традиційних апаратно-програмних засобів автоматизації не дозволяє забезпечити адаптивність АСК слабо формалізованими розподіленими технологічними процесами за умов апріорної невизначеності, технічних, інформаційних і ресурсних обмежень.

Особливо актуальною проблема підвищення функціональної ефективності АСК є для керованих розподілених технологічних процесів за умов відсутності або технічної неможливості здійснення жорсткого вхідного контролю сировини, матеріалів та інших факторів.

Загальну постановку задачі оптимізації параметрів функціонування інтелектуальної СППР для автоматизації керованих процесів можна сформулювати так: є деякий процес, на якість перебігу якого можна деякою мірою впливати, вибираючи тим чи іншим способом залежні параметри x_1, x_2, \dots , а ефективність процесу характеризується показником J , який має бути перетворений в максимум (мінімум).

Для розв'язання прямої задачі математична модель процесу повинна дозволяти обчислити значення J за будь-яким вибраним рішенням і для будь-якої сукупності умов a_1, a_2, \dots . За умови, що відомі фактори a_1, a_2, \dots , на які впливати не можна (зокрема, відомі обмеження, що накладаються на рішення) і є залежні фактори (параметри) x_1, x_2, \dots , які в певних межах можна вибирати, показник ефективності процесу оптимізації має вигляд [18]

$$J = J(a_1, a_1, \dots; x_1, x_2, \dots). \quad (1.2.1)$$

Якщо вид залежності (1.2.1) відомий, то за допомогою такої математичної моделі можна обчислити J при будь-яких заданих значеннях $a_1, a_2, \dots, x_1, x_2, \dots$ (тобто пряма задача є розв'язаною).

Обернена задача формулюється так: за заданих умов a_1, a_2, \dots треба знайти такі параметри x_1, x_2, \dots процесу, які спрямовують показник J в максимум (мінімум). Це типова математична задача, яка належить до класу варіаційних задач. Методи розв'язання останніх є добре розробленими. Найпростіші з них – дуже відомі: вони потребують для визначення екстремуму функції диференціювання її за аргументами, прирівнювання похідних до нуля та розв'язування здобутої системи рівнянь як це показано, наприклад, у праці [22]. Однак, при збільшенні кількості аргументів задача пошуку екстремуму стає складнішою, при цьому екстремум функції J часто досягається не в точці, де похідна перетворюється в нуль (такої точки взагалі може не існувати), а десь на межі області змін аргументів. Виникають специфічні труднощі багатовимірної варіаційної задачі при наявності обмежень, яку через її складність не під силу розв'язати іноді навіть сучасній ЕОМ.

Крім того, якщо функція J взагалі не має похідних, а є цілочисловою чи заданою тільки при цілочислових значеннях її аргументів, то вибираються методи цілочислового програмування. Наприклад, якщо функція J лінійно залежить від параметрів x_1, x_2, \dots процесу, а обмеження, що накладаються на ці параметри, мають вигляд лінійних рівностей або нерівностей, виникає задача лінійного програмування. Якщо функція J є опуклою, то застосовуються спеціальні методи опуклого програмування з їх різновидом – методом квадратичного програмування.

Досі йшлося про задачі, в яких фігурують тільки дві групи факторів: умови a_1, a_2, \dots і параметри (елементи розв'язку) x_1, x_2, \dots процесу. Реальні ж задачі, що розв'язуються під час моделювання виробничих систем, частіше зводяться до таких, де наявною є третя група факторів – невідомі

неконтрольовані фактори f_1, f_2, \dots , значення яких не можна завбачити заздалегідь, і замість виразу (1.2.1) має місце

$$J = J(a_1, a_1, \dots; x_1, x_2, \dots; f_1, f_2, \dots). \quad (1.2.2)$$

У цьому випадку задачу оптимізації процесу може бути сформульовано так: за заданих умов a_1, a_2, \dots з урахуванням невідомих факторів f_1, f_2, \dots знайти такі значення параметрів (елементів розв'язку) x_1, x_2, \dots процесу, які забезпечують максимальне (мінімальне) значення показника ефективності J . Через наявність невідомих факторів задача оптимізації в цьому випадку набуває нової якості: вона перетворюється в задачу про вибір розв'язку за умов апріорної невизначеності.

На практиці доволі часто зустрічається задача автоматизації процесу керування, коли використовується не один показник ефективності J для оцінювання технологічного процесу, а має місце багатокритеріальна задача прийняття рішень [24]. У праці [25] запропоновано такий спосіб складання узагальненого показника ефективності: беруть окремі показники і приписують їм ваги (b_1, b_2, \dots) ; множать кожний показник на відповідну вагу і додають. Ті показники, що мають бути максимізовані, входять до суми зі знаком «плюс», а мінімізовані – зі знаком «мінус». Проблема лише в тому, що вагові коефіцієнти не є сталими, а залежать як від величин a_1, a_2, \dots , так і від професійного досвіду ОПР.

Історично постановка та розробка методів інформаційного синтезу для АСК спочатку відбувалася для класу лінійних динамічних систем як це мало місце, наприклад, у працях [26-28]. Загальним для цих праць є конструювання функціоналу, чутливого до впливу як детермінованої складової – оператора, так і інформаційної міри невизначеності на вході та виході системи:

$$\Phi\{W(s), J[p(\mu_m/\gamma_l)]\}, \quad (1.2.3)$$

де $W(s)$ – передатна функція системи;

$J[p(\mu_m/\gamma_l), t]$ – інформаційна міра, яка, у загальному випадку, є функцією часу t .

У виразі (1.2.3) $p(\mu_m/\gamma_l)$ – апостеріорна ймовірність появи події μ_m за умови прийняття гіпотези γ_l , де $m, l = \overline{1, M}$ – змінні числа функціональних станів системи. Тоді розв’язання задачі інформаційного синтезу полягає у знаходженні екстремальних параметрів функціонування системи $\{g_\xi\}$, $\xi = \overline{1, \Xi}$, які забезпечують максимум функціоналу

$$\Phi_m^* = \max_G \Phi\{W(s), J[p(\mu_m/\gamma_l), t]\},$$

де G – область допустимих значень параметрів функціонування.

Вперше задачу інформаційного синтезу АСК було розглянуто у 1971 році у монографії І. В. Кузьміна [29], в якій пропонувалося оцінювати ефективність системи за узагальненим функціонально-статистичним критерієм. Цей критерій враховує як інформаційну спроможність, яка визначає найважливішу складову загальної ефективності системи – функціональну ефективність, так і зведені витрати на забезпечення прийнятних показників інших складових ефективності. При цьому узагальнений функціонально-статистичний критерій ефективності має вигляд

$$E_{I,C} = K_I/K_{I0}, \quad (1.2.4)$$

де $K_I = I_{max}/C$ – узагальнена функціонально-статистична характеристика

робочої системи; $K_{10} = I_{\max}^0 / C_{\min}$ – узагальнена функціонально-статистична характеристика потенційної (ідеальної) системи; I_{\max} – максимальна інформаційна спроможність системи; C – зведені витрати на створення, експлуатацію та ліквідацію системи; I_{\max}^0 – максимальна інформаційна спроможність потенційної системи; C_{\min} – зведені витрати для потенційної системи.

Потенційна система повинна відповідати таким умовам:

- апріорні (безумовні) гіпотези є рівноймовірними, оскільки потенційна система повинна приймати рішення в найгірших у статистичному розумінні умовах ;

- умовні імовірності прийняття гіпотези γ_l при появі події μ_m , де $l, m = \overline{1, M}$, розподілені за нормальним законом, при якому інформаційна міра є максимальною;

Використання узагальненого критерію (1.2.4) є доцільним на етапі апріорного моделювання, коли оцінюється ступінь близькості системи, що проектується, до потенційної, і при структурному синтезі системи. При цьому має місце подвійність розв'язання цієї задачі [30]:

1) задана нижня допустима межа ефективності системи. Потрібно синтезувати систему із заданою ефективністю та мінімальною вартістю;

2) задана верхня межа вартості системи. Потрібно синтезувати систему із максимальною ефективністю, але не дорожче заданої вартості.

На етапі апостеріорного моделювання найбільшої актуальності набуває оцінка функціональної ефективності системи. Тому вартісна складова критерію ефективності може не розглядатися.

Незважаючи на те, що всі сучасні комп'ютерні системи мають убудовані модулі відліку часу, на практиці реалізуються різні принципи збирання та надходження інформації від зовнішніх джерел: періодичне опитування датчиків; незалежне (асинхронне) надходження даних у міру змін параметрів і комбінований підхід.

Інтелектуальні АСК реального часу застосовуються, як правило, в задачах автоматизованого керування слабо формалізованими процесами за умов апріорної невизначеності, нечітких даних, інформаційних і ресурсних обмежень. Організація замкнутого контуру із застосуванням інтелектуальної СППР в системах з оператором передбачає участь ОПР в замкнутому контурі керування; наявність клавіші «Виконання»; набір команд замкнутого керування; специфікацію передумов виконання команд і команди скасування дій.

За аналогією з працею [25] можна виділити ряд властивостей, які повинні мати інтелектуальні СППР (далі просто СППР), що працюють в реальному часі:

- динамічне призначення пріоритетів розв'язуваних задач;
- застосування нових ефективних методів попереднього оброблення інформації як з метою її нормалізації, так і скорочення обсягу та складності даних.
- оптимальне співвідношення між точністю розв'язання (рішення), з одного боку, і витратами (в тому числі й часовими) на його одержання;
- оцінка функціональної ефективності навчання СППР;
- адаптивність АСК за умов невизначеності, інформаційних і ресурсних обмежень;
- пошук в БЗ при наявності часових обмежень;
- використання методів виведення з урахуванням часових відношень

Основним недоліком СППР, що не здатна навчатися, є значне інформаційно-процедурне навантаження, яке припадає на ОПР. Ці навантаження пов'язані із рівнем професійної підготовки ОПР у конкретно-предметній галузі, проблемою подання (формалізації знань), недосконалістю механізму виведення нових знань на базі існуючих, багатofакторністю та нестационарністю слабо формалізованих процесів та об'єктів, які функціонують за умов невизначеності, і аналіз яких обмежується фізико-біологічними та психічними можливостями ОПР. Автори праць [31,32]

пов'язують подолання існуючих методологічних ускладнень в інженерії знань саме завдяки застосуванню методів і алгоритмів навчання, розроблених в теорії автоматичної класифікації.

Структура знань, які відображаються в інтелектуальній СППР, залежить від сфери їх застосування, містить різні факти з предметної галузі, взаємозв'язки між ними, правила дій та інше. Вона має містити так само знання, що стосуються способу включення знань в СППР. Складність і подібність структур знань спричинили появу кількох різних способів подання знань, з яких слід виділити логічні методи [33], фреймові [34] та продукційні системи [35], семантичні мережі [36]. Кожен із цих способів має свої переваги та недоліки і тяжіє до певної структури знань.

Особливість систем подання знань полягає в тому, що вони моделюють діяльність людини, яка часто здійснюється в неформальному вигляді. Якщо методи розв'язування обчислювальних задач ґрунтуються на чітких алгоритмах, побудованих з використанням поняття збігу, то в основі моделей подання знань лежить інформація, здобута від експертів, яка часто має якісний і до того ж суперечливий характер. Однак, виходячи зі специфіки функціонування ЕОМ, така інформація має бути зведеною до однозначного формального вигляду, що здійснюється завдяки ідеям багатозначної логіки, теорії нечітких множин і аналітичних математичних моделей.

1.3. Аналітичний огляд методів автоматичної класифікації

Становлення теорії автоматичної класифікації, починаючи з другої половини ХХ століття, відбувається, головним чином, шляхом подальшого розвитку ідей і методів як розпізнавання образів [37-39], так і статистичної теорії прийняття рішень [40].

В основу класичних статистичних методів розпізнавання образів покладено методи та критерії статистичної теорії прийняття рішень [41-43]. Найбільш суттєві результати аналізу та синтезу складних систем за

статистичними методами одержано у тих галузях науки і техніки, де є можливість набору репрезентативної статистики за умови забезпечення статистичних однорідності та стійкості експериментальних даних. В АСК, які містять контур компаративного розпізнавання зображень за еталоном, знайшли широке використання, наприклад, при ідентифікації кадрів та у задачах навігації кореляційно-екстремальні методи розпізнавання образів [44], які як порогові методи відрізняються високою заводо захищеністю. При цьому системи, які реалізують кореляційно-екстремальні методи розпізнавання, вимагають значних обчислювальних ресурсів, оскільки необхідно зберігати у пам'яті АСК набір еталонів, що перекривають усі можливі зміни на зображеннях. Але основним недоліком кореляційно-екстремальних методів, є відсутність адаптивного вибору порогу порівняння кореляційної функції або згортки. Аналіз сучасного стану розробок в області створення систем компараторного розпізнавання і спостереження за різного роду об'єктами у реальному масштабі часу показує явну тенденцію поступового переходу від систем, заснованих на кореляційних підходах, до систем на основі навчання (самонавчання). Розвиток теорії машинного навчання, основи якої закладено у працях [45-47], відбувався в рамках статистичного підходу до розпізнавання образів, що є його головною заслугою. Зміст усіх статистичних методів навчання полягає у знаходженні роздільної функції, спосіб побудови якої задає спосіб розбиття простору ознак на класи розпізнавання.

Об'єднати переваги як детермінованого, так і статистичного підходів до розпізнавання образів та обійти їх недоліки – така задача розв'язується у рамках детерміновано-статистичного підходу, наприклад, у праці [48], де запропоновано будувати класифікатор у два етапи. На першому етапі будується найбільш просте детерміноване вирішальне правило, яке забезпечує безпомилкове розпізнавання реалізацій навчальної матриці. Для мінімізації кількості помилок і прогнозу їх величин при розпізнаванні на другому етапі здійснюється непараметрична статистична корекція

вирішального правила. Для такої корекції використовують два одномірних параметри, оцінку яких можна отримати за класифікованою навчальною матрицею. Це максимально можливий викид реалізацій за “свою” еталонну оболонку та максимально-можливе проникнення у глибину еталонних оболонок реалізацій “чужих” образів. Подальшим розвитком методу еталонів [48] став метод еталонного опису [49], в якому з метою зменшення кількості еталонів, а отже і спрощення мінімізації покриття класу, акцент зроблено на відбір ефективних ознак розпізнавання шляхом оцінки їх корисності. Модельність як методу еталонів, так і методу еталонного опису полягає в ігноруванні загального випадку перетину класів. А загальним недоліком цих та інших відомих методів розпізнавання, який знижує їх практичну цінність, є відсутність алгоритмів оптимізації процесу навчання за прямим КФЕ.

Основний недолік більшості відомих методів автоматичної класифікації полягає в ігноруванні перетину класів розпізнавання, що має місце на практиці в задачах керування слабо формалізованими процесами з нечіткими даними. Тому останнім часом спостерігається збільшення уваги до розробки алгоритмів автоматичної класифікації у рамках теорії нечітких множин [50]. Але при цьому все ще поза увагою досліджень залишається таке важливе питання, як оптимізація просторово-часових параметрів навчання інтелектуальної системи.

В останні роки значного поширення в задачах адаптивного пошуку та багатопараметричної оптимізації набули методи, які базуються на генетичних алгоритмах [51-53]. Ці алгоритми, які є методами випадкового пошуку глобального екстремуму критерію оптимізації, копіюють механізми як генетичної спадковості, так і природного відбору. Але дійсна практична цінність генетичних алгоритмів полягає у поєднанні їх з іншими методами розпізнавання образів, які безпосередньо вирішують проблему побудови оптимального класифікатора.

Незважаючи на великий обсяг публікацій у галузі автоматичної класифікації, значних зрушень у підвищенні ефективності навчання АСК

слабо формалізованими процесами все ще не відбувається. Основними причинами такого стану є:

- модельний характер переважної більшості відомих методів автоматичної класифікації, що становить певну теоретико-методологічну цінність, але є непридатним для практичного застосування;
- незавершеність теорії ефективного машинного навчання інтелектуальних систем;
- нелінійність, нестационарність і багатофакторність виробничих технологічних процесів, які відбуваються за умов апріорної невизначеності, що й обумовлює їх слабку формалізованість.

Значним досягненням за останні роки у розвитку теорії автоматичної класифікації стало обґрунтування В. І. Васильєвим принципу редукції простору ознак розпізнавання, що обумовлює доцільність побудови відносно простих вирішальних правил [54].

На усунення недоліків відомих методів автоматичної класифікації спрямовано інформаційно-екстремальну інтелектуальну технологію (ІЕІтехнологія) аналізу та синтезу здатних навчатися АСК [31], яка ґрунтується на прямій оцінці інформаційної спроможності системи і гіпотезі нечіткої компактності реалізацій образу, що допускає перетин класів розпізнавання.

Основні концептуальні положення ІЕІ-технології такі:

- ІЕІ-технологія ґрунтується на прямій оцінці інформаційної спроможності АСК, що навчається;
- ІЕІ-технологія дозволяє оптимізувати просторово-часові параметри функціонування СППР, що навчається;
- прийняття рішень у рамках ІЕІ-технології здійснюється в рамках детерміновано-статистичного підходу шляхом побудови відносно простого детермінованого класифікатора, статистична корекція якого здійснюється в процесі навчання з метою побудови безпомилкових, за навчальними матрицями, вирішальних правил;

- методи ІЕІ-технології ґрунтуються на застосуванні гіпотез як чіткої, так і нечіткої компактності реалізацій образу, тобто є працездатним за умови перетину класів розпізнавання; що має місце в практичних задачах контролю та керування;

- методи ІЕІ-технології є об'єктно-структурованими, що дозволяє їх розвивати для вирішення проблеми інформаційного синтезу широкого класу АСК, що навчаються, з успадкуванням властивостей структурованих об'єктів вищого ієрархічного рівня і до визначенням їх методів;

- методи ІЕІ-технології базуються на вибірковому підході математичної статистики і орієнтований на застосування прийнятних з практичних міркувань обсягів репрезентативних навчальних вибірок;

- методи ІЕІ-технології є універсальними для проектування здатної навчатися АСК будь-якої природи і дозволяють вирішувати як загальну, так і частинні задачі її інформаційного синтезу.

Оскільки доцільність побудови “точного” контейнера складної геометричної форми залишається дискусійною через значні обчислювальні ускладнення при реалізації такого контейнера і випадковість значень ознак розпізнавання на екзамені, то виглядає перспективною ідея про доцільність у рамках детерміновано-статистичного підходу побудови на етапі навчання СППР усередненого контейнера спрощеної форми, наближеного до “точного” деяким оптимальним способом і який забезпечує на екзамені достовірність розпізнавання, близьку до максимальної асимптотичної.

У загальному випадку математичну (категорійну) модель будь-якої АСК, яка розв'язує задачу класифікаційного аналізу, подамо, аналогічно праці [23], у вигляді діаграми відображень відповідних множин:

$$\begin{array}{ccccccc}
 \mathbf{G} \times \mathbf{T} \times \mathbf{\Omega} \times \mathbf{Z} & \xrightarrow{\Phi} & \mathbf{Y} & \xrightarrow{\theta_1} & \mathbf{X} & \xrightarrow{\theta_2} & \tilde{\mathcal{R}}^{|\mathbf{M}|} & \xrightarrow{\Psi_H} & \mathbf{I}^{//} \\
 & & \searrow \nu & & \swarrow \varkappa & & \uparrow \xi & & \\
 & & & & \mathbf{L} & & & &
 \end{array}
 , \quad (1.3.1)$$

де G – простір вхідних сигналів (факторів), які діють на СППР; T – множина моментів часу зняття інформації; Ω – простір ознак розпізнавання; Z – простір можливих станів АСК; Y – вибіркова множина (вихідна навчальна матриця $\|y_{mi}^{(j)}\|$); $\Phi: G \times T \times \Omega \times Z \rightarrow Y$ – оператор формування вибіркової множини Y на вході СППР; X – вибіркова множина, яка утворює экзаменаційну матрицю $\|x_i^{(j)} | i = \overline{1, N}, j = \overline{1, n}\|$, аналогічну за структурою, процедурою та параметрами формування навчальній матриці; $\tilde{\mathfrak{R}}^{|M|}$ – нечітке розбиття ($\tilde{\mathfrak{R}}^{|M|} \subset \Omega$); ψ_H – оператор класифікації; $I^{|l|}$ – множина допустимих гіпотез (рішень); l – кількість статистичних гіпотез; L – покриття, що будується для заданого алфавіту класів розпізнавання $\{X^o\}$; $v: Y \rightarrow L$ – оператор побудови апіорного покриття L ; $\alpha: \mathfrak{R}^{|M|} \rightarrow L$ – оператор побудови апостеріорного покриття L ; ξ – оператор зворотного зв'язку; θ_1 – оператор відображення вибіркової множини Y на X ; θ_2 – оператор відображення вибіркової множини X на нечітке розбиття $\tilde{\mathfrak{R}}^{|M|}$.

У діаграмі (1.3.1) оператор зворотного зв'язку ξ корегує геометричні параметри розбиття $\tilde{\mathfrak{R}}^{|M|}$ з метою покращення точнісних характеристик СППР. Саме спосіб реалізації оператора ξ породжує різні методи та підходи до розв'язання проблеми адаптивного навчання АСК.

При обґрунтуванні гіпотези нечіткої компактності має місце нечітке розбиття $\tilde{\mathfrak{R}}^{|M|} \subset \Omega$. Застосуємо оператор θ нечіткої факторізації простору ознак: $\theta: Y \rightarrow \tilde{\mathfrak{R}}^{|M|}$. Нехай оператор класифікації $\psi_H: \tilde{\mathfrak{R}}^{|M|} \rightarrow I^{|l|}$ перевіряє основну статистичну гіпотезу про належність реалізацій $\{x_m^{(j)} | j = \overline{1, n}\}$ класу X_m^o . Оператор $\gamma: I^{|l|} \rightarrow \mathfrak{S}^{|q|}$ шляхом оцінки статистичних гіпотез формує множину точнісних характеристик $\mathfrak{S}^{|q|}$, де $q = l^2$ – кількість точнісних характеристик. Оператор $\varphi: \mathfrak{S}^{|q|} \rightarrow E$ обчислює множину значень інформаційного КФЕ, який є функціоналом точнісних характеристик. Контур оптимізації геометричних параметрів розбиття $\tilde{\mathfrak{R}}^{|M|}$ шляхом пошуку

максимуму КФЕ навчання розпізнаванню реалізацій класу X_m^o замикається оператором $r: \mathbf{E} \rightarrow \tilde{\mathfrak{R}}^{|\mathcal{M}|}$. Структурна діаграма процесу навчання за базовим алгоритмом ІЕІ-технології для апріорно нечіткого розбиття має вигляд [30]:

$$\begin{array}{c}
 \xrightarrow{\quad U \quad} \mathbf{E} \xleftarrow{\quad \varphi \quad} \mathfrak{Z}^{|\mathcal{Q}|} \\
 \uparrow \quad \uparrow \\
 \mathbf{G} \times \mathbf{T} \times \mathbf{\Omega} \times \mathbf{Z} \xrightarrow{\quad \Phi \quad} \mathbf{Y} \xrightarrow{\quad \theta_1 \quad} \mathbf{X} \xrightarrow{\quad \theta_2 \quad} \tilde{\mathfrak{R}}^{|\mathcal{M}|} \xrightarrow{\quad \Psi_H \quad} \mathbf{I}^{||} \\
 \uparrow \quad \uparrow \\
 \mathbf{r} \quad \gamma
 \end{array}
 \quad (1.3.2)$$

При цьому у діаграмі (1.3.2) контур операторів

$$\boxed{\Psi_H \longrightarrow \gamma \longrightarrow \varphi \longrightarrow r}
 \quad (1.3.3)$$

безпосередньо оптимізує геометричні параметри розбиття $\tilde{\mathfrak{R}}^{|\mathcal{M}|}$. Оператор $U: \mathbf{E} \rightarrow \mathbf{G} \times \mathbf{T} \times \mathbf{\Omega} \times \mathbf{Z}$ регламентує процес навчання і дозволяє оптимізувати параметри його плану, які визначають, наприклад, обсяг і структуру випробовувань, черговість розгляду класів розпізнавання та інше. Серед параметрів навчання, які суттєво впливають на достовірність класифікатора, розглядаються поля контрольних допусків $\{\delta_{k,i} \mid i = \overline{1, N}\}$ на значення ознак розпізнавання, рівні селекції $\{\rho_m\}$ координат еталонних двійкових векторів і крок квантування в часі τ реалізацій образу.

Діаграма відображень множин на екзамені має такі відмінності від діаграм оптимізаційного навчання в рамках ІЕІ-технології:

- зворотний зв'язок у діаграмі не містить контурів оптимізації параметрів функціонування АСК, а призначенням оператора U_E є регламентація екзамену;

- замість оператора θ вводиться оператор P відображення вибіркової множини $\mathbf{X} \subset \mathbf{\Omega}_B$, що розпізнається, на побудоване на етапі навчання розбиття $\tilde{\mathfrak{R}}^{|\mathcal{M}|}$;

- комутативне кільце утворюється між розбиттям $\tilde{\mathfrak{R}}^{|M|}$, множиною гіпотез $I^{|M+1|}$ і покриттям $L^{|M|}$;

- оператор класифікації ψ утворює композицію двох операторів: $\psi_1: \tilde{\mathfrak{R}}^{|M|} \rightarrow F$, де F – множина функцій належності, і оператор $\psi_2: F \rightarrow I^{|M+1|}$, який вибирає гіпотезу за максимальним значенням функції належності.

З урахуванням наведених відмінностей діаграма відображень множин на екзамени набуває вигляду

$$\begin{array}{c}
 \xrightarrow{\hspace{10em} U_E \hspace{10em}} \\
 \mathbf{G} \times \mathbf{T} \times \mathbf{\Omega} \times \mathbf{Z} \xrightarrow{\Phi} \mathbf{X} \xrightarrow{P} \tilde{\mathfrak{R}}^{|M|} \xrightarrow{\psi_1} \mathbf{F} \xrightarrow{\psi_2} \mathbf{I}^{|M+1|} \\
 \searrow v \qquad \swarrow \varepsilon \\
 \mathbf{L}
 \end{array}
 \quad . \quad (1.3.4)$$

У діаграмі (1.3.4) оператор Φ відображає універсум випробувань на вибірку множини X , яка утворює бінарну екзаменаційну матрицю $\|x_i^{(j)} \mid i = \overline{1, N}, j = \overline{1, n}\|$, аналогічну за структурою та параметрами формування навчальній матриці.

Суттєвою перевагою категорійних моделей у вигляді діаграм відображень множин типу (1.3.2) є те, що вони дозволяють на етапі системного аналізу не тільки встановлювати відношення між елементами інформаційного забезпечення та інформаційними потоками оброблення інформації, але і суттєво полегшують розробку структур алгоритмів різних режимів функціонування АСК, що навчається.

Таким чином, аналіз сучасного стану автоматизації розподілених технологічних процесів показує, що традиційні апаратно-програмні методи за детермінованими алгоритмами керування не забезпечують адаптивність АСК за умов апіорної невизначеності, а переважна більшість відомих методів автоматичної класифікації носить модельний характер, оскільки вони орієнтовані апіорно на чітке розбиття класів розпізнавання. Як один із перспективних напрямів аналізу і синтезу інтелектуальних АСК є ІЕІТ, що

ґрунтується на максимізації інформаційної спроможності СППР шляхом введення на етапі навчання додаткових інформаційних обмежень.

1.4 Основні висновки та постановка задачі дослідження

У результаті аналізу сучасного стану існуючих методів автоматичної класифікації, які застосовуються в АСК, що навчаються, та тенденції їх розвитку можна зробити такі висновки:

1. Аналіз керованих складних технологічних процесів (на прикладі виробництва складного мінерального добрива НКР у ВАТ “Сумхімпром”) показав, що застосування ПД-регуляторів, частотних регуляторів та інших сучасних апаратно-програмних засобів керування дозволяє вирішувати питання керування локальними процесами, в яких існує однозначна залежність між входом і виходом, але не забезпечує високу функціональну ефективність керування розподіленими в просторі і часі технологічними процесами, що відбуваються за умов апріорної невизначеності, нечітких даних, інформаційних і ресурсних обмежень.

2. На практиці в задачах контролю та керування слабо формалізованими технологічними процесами існують ускладнення, пов’язані, по-перше, з забезпеченням статистичної сталості вхідних даних і, по-друге, безпосередньо з формалізацією процесу навчання, що обмежує використання відомих параметричних і непараметричних методів автоматичної класифікації, які ґрунтуються на теорії статистичних рішень.

3. Алгоритми автоматичної класифікації, розроблені з використанням елементів теорії нечітких множин дозволяють усунути основний недолік класичних методів автоматичної класифікації, який полягає в ігноруванні перетину класів розпізнавання, що має місце на практиці в задачах керування слабо формалізованими процесами з нечіткими даними, але мають принципові ускладнення при конструюванні критеріїв оптимізації параметрів функціонування інтелектуальної СППР.

4. Застосування ідей теорії редукції, створеної українською школою В.І.Васильєва, надає можливість зокрема реалізувати механізм цілеспрямованої трансформації розподілів векторів-реалізацій образів у просторі ознак для побудови відносно простих вирішальних правил, відкриває перспективу вирішення у рамках єдиної методології проблеми синтезу високо достовірних алгоритмів прийняття рішень і оптимізації характеристик функціонування АСК, що навчаються.

5. До теперішнього часу найважливіша проблема інформаційного синтезу АСК, що навчаються – оптимізація їх просторово-часових характеристик за оцінками точнісних характеристик СППР все ще залишається в науково-технічній літературі мало дослідженою.

6. Одним із нових перспективних напрямів аналізу і синтезу здатних навчатися (самонавчатися) АСКТП є застосування розробленої колективом науково-дослідницької лабораторії інтелектуальних систем Сумського державного університету під науковим керівництвом професора Довбиша А.С. ІЕІ-технології, що ґрунтується на максимізації інформаційної спроможності системи шляхом введення на етапі навчання додаткових інформаційних обмежень [31].

На основі аналізу сучасного стану проектування та використання інтелектуальних СППР для автоматизації розподілених в просторі і часі технологічних процесів і мети дисертаційної роботи сформуємо узагальнену задачу дослідження в дисертаційній роботі.

Нехай відомі алфавіт класів розпізнавання $\{X_m^o \mid m = \overline{1, M}\}$, який характеризує M функціональних станів технологічного процесу, і навчальна матриця типу «об'єкт-властивість» $\|y_{m,i}^{(j)}\|, i = \overline{1, N}, j = \overline{1, n}$, де N, n – кількість ознак розпізнавання і випробувань відповідно. Нехай базовий клас X_1^o характеризує найбільш бажаний функціональний стан технологічного процесу. Необхідно:

1) На етапі навчання побудувати оптимальне нечітке розбиття $\tilde{\mathfrak{R}}^{|M|}$, $M = \text{Card } \mathfrak{R}^{|M|}$, бінарного простору ознак Ω_B на класи розпізнавання за умов:

$$(\forall X_m^o \in \tilde{\mathfrak{R}}^{|M|}) [X_m^o \neq \emptyset]; \quad (1.4.1)$$

$$(\exists X_k^o \in \tilde{\mathfrak{R}}^{|M|}) (\exists X_l^o \in \tilde{\mathfrak{R}}^{|M|}) [X_k^o \neq X_l^o \rightarrow X_k^o \cap X_l^o \neq \emptyset]; \quad (1.4.2)$$

$$(\forall X_k^o \in \tilde{\mathfrak{R}}^{|M|}) (\forall X_l^o \in \tilde{\mathfrak{R}}^{|M|}) [X_k^o \neq X_l^o \rightarrow \text{Ker} X_k^o \cap \text{Ker} X_l^o = \emptyset]; \quad (1.4.3)$$

$$(\forall X_m \in \mathfrak{R}^{|\Lambda|}) (\forall X_c \in \mathfrak{R}^{|\Lambda|}) [\text{if } E_m \in G_E \text{ then } E_m^* = \max_{\{m\}} E_m], m, c = \overline{1, M}; \quad (1.4.4)$$

$$\bigcup_{X_m^o \in \tilde{\mathfrak{R}}} X_m^o \subseteq \Omega_B. \quad (1.4.5)$$

де E_m – інформаційний критерій функціональної ефективності (КФЕ)

навчання СППР розпізнавати реалізації класу X_m^o .

2) На етапі екзамену (безпосередньо у робочому режимі СППР) за максимальним значенням функції належності дефазифікувати функціональний стан АСКТП.

3) За умови невідповідності поточного функціонального стану технологічного процесу класу X_1^o реалізувати оператор корекції для стабілізації АСКТП.

Для досягнення поставленої в дисертаційній роботі мети необхідно вирішити такі **завдання**:

- формування вхідного математичного опису інтелектуальної СППР (далі під терміном СППР буде розумітися тільки інтелектуальна СППР, що здатна навчатися (самонавчатися));

- розроблення математичних моделей і ефективних алгоритмів оптимізації параметрів функціонування СППР, що навчається в режимі факторного кластер-аналізу;

- оцінка функціональної ефективності навчання і дослідження впливу параметрів навчання на функціональну ефективність СППР в режимі кластер-аналізу;
- корекція асимптотичних точнісних характеристик СППР в процесі навчання;
- визначення мінімального обсягу репрезентативної навчальної вибірки;
- оцінка статистичної стійкості та статистичної однорідності навчальної вибірки для технологічних режимів;
- розроблення засобів інформаційної технології для проектування СППР, що навчається в режимі кластер-аналізу за умов апіорної невизначеності та ресурсних обмежень.

РОЗДІЛ 2

НАУКОВО-МЕТОДОЛОГІЧНІ ОСНОВИ АНАЛІЗУ І СИНТЕЗУ АСК, ЩО ФУНКЦІОНУЮТЬ У РЕЖИМІ ФАКТОРНОГО КЛАСТЕР-АНАЛІЗУ

2.1 Основні принципи, поняття та визначення інформаційно-екстремального кластер-аналізу

У практичних задачах класифікаційного керування розподілених у часі і просторі слабо формалізованих технологічних процесів виникає необхідність донавчання АСК, що обумовлено появою нових функціональних станів керованого процесу, які не було враховано на стадії апріорного навчання. Ця проблема неодмінно виникає при формуванні відкритих інформаційних фондів, наприклад, у задачах медичного діагностування, розпізнавання електронограм матеріалів різних зразків, а так само при керуванні слабо формалізованими технологічними процесами, які змінюють свій функціональний стан під впливом різних контрольованих та неконтрольованих факторів. У працях [55-56] розглядалися задачі кластер-аналізу, заснованого на відношеннях подоби з метою побудови в просторі ознак розпізнавання нових представництв, що є афінними різноманіттями, навколо яких агрегатуються нові класи. Запропонований і розвинений у цих роботах підхід до розв'язання задачі кластер-аналізу носить модельний характер, оскільки базується на гіпотезі чіткої компактності реалізацій образу і відповідно на дистанційних критеріях розпізнавання образів.

На практиці в задачах автоматизації технологічних процесів через вплив на АСКТП зовнішніх і внутрішніх випадкових неконтрольованих факторів, що призводить до довільних початкових умов в моменти зняття інформації із датчиків сигналів виправданою є гіпотеза нечіткої компактності. Це обумовлює необхідність нормалізації апріорно деформованих образів з метою застосування їх для подальшого навчання інтелектуальної СППР, що потребує значних матеріальних витрат і відповідного інформаційного забезпечення. Одними із ефективних методів

автоматичної класифікації, як показано у першому розділі, є інформаційно-екстремальні методи, які дозволяють здійснювати нормалізацію образів безпосередньо в процесі навчання системи шляхом цілеспрямованої ітераційної процедури пошуку глобального екстремуму інформаційного КФЕ навчання АСК [31]. Об'єктно-орієнтована концепція ІЕІ-технології дозволяє реалізувати в рамках єдиної методології аналізу і синтезу СППР як алгоритми навчання за апріорно класифікованими навчальними матрицями (навчання з „учителем”), так і алгоритми кластер-аналізу за апріорно некласифікованими навчальними матрицями. Це дозволяє при розробці здатних навчатися АСК різного призначення досягти в рамках ІЕІ-технології спадковості, поліморфізму та віртуалізації процедур, що на практиці спрощує розробку складних програмних систем класифікаційного керування.

Основними принципами кластер-аналізу у рамках ІЕІ-технології є:

- принцип максимізації інформації, який реалізується шляхом введення додаткових інформаційних і ресурсних обмежень [31]
- принцип дуальності оптимального керування А.О. Фельдбаума за яким апостеріорна модель АСК удосконалюється в процесі навчання;
- принцип зовнішнього доповнення [57];
- розширений принцип редукції, який полягає у цілеспрямованому спрощенні як простору ознак розпізнавання, так і гіпотетично існуючої найкращої роздільної функції складного вигляду у роздільну функцію більш простого вигляду, реалізація якої не викликає значної обчислювальної трудоемності;
- принцип квантовості подання та виведення знань І.Б. Сіроджи [58], що обумовлює необхідність покрокового оброблення інформації в процесі навчання СППР;
- принцип апріорної недостатності обґрунтування гіпотез (принцип Бернуллі-Лапласа), за яким для оцінки ефективності функціонування АСК, що навчається, виправдано прийняття рівноймовірних гіпотез, оскільки апріорна інформація є неповною;

- принцип “найближчого сусіди”, який згідно з гіпотезою компактності (чіткою або нечіткою) виключає вплив реалізацій далеких класів на геометричні параметри контейнерів, що оптимізуються на етапі навчання АСК;

- принцип K – найближчих сусідів [59], який реалізується безпосередньо на етапі прийняття рішень – екзамені;

- принцип рандомізації (принцип приведення до випадковості) вхідних даних. Цей принцип дозволяє оцінювати точнісні характеристики процесу навчання і обчислювати інформаційну спроможність АСК, що навчається за статистичними інформаційними критеріями;

- дистанційно-мінімальний принцип, який обумовлює необхідність мінімізації середньої відстані реалізацій образу від центру їх розсіювання;

- дистанційно-максимальний принцип середньої міжкласової відстані для заданого алфавіту класів розпізнавання.

Зрозуміло, що на методи ІЕІ-технології поширюються й інші принципи теорії прийняття рішень і розпізнавання образів [60].

Реалізація дистанційно-мінімального принципу при відновленні в процесі навчання в радіальному базисі оптимальних в інформаційному розумінні контейнерів класів розпізнавання вимагає доповнення умови оптимізації апріорного нечіткого розбиття дискретного простору ознак на класи розпізнавання (1.4.3) таким обмеженням:

$$\begin{aligned} & (\forall X_k^o \in \tilde{\mathfrak{R}}^{|M|}) (\forall X_l^o \in \tilde{\mathfrak{R}}^{|M|}) [X_k^o \neq X_l^o \rightarrow (d_k^* < d(x_k \oplus x_l)) \& \\ & \& (d_l^* < d(x_k \oplus x_l))], \end{aligned} \quad (2.1.1)$$

де d_k^* – оптимальний радіус контейнера K_k^o ;

d_l^* – оптимальний радіус контейнера K_l^o .

Основна ідея методів кластер-аналізу за ІЕІ-технологією полягає в паралельному за часом функціонуванні інтелектуальних агентів СППР, що

реалізують задачі навчання та екзамену та мають між собою динамічний зв'язок, спрямований на статистичну корекцію детермінованих вирішальних правил з метою побудови безпомилкового за навчальною вибіркою класифікатора. При цьому на етапі навчання за поточними вхідними даними і результатами екзамену, здійснюється цілеспрямовано трансформація поточного нечіткого розбиття простору ознак розпізнавання у чітке розбиття еквівалентності, а на етапі екзамену за статистичним критерієм стійкості та однорідності послідовне збільшення потужності відповідного алфавіту класів розпізнавання $\{X_m^0\}^\wedge$ та формування повної навчальної матриці, яка поступає на вхід агента „Навчання” для донавчання системи.

Алгоритм навчання за ІЕІ-технологією полягає в реалізації багато циклическої ітераційної процедури оптимізації структурованих просторово-часових параметрів функціонування СПР шляхом пошуку глобального максимуму усередненого за алфавітом $\{X_m^o\}$ значення КФЕ навчання в робочій (допустимій) області визначення його функції.

Основними концептуальними положеннями методів кластер-аналізу у рамках ІЕІ-технології є:

- методи кластер-аналізу ґрунтуються на прямій оцінці інформаційної спроможності АСК, що навчається;
- прийняття рішень здійснюється в рамках детерміновано-статистичного підходу шляхом побудови відносно простого детермінованого класифікатора, статистична корекція якого здійснюється в процесі навчання з метою підвищення достовірності рішень, що приймаються на екзамені;
- методи ґрунтуються на застосуванні гіпотез як чіткої, так і нечіткої компактності реалізацій образу, тобто є працездатним за умов нечіткого апріорного розбиття простору ознак на класи розпізнавання, що має місце в практичних задачах контролю та керування;
- методи є об'єктно-структурованими, що дозволяє їх розвивати для вирішення проблеми інформаційного синтезу широкого класу АСК з

успадкуванням властивостей структурованих об'єктів вищого ієрархічного рівня і до визначення їх методів;

- методи базуються на вибірковому підході математичної статистики і орієнтовані на застосування прийнятних з практичних міркувань мінімальних обсягів репрезентативних навчальних вибірок;

- процес навчання та екзамену при функціонуванні АСК в режимі кластер-аналізу здійснюється паралельно;

- структура АСК містить як обов'язковий функціональний елемент (інформаційний агент) блок розвідувального аналізу, призначенням якого є оцінка статистичної сталості та статистичної однорідності навчальних вибірок при формуванні алфавіту $\{X_m^0\}^\Lambda$;

- методи повинні бути універсальними для проектування здатної навчатися АСК будь-якої природи і дозволяти вирішувати як загальну, так і часткові задачі її інформаційного синтезу.

До теперішнього часу серед фахівців в галузі автоматичної класифікації не існує єдиної думки про завдання та зміст кластер-аналізу. Можна запропонувати три задачі кластер-аналізу, які становлять практичний інтерес в задачах керування слабо формалізованими процесами:

- кластер-аналіз першого роду, або факторний кластер-аналіз (ФКА);
- кластер-аналіз другого роду, або традиційний кластер-аналіз;
- кластер-аналіз третього роду (самонавчання).

Визначення 2.1.1. Кластер-аналіз першого роду – класифікація реалізацій образу при частково невизначеному алфавіті класів розпізнавання, яка допускає апріорне завдання, по крайній мірі, базового класу X_1^0 .

Ця задача є типовою для СППР, основною функцією яких є оцінка поточного стану технологічного процесу з метою вироблення корегуючих команд. При цьому за базовий клас X_1^0 природньо приймається функціональний стан технологічного процесу, який характеризується найбільшою функціональною ефективністю АСКТП. Для забезпечення умов

ефективного функціонування СППР в режимі ФКА необхідно забезпечення статистичної стійкості й однорідності вибірових значень ознак розпізнавання при зміні функціональних станів технологічного процесу.

Визначення 2.1.2. Кластер-аналіз другого роду – класифікація реалізацій образу при повністю невизначеному алфавіті класів розпізнавання і незмінній структурі та потужності словника ознак розпізнавання. Саме цей тип задачі в літературі найчастіше подається як традиційна задача кластер-аналізу.

При автоматизації керованих технологічних процесів цей тип задач має місце, наприклад при формуванні апріорних нечітких класів розпізнавання, що дозволяє СППР перебрати на себе функцію розробника вхідного математичного опису.

Визначення 2.1.3. Кластер-аналіз третього роду, або самонавчання – класифікація при повністю невизначеному алфавіті класів розпізнавання з оптимізацією параметрів і потужності словника ознак розпізнавання.

Таким чином, специфіка самонавчання СППР у рамках ІЕІТ пов'язана як з розв'язанням задачі традиційного кластер-аналізу, так і з оцінкою інформативності ознак розпізнавання та перевіркою статистичних гіпотез про параметри та закони розподілу ймовірностей випадкових величин за методами багатовимірного статистичного аналізу.

2.2 Постановка задачі та математична модель інформаційно-екстремального факторного кластер-аналізу

Нехай дано $\{X_m^o \mid m = \overline{1, M}\}^\Lambda$ – відкритий алфавіт класів розпізнавання, який в процесі функціонування СППР, що навчається, змінює свою потужність, і відкриту навчальну матрицю, де N, n – кількість ознак розпізнавання та випробувань відповідно, Λ – символ відкритості множини. На етапі навчання СППР за апріорно класифікованими реалізаціями образів у

рамках ІЕІТ побудовано оптимальне в інформаційному розумінні чітке розбиття $\mathfrak{R}^{|M|}$ дискретного простору ознак Ω_B на M класів розпізнавання.

Треба на етапі екзамену за алгоритмом ФКА для нового класу X_{M+1}^0 сформуувати статистично стійку та однорідну навчальну матрицю $\{x_{M+1}^{(j)} \mid j = \overline{1, n}\} \in \{x_m^{(j)} \mid m = \overline{1, m+1}\}^\Lambda$ за таким предикатним виразом:

$$(\forall x^{(j)} \in \tilde{\mathfrak{R}}^\Lambda) [if\ x^{(j)} \notin \{X_m^o\}^\Lambda\ then\ x^{(j)} \in X_{M+1}^o],$$

де $x^{(j)}$ – двійковий вектор-реалізація образу, що розпізнається;

$\tilde{\mathfrak{R}}^\Lambda$ – відкрите нечітке розбиття простору Ω_B ,

та здійснити донавчання системи за ІЕІ-технологією так, щоб максимізувати усереднене значення інформаційного КФЕ навчання АСК:

$$\bar{E}^* = \frac{1}{M+1} \sum_{m=1}^{M+1} \max_{\{d\}} E_m, \quad (2.2.1)$$

де E_m – інформаційний КФЕ навчання АСК розпізнавати реалізації класу X_m^0 ;

$\{d\}$ – множина кроків навчання.

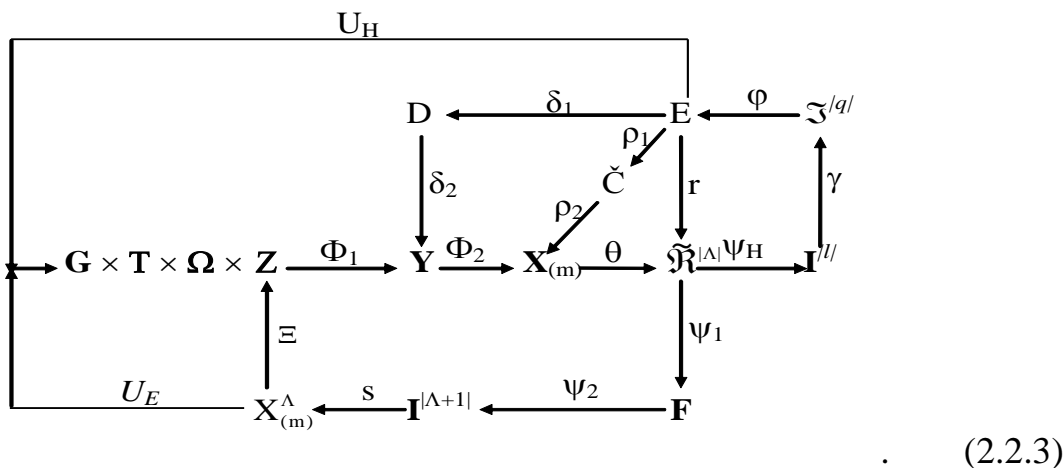
Оскільки достовірність керуючих рішень залежить від забезпечення статистичних стійкості та однорідності багатовимірної навчальної матриці, то важливим завданням дослідження є розробка ієрархічного алгоритму оцінки статистичних властивостей навчальних вибірок за умов їх нормальності у динамічному режимі зміни функціональних станів технологічного процесу. Умова нормальності таких вибірок на практиці забезпечує обґрунтування гіпотези компактності реалізацій образу (чіткої або нечіткої), що має місце в практичних задачах контролю та керування.

На рівні системного аналізу узагальнена математична модель ФКА повинна включати як обов'язкову складову частину вхідний математичний опис, який подамо у вигляді такої теоретико-множинної структури:

$$\Delta_B = \langle G, T, \Omega, Z, Y, \Xi, \Phi \rangle, \tag{2.2.2}$$

де G – простір вхідних сигналів (факторів), які діють на АСК; T – множина моментів часу зняття інформації; Ω – простір ознак розпізнавання; Z – простір можливих станів АСК; Y – навчальна матриця $\|y_{m,i}^{(j)}\|$; $\Xi: G \times T \times \Omega \rightarrow Z$ – оператор формування нових функціональних станів; $\Phi: G \times T \times \Omega \times Z \rightarrow Y$ – оператор формування вибіркової множини Y на вході СППР.

Математична модель ФКА у вигляді діаграми відображень множин, що використовуються в процесі прийняття рішень за ІЕІ-технологією має вигляд



У діаграмі (2.2.3) оператор формування вхідного математичного опису Φ подано як композицію двох операторів: $\Phi = \Phi_1 \circ \Phi_2$, де Φ_1 – оператор формування цілої (або дійсної) навчальної матриці Y , а ψ_2 – оператор формування бінарної навчальної матриці X шляхом допустимих перетворень матриці Y . Оператор $\theta: Y \rightarrow \tilde{\mathfrak{R}}^{|M|}$ будує у загальному випадку нечітке розбиття $\tilde{\mathfrak{R}}^{|M|}$, яке допускає перетин класів розпізнавання, що є

характерним для практичних задач контролю та керування. Оператор $\Psi: \tilde{\mathfrak{R}}^{|M|} \rightarrow I^{|l|}$ перевіряє основну статистичну гіпотезу $\gamma_1: y_{m,i}^{(j)} \in X_m^o$, де $I^{|l|}$ – множина гіпотез, яка для двохальтернативних рішень крім основної містить альтернативну гіпотезу $\gamma_2: y_{m,i}^{(j)} \notin X_m^o$, l – кількість статистичних гіпотез. Оператор γ визначає множину точнісних характеристик $\mathfrak{S}^{|q|}$. Тут $q=l^2$. Оператор ϕ обчислює терм-множину E значень інформаційного критерію оптимізації, який є функціоналом точнісних характеристик. Оператор r корегує розбиття $\tilde{\mathfrak{R}}^{|M|}$ залежно від значень критерію E .

У порівнянні з базовим алгоритмом навчання (1.3.2) діаграма (2.2.3) містить контур

$$\boxed{\rightarrow \Phi_2 \rightarrow \theta \rightarrow \Psi_H \rightarrow \gamma \rightarrow \phi \rightarrow \delta_1 \rightarrow \delta_2}, \quad (2.2.4)$$

який оптимізує за інформаційним КФЕ навчання систему контрольних допусків D шляхом цілеспрямованої зміни параметра поля допусків $\delta \in [1; \delta_n/2]$, де δ_n – нормоване (експлуатаційне) поле допусків на ознаки розпізнавання. Контур

$$\boxed{\rightarrow \theta \rightarrow \Psi_H \rightarrow \gamma \rightarrow \phi \rightarrow \rho_1 \rightarrow \rho_2}, \quad (2.2.5)$$

оптимізує рівні селекції координат двійкових еталонних векторів-реалізацій шляхом цілеспрямованої зміни рівня селекції $\rho_m \in [0;1]$.

Оператори U_H і U_E регламентують процес навчання й екзамену відповідно, а оператор класифікації реалізацій образу в режимі екзамену утворює композицію $\psi = \psi_1 \cdot \psi_2$, де ψ_1 – оператор обчислення функції належності реалізації образу відповідному контейнеру; ψ_2 – оператор реалізації вирішальних правил.

Необхідною та достатньою умовою реалізації ФКА у рамках ІЕІ-технології є виконання нерівності

$$\bar{\mu}_m = \frac{1}{n} \sum_{j=1}^n \mu_{m,j} \leq c, \quad (2.2.6)$$

де $\bar{\mu}_m$ – усереднена функція належності контейнеру $K_m^0 \subset X_m^0$

реалізації образу, що розпізнається;

c – порогове значення, що обумовлює прийняття гіпотези відмови

від класифікації $\gamma_{\Lambda+1} \in I^{|\Lambda+2|}$.

Тут $I^{|\Lambda+2|}$ – множина гіпотез для відкритої абетки, де $\gamma_{\Lambda+2}$ – гіпотеза, що дозволяє формування навчальної матриці нового класу X_{Λ}^0 і відповідно донавчання системи.

У рамках ФКА за ІЕІ-технологією ідея агрегування нового класу при незмінному словнику ознак розпізнавання полягає у формуванні оператором Ξ додаткової навчальної матриці $\|x_{\Lambda}^{(j)}\|$, де $\Lambda = M + 1$, яка складається з реалізацій екзаменаційної матриці, що дали негативні для всіх класів значення функції належності. При досягненні необхідної репрезентативності матриці $\|x_{\Lambda}^{(j)}\|$ оператор ζ запускає процес донавчання АСК з метою побудови нового розбиття простору ознак.

Таким чином, необхідність реалізації алгоритмів ФКА в сучасних АСК обумовлено низькою достовірністю оцінки функціональних станів керованих слабо формалізованих технологічних процесів, що відбуваються за умов апріорної невизначеності. Крім того, на достовірність керуючих рішень, що виробляє здатна навчатися АСК, суттєво впливає крім відносно великої дисперсії значень ознак розпізнавання в екзаменаційній матриці і дрейфу у часі вершин векторів-реалізацій образів забезпечення статистичних стійкості

й однорідності навчальної вибірки нового класу розпізнавання, що формується безпосередньо в процесі функціонування СППР в режимі ФКА.

2.3 Забезпечення статистичних стійкості й однорідності навчальної вибірки у процесі функціонування СППР в режимі ФКА

Необхідною умовою застосування на виробництві розроблених за ІЕІ-технологією інтелектуальних СППР, що реалізують алгоритми кластер-аналізу, є забезпечення статистичної стійкості та однорідності навчальної вибірки. Тому одним із завдань дисертаційної роботи є розробка методу оцінки статистично стійких та однорідних навчальних вибірок нових класів, які є множиною випадкових значень ознак розпізнавання функціональних станів керованого технологічного процесу, що можуть поступово змінюватися під впливом некерованих випадкових факторів.

Нехай дано $\{X_m^o \mid m = \overline{1, M}\}^\Lambda$ – відкритий алфавіт класів розпізнавання, який в процесі функціонування АСК, що навчається, змінює свою потужність, і відкриту навчальну матрицю $\|y_{m,i}^{(j)}\|$, де N, n – кількість ознак розпізнавання та випробувань відповідно, Λ – символ відкритості множини. На етапі навчання АСК за апріорно класифікованими реалізаціями образів у рамках ІЕІТ побудовано оптимальне в інформаційному розумінні чітке розбиття $\mathfrak{R}^{|M|}$ дискретного простору ознак Ω_B на M класів розпізнавання. Треба на етапі екзамену за алгоритмом ФКА для нового класу X_{M+1}^o сформувати статистично стійку та однорідну навчальну матрицю $\{x_{M+1}^{(j)} \mid j = \overline{1, n}\} \in \|x_m^{(j)} \mid m = \overline{1, M+1}\|^\Lambda$ за таким предикатним виразом:

$$(\forall x^{(j)} \in \tilde{\mathfrak{R}}^\Lambda) [if\ x^{(j)} \notin \{X_m^o\}\ then\ x^{(j)} \in X_{M+1}^o], \quad (2.3.1)$$

де $x^{(j)}$ – двійкова вектор-реалізація образу, що розпізнається,

та здійснити донавчання системи за ІЕІ-технологією так, щоб максимізувати усереднене значення інформаційного КФЕ навчання АСК (2.2.1):

Оскільки достовірність керуючих рішень залежить від забезпечення статистичних стійкості та однорідності багатовимірної навчальної матриці, то важливим завданням дослідження є розробка ієрархічного алгоритму оцінки статистичних властивостей навчальних вибірок за умов їх нормальності у динамічному режимі зміни функціональних станів технологічного процесу. Умова нормальності таких вибірок на практиці забезпечує обґрунтування гіпотези компактності реалізацій образу (чіткої або нечіткої), що має місце в практичних задачах контролю та керування.

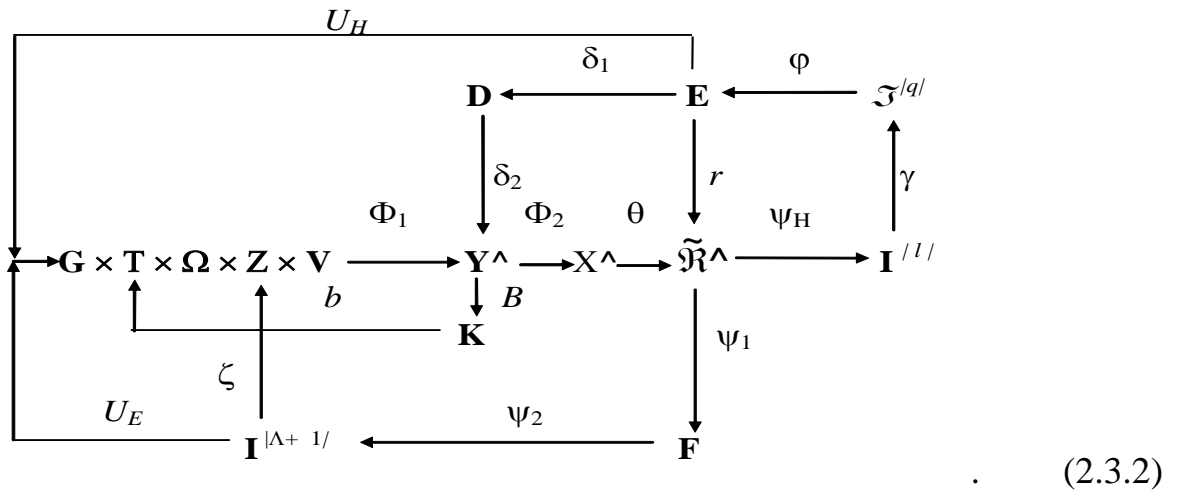
Розглянемо в рамках ІЕІ-технології розв'язання задачі ФКА з метою побудови в просторі ознак розпізнавання нових представництв, що є афінними різноманіттями, навколо яких агрегуються нові класи. Необхідною та достатньою умовою реалізації ФКА за ІЕІ-технологією є виконання нерівності (2.2.6), що обумовлює прийняття гіпотези $\gamma_{\Lambda+1} \in I^{|\Lambda+2|}$ – відмова від класифікації j -ї реалізації образу. Тут $I^{|\Lambda+2|}$ – множина гіпотез для відкритої абетки, де $\gamma_{\Lambda+2}$ – гіпотеза, що дозволяє формування навчальної матриці нового класу X_{Λ}^o і відповідно донавчання системи. Нехай вхідний математичний опис СППР в режимі ФКА має структуру:

$$\Delta_B = \langle G, T, \Omega, Z, V, Y^{\Lambda}, X^{\Lambda}, K; \Phi_1, \Phi_2, B \rangle,$$

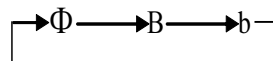
де V – множина типів вирішальних правил; Y^{Λ} – вхідна навчальна матриця $\| y_m^{(j)} \|^{\Lambda}$; X^{Λ} – відкрита дискретна вибіркова множина, що утворює бінарну навчальну матрицю $\| x_m^{(j)} \|^{\Lambda}$; K – терм-множина статистичних оцінок навчальної матриці; $\Phi_1: G \times T \times \Omega \times Z \times V \rightarrow Y$ – оператор формування

вибіркової множини Y на вході СППР; $\Phi_2: Y \rightarrow X^\wedge$ оператор формування дискретної множини X^\wedge ; B – оператор формування множини K .

Узагальнену математичну модель ФКА за ІЕІ-технологією подамо у вигляді категорійної моделі – діаграми відображень множин, що застосовуються на практиці при функціонування СППР для керування складними технологічними процесами:



У діаграмі (2.3.2) контур



за результатами розвідувального аналізу забезпечує блокування алгоритму навчання СППР у випадку, якщо не забезпечуються статистичні стійкість та однорідність навчальних вибірок для всіх ознак розпізнавання. Оператор θ формує апіорне розбиття $\tilde{\mathfrak{R}}^{|\Lambda|}$, яке у загальному випадку може бути нечітким, а оператор $\Psi_H: \tilde{\mathfrak{R}}^{|\Lambda|} \rightarrow I^{|l|}$, де $I^{|l|} = \{\gamma_1, \dots, \gamma_l\}$ – множина статистичних гіпотез, перевіряє основну статистичну гіпотезу: $\gamma_1: x_m^{(j)} \in X_m^o$.

Після формування терм-множини точністних характеристик навчання $\mathfrak{Z}^{|q|}$, де $q = l^2$, оператор ϕ обчислює значення інформаційного КФЕ (терм-множина E). Контур операторів

$$\boxed{\Psi_H \rightarrow \gamma \rightarrow \varphi \rightarrow \Gamma}, \quad (2.3.3)$$

який оптимізує геометричні параметри розбиття $\tilde{\mathfrak{R}}^{|\Lambda|}$, і контур

$$\boxed{\theta \rightarrow \Psi_H \rightarrow \gamma \rightarrow \varphi \rightarrow \delta_1 \rightarrow \delta_2}, \quad (2.3.4)$$

який оптимізує систему контрольних допусків D , реалізують базовий алгоритм навчання у рамках ІЕІ-технології [31]. Оператор U_H регламентує навчання і дозволяє оптимізувати параметри плану навчання. Оператори Ψ_1, Ψ_2 і U_E реалізують алгоритм екзамену, який при кластер-аналізі функціонує паралельно з алгоритмом навчання. Тут оператор Ψ_1 обчислює терм-множину функцій належності для заданого типу класифікатора, оператор Ψ_2 здійснює дефазифікацію, а оператор U_E регламентує процес екзамену. Оператор ζ формує додаткову навчальну матрицю і дає дозвіл на донавчання системи.

Ієрархічна структура критеріїв оцінки статистичних стійкості та однорідності, наведених, наприклад, в працях [98-100] навчальної вибірки наведено на рис. 2.1. Аналіз рис.2.1 показує, що невиконання основної гіпотези навіть за одним будь-яким критерієм призводить до заблокування процесу формування додаткової навчальної матриці $\|x_{\Lambda}^{(j)}\|$, а тим самим і донавчання СППР.

У загальному випадку блоки навчання та екзамену припиняють функціонування СППР за таких причин:

- перехідний процес при зміні функціональних станів технологічного процесу під впливом як керованих, так і некерованих факторів;
- відмова технологічного обладнання;
- неправильне функціонування датчиків інформації.



Рисунок 2.1 – Ієрархічна структура критеріїв оцінки статистичних стійкості та однорідності навчальної вибірки



Після того як блок розвідувального аналізу, що перевіряє основні статистичні гіпотези за відповідними критеріями, зафіксує нормальність розподілу та статистичну стійкість вхідних даних, СППР переводиться в режим штатного функціонування.

2.4 Критерії оптимізації процесу навчання СППР

Центральним питанням інформаційного синтезу СППР є оцінка функціональної ефективності процесу навчання, яка визначає максимальну асимптотичну достовірність рішень, що приймаються на екзамені. Як КФЕ в

методах ІЕІ-технології можуть використовуватися різні критерії, які задовольняють таким властивостям інформаційних мір:

- інформаційна міра є величина дійсна і знако-додатна як функція від імовірності;

- кількість інформації для детермінованих змінних ($p_i = 1$ або $p_i = 0$) дорівнює нулю;

- інформаційна міра має екстремум при значенні ймовірності $p_i = \frac{1}{m}$,

де m – кількість якісних ознак розпізнавання.

Серед інформаційних мір для оцінки функціональної ефективності СППР, що навчається, перевагу слід віддавати статистичним логарифмічним критеріям, які дозволяють працювати з навчальними вибірками відносно малих обсягів [61]. Серед таких критеріїв найбільшого використання знайшли ентропійні міри [62,63] та інформаційна міра Кульбака [64].

Подамо нормований ентропійний КФЕ [24] навчання СППР розпізнавати реалізації класу X_m^o у вигляді:

$$E_m^{(k)} = \frac{I_m^{(k)}}{I_{\max}^{(k)}} = \frac{H_m^{(k)} - H_m^{(k)}(\gamma)}{H_m^{(k)}}, \quad (2.4.1)$$

де $I_m^{(k)}$ – кількість умовної інформації, що обробляється на k -му кроці навчання СППР розпізнавати реалізації класу X_m^o ; $I_{\max}^{(k)}$ – максимально можлива кількість умовної інформації, що одержана на k -му кроці навчання СППР розпізнавати реалізації одного із класів із заданого алфавіту $\{X_m^o\}$, $m = \overline{1, M}$;

$$H_m^{(k)} = - \sum_{l=1}^M p(\gamma_{l,k}) \log_2 p(\gamma_{l,k}) - \quad (2.4.2)$$

априорна (безумовна) ентропія, що існує на k -му кроці навчання системи розпізнавати реалізації класу X_m^o ;

$$H_m^{(k)}(\gamma) = -\sum_{l=1}^M \sum_{m=1}^M p(\gamma_{l,k}) p(\mu_{m,k} / \gamma_{l,k}) \log_2 p(\mu_{m,k} / \gamma_{l,k}) - \quad (2.4.3)$$

апостеріорна (умовна) ентропія, що характеризує залишкову невизначеність після k -го кроку навчання системи розпізнавати реалізації класу X_m^o ; $p(\gamma_{l,k})$ – безумовна ймовірність прийняття на k -му кроці навчання гіпотези $\gamma_{l,k}$; $p(\mu_{m,k} / \gamma_{l,k})$ – апостеріорна ймовірність прийняття на k -му кроці навчання рішення $\mu_{m,k}$ за умови, що прийнята гіпотеза $\gamma_{l,k}$.

Для двохальтернативної системи оцінок ($M = 2$) і рівноймовірних гіпотез, що характеризує найбільш важкий у статистичному сенсі випадок прийняття рішень, після відповідної підстановки ентропій (2.4.2) і (2.4.3) у вираз (2.4.1) та заміни відповідних апостеріорних ймовірностей на априорні за формулою Байєса ентропійний критерій набуває вигляду:

$$E_m^{(k)} = 1 + \frac{1}{2} \left(\frac{\alpha_m^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} \log_2 \frac{\alpha_m^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} + \frac{\beta_m^{(k)}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} \log_2 \frac{\beta_m^{(k)}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} + \frac{D_{1,m}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} \log_2 \frac{D_{1,m}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} + \frac{D_{2,m}^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} \log_2 \frac{D_{2,m}^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} \right), \quad (2.4.4)$$

де $\alpha_m^{(k)}(d)$ – помилка першого роду – точнісна характеристика рішення на k -му кроці навчання; $\beta_m^{(k)}(d)$ – помилка другого роду; $D_{1,m}^{(k)}(d)$ – перша достовірність; $D_{2,m}^{(k)}(d)$ – друга достовірність; d – дистанційна міра, яка

визначає радіуси гіперсферичних контейнерів, побудованих в радіальному базисі простору Хеммінга.

Оскільки точнісні характеристики є функціями відстані вершин еталонних векторів-реалізацій образу від геометричних центрів контейнерів відповідних класів розпізнавання, то критерій (2.4.4) в ІЕІ-технології слід розглядати як нелінійний і взаємно-неоднозначний функціонал від точнісних характеристик, що потребує знаходження в процесі навчання робочої (допустимої) області для його визначення.

Розглянемо модифікацію диференційної інформаційної міри Кульбака, яка подається як добуток відношення правдоподібності Λ на міру відхилень відповідних розподілів імовірностей [65].

У праці [31] розглядається логарифмічне відношення повної ймовірності $P_{t,m}^{(k)}$ правильного прийняття рішень про належність реалізацій класів X_m^o і X_c^o контейнеру $K_{m,k}^o \in X_m^o$ до повної ймовірності помилкового прийняття рішень $P_{f,m}^{(k)}$, яке для двохальтернативної системи оцінок рішень має такий вигляд:

$$\Lambda = \log_2 \frac{P_{t,m}^{(k)}}{P_{f,m}^{(k)}} = \log_2 \frac{p(\mu_m)p(\gamma_{1,k}/\mu_m) + p(\mu_c)p(\gamma_{2,k}/\mu_c)}{p(\mu_m)p(\gamma_{2,k}/\mu_m) + p(\mu_c)p(\gamma_{1,k}/\mu_c)}, \quad (2.4.5)$$

де $p(\mu_m)$ –безумовна ймовірність появи реалізації класу X_m^o ; $p(\mu_c)$ –безумовна ймовірність появи реалізації класу X_c^o ; $\gamma_{1,k}$ – гіпотеза про належність контейнеру $K_{m,k}^o \in X_m^o$ реалізації класу X_m^o ; $\gamma_{2,k}$ – альтернативна гіпотеза.

Із врахуванням (2.4.5) при допущенні згідно з принципом Лапласа-Бернуллі, що $p(\mu_m) = p(\mu_c) = 0,5$, і після переозначення апріорних умовних

імовірностей відповідними точністними характеристиками [99], загальна міра Кульбака остаточно набуває вигляду:

$$\begin{aligned}
 J_m^{(k)} &= \log_2 \frac{P_{t,m}^{(k)}}{P_{f,m}^{(k)}} * [P_{t,m}^{(k)} - P_{f,m}^{(k)}] = \\
 &= 0,5 \log_2 \left(\frac{D_{1,m}^{(k)}(d) + D_{2,m}^{(k)}(d)}{\alpha_m^{(k)}(d) + \beta_m^{(k)}(d)} \right) \left[(D_{1,m}^{(k)}(d) + D_{2,m}^{(k)}(d)) - (\alpha_m^{(k)}(d) + \beta_m^{(k)}(d)) \right] = \\
 &= \log_2 \left(\frac{2 - (\alpha_m^{(k)}(d) + \beta_m^{(k)}(d))}{\alpha_m^{(k)}(d) + \beta_m^{(k)}(d)} \right) \left[1 - (\alpha_m^{(k)}(d) + \beta_m^{(k)}(d)) \right]. \quad (2.4.6)
 \end{aligned}$$

Нормовану модифікацію критерію (2.4.6) можна подати у вигляді:

$$E_{K,m}^{(k)} = \frac{J_m^{(k)}}{J_{\max}^{(k)}},$$

де $J_{\max}^{(k)}$ – значення критерію при $D_{1,m}^{(k)}(d) = D_{2,m}^{(k)}(d) = 1$ і відповідно $\alpha_m^{(k)}(d) = \beta_m^{(k)}(d) = 0$ для формули (2.4.6).

У задачах оптимізації параметрів функціонування АСК, що навчається в процесі навчання за ІЕІ-технологією нормування критеріїв оптимізації не є обов'язковим, оскільки тут розв'язується задача пошуку екстремальних значень параметрів навчання, які відповідають глобальному максимуму КФЕ у робочій області його визначення. Але нормування критеріїв оптимізації є доцільним при порівняльному аналізі результатів досліджень і при оцінці ступеню близькості реальної системи керування до потенційної.

Розглянемо процедуру обчислення модифікації ентропійного інформаційного КФЕ за Шенноном для двохальтернативного рішення при рівноймовірних гіпотезах згідно з формулою (2.4.4). Оскільки інформаційний критерій є функціоналом від точністних характеристик, то при мінімальному обсязі репрезентативної навчальної вибірки слід користуватися їх оцінками:

$$D_{1,m}^{(k)}(d) = \frac{K_{1,m}^{(k)}}{n_{\min}}; \alpha_m^{(k)}(d) = \frac{K_{2,m}^{(k)}}{n_{\min}}; \beta_m^{(k)}(d) = \frac{K_{3,m}^{(k)}}{n_{\min}}; D_{2,m}^{(k)}(d) = \frac{K_{4,m}^{(k)}}{n_{\min}}, \quad (2.4.7)$$

де $K_{1,m}^{(k)}$ – кількість подій, які означають належність реалізацій образу контейнеру $K_{1,mk}^o$, якщо дійсно $\{x_1^{(j)}\} \in X_1^o$; $K_{2,m}^{(k)}$ – кількість подій, які означають неналежність реалізацій образу контейнеру $K_{1,m}^o$, якщо дійсно $\{x_1^{(j)}\} \in X_1^o$; $K_{3,m}^{(k)}$ – кількість подій, які означають належність реалізацій образу контейнеру $K_{1,m}^o$, якщо вони насправді належать класу X_2^o ; $K_{4,m}^{(k)}$ – кількість подій, які означають неналежність реалізацій образу контейнеру $K_{1,m}^o$, якщо вони насправді належать класу X_2^o ; n_{\min} – мінімальний обсяг репрезентативної навчальної вибірки.

Після підстановки відповідних позначень (2.4.7) в (2.4.4) отримаємо робочу модифіковану формулу для обчислення в рамках ІЕІ-технології ентропійного інформаційного КФЕ навчання АСК розпізнаванню реалізацій класу X_1^o

$$E_m^{(k)} = 1 + \frac{1}{2} \left(\frac{K_{1,m}^{(k)}}{K_{1,m}^{(k)} + K_{3,m}^{(k)}} \log_2 \frac{K_{1,m}^{(k)}}{K_{1,m}^{(k)} + K_{3,m}^{(k)}} + \frac{K_{2,m}^{(k)}}{K_{2,m}^{(k)} + K_{4,m}^{(k)}} \log_2 \frac{K_{2,m}^{(k)}}{K_{2,m}^{(k)} + K_{4,m}^{(k)}} + \frac{K_{3,m}^{(k)}}{K_{1,m}^{(k)} + K_{3,m}^{(k)}} \log_2 \frac{K_{3,m}^{(k)}}{K_{1,m}^{(k)} + K_{3,m}^{(k)}} + \frac{K_{4,m}^{(k)}}{K_{2,m}^{(k)} + K_{4,m}^{(k)}} \log_2 \frac{K_{4,m}^{(k)}}{K_{2,m}^{(k)} + K_{4,m}^{(k)}} \right). \quad (2.4.8)$$

Робоча модифікація критерію Кульбака після відповідної підстановки оцінок (2.4.7) у вираз (2.4.6) набуває вигляду

$$J_m^{(k)} = \frac{1}{n} \log_2 \left\{ \frac{2n + 10^{-r} - [K_{2,m}^{(k)} + K_{3,m}^{(k)}]}{[K_{2,m}^{(k)} + K_{3,m}^{(k)}] + 10^{-r}} \right\} * [n - (K_{2,m}^{(k)} + K_{3,m}^{(k)})], \quad (2.4.9)$$

де $K_{2,m}^{(k)}$ – число реалізацій класу X_m^o , які не знаходяться в побудованому на k -му кроці навчання контейнері цього класу; $K_{3,m}^{(k)}$ – число реалізацій класу X_c^o , які знаходяться в побудованому на k -му кроці навчання контейнері класу X_m^o .

Розглянемо схему обчислення коефіцієнтів $K_2^{(k)}$ і $K_3^{(k)}$ у формулах (2.4.8) і (2.4.9). На рис. 3.3 показано структуру навчальної матриці при побудові оптимального контейнера для класу X_1^o . Навчальна матриця послідовно складається з векторів реалізацій $\{x_1^{(j)}\} \in X_1^o$ і $\{x_2^{(j)}\} \in X_2^o$ відповідно. Структура навчальної матриці для обчислення, наприклад, коефіцієнтів $K_2^{(k)}$ і $K_3^{(k)}$ у формулі (2.4.9) представлена на рис. 2.2.

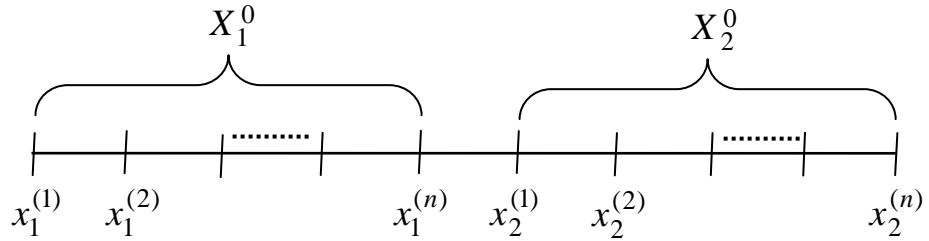


Рисунок 2.2– Структура навчальної матриці

Тестовий алгоритм обчислення коефіцієнтів $K_2^{(k)}$ і $K_3^{(k)}$ має вигляд

$$(\forall X_1^o \in \mathfrak{R}^{|\Lambda|})(\forall X_2^o \in \mathfrak{R}^{|\Lambda|})[\text{if } x_1^{(j)} \in X_1 \text{ then } K_1(j) := K_1(j-1) + 1 \\ \text{else } K_2(j) := K_2(j-1) + 1];$$

$$(\forall X_1^o \in \mathfrak{R}^{|\Lambda|})(\forall X_2^o \in \mathfrak{R}^{|\Lambda|})[\text{if } x_2^{(j)} \in X_1 \text{ then } K_3(j) := K_3(j-1) + 1 \\ \text{else } K_4(j) := K_4(j-1) + 1]. \quad (2.4.10)$$

Таким чином, інформаційні критерії (2.4.4) і (2.4.6) є функціоналами як від точнісних характеристик рішень, що приймаються, так і від дистанційних критеріїв, тобто є узагальненням відомих статистичних і детермінованих критеріїв оптимізації параметрів функціонування системи прийняття рішень.

2.5 Визначення мінімального обсягу репрезентативної навчальної вибірки

Однією з переваг методів оптимізації АСК, що навчаються, у рамках ІЕІ-технології є використання навчальних вибірок супермалого обсягу, що пояснюється згладжуючим ефектом логарифмічних інформаційних КФЕ [61]. Тому важливого значення набуває задача визначення мінімального обсягу репрезентативної навчальної вибірки, яка забезпечує гарантовану статистичну похибку $\varepsilon = |p_i - \frac{n_i}{n}|$, де p_i , $\frac{n_i}{n}$ – імовірність та емпірична частота знаходження i -ї ознаки розпізнавання в своєму полі контрольних допусків. Для двійкової навчальної вибірки верхня оцінка статистичної похибки Sup_{ε} визначається за відомою теоремою Муавра-Лапласа [99,100]:

$$P\left\{\left|\frac{k_i}{n} - p_i\right| \geq \varepsilon\right\} = P\left\{\left|\frac{k_i - np}{\sqrt{np_iq_i}}\right| - \frac{\varepsilon\sqrt{n}}{\sqrt{p_iq_i}} \geq 0\right\} = 2\Phi\left(-\frac{\varepsilon\sqrt{n}}{\sqrt{p_iq_i}}\right) \geq 2\Phi(-2\varepsilon\sqrt{n}), (2.5.1)$$

де k_i – кількість подій, при яких значення i -ї ознаки знаходиться в полі допусків $\delta_{k,i}$; $q_i = 1 - p_i$ – імовірність неналежності значення i -ї ознаки полю контрольних допусків $\delta_{k,i}$; $\Phi(\dots)$ – функція Лапласа.

Для визначення залежності статистичної похибки від обсягу навчальної вибірки $\{x_{m,i}^{(j)} \mid j = \overline{1, n}\}$ прирівняємо верхню оцінку статистичної похибки вибраному рівню значущості Q :

$$2\Phi(-2\varepsilon_Q\sqrt{n})=Q. \quad (2.5.2)$$

Із урахуванням властивості функції Лапласа перетворимо (2.5.2) до вигляду:

$$\Phi(2\varepsilon_Q\sqrt{n})=1-\frac{Q}{2} \quad \Phi(2\varepsilon_Q\sqrt{n})=1-\frac{Q}{2}. \quad (2.5.3)$$

Для визначення мінімального обсягу n_{\min} репрезентативної навчальної вибірки з гарантованою максимальною статистичною похибкою $\varepsilon_Q(n)$ побудуємо динамічний довірчий інтервал, який оцінює імовірність p_i знаходження i -ї ознаки в полі контрольних допусків з імовірністю довіри $1-Q$:

$$P\left\{\frac{k_i}{n}-\varepsilon_Q \leq p_i \leq \frac{k_i}{n}+\varepsilon_Q\right\}=1-Q, \quad (2.5.4)$$

де Q – рівень значущості (будь-яке наближене до нуля позитивне число).

Наприклад, для $Q = 0.05$ за таблицею значень функції Лапласа [100, 107, 108] з урахуванням (2.5.3) для $\Phi(x) = 1 - Q/2 = 0.975$, значення аргументу функції Лапласа дорівнює $\arg \Phi(1 - \frac{Q}{2}) = 2\varepsilon_Q\sqrt{n} = 1,98$, тобто похибка ε_Q змінюється залежно від обсягу навчальної вибірки n за гіперболічним законом:

$$\mathop{\text{Sup}}_n \varepsilon_Q = \frac{\arg \Phi(1 - \frac{Q}{2})}{2\sqrt{n}} = \frac{0,98}{\sqrt{n}}; \forall n > 1. \quad (2.5.5)$$

На рис. 2.3 наведено графік функції $\varepsilon_Q = f(n)$ (крива 1) і умовно виділено три області значень аргументу, які відрізняються крутизною функції. Тут область I є забороненою областю, оскільки похибка перебільшує допустиму. Область III характеризується значними економічними втратами при малій швидкості зменшення похибки ε_Q . Область II є компромісною і охоплює інтервал приблизно від 40 до 100 випробувань. Легко довести, що при різних значеннях рівня звчущості Q графік функції $\varepsilon_Q = f(n)$ буде переміщуватися паралельно по вертикалі, не змінюючи свого вигляду.

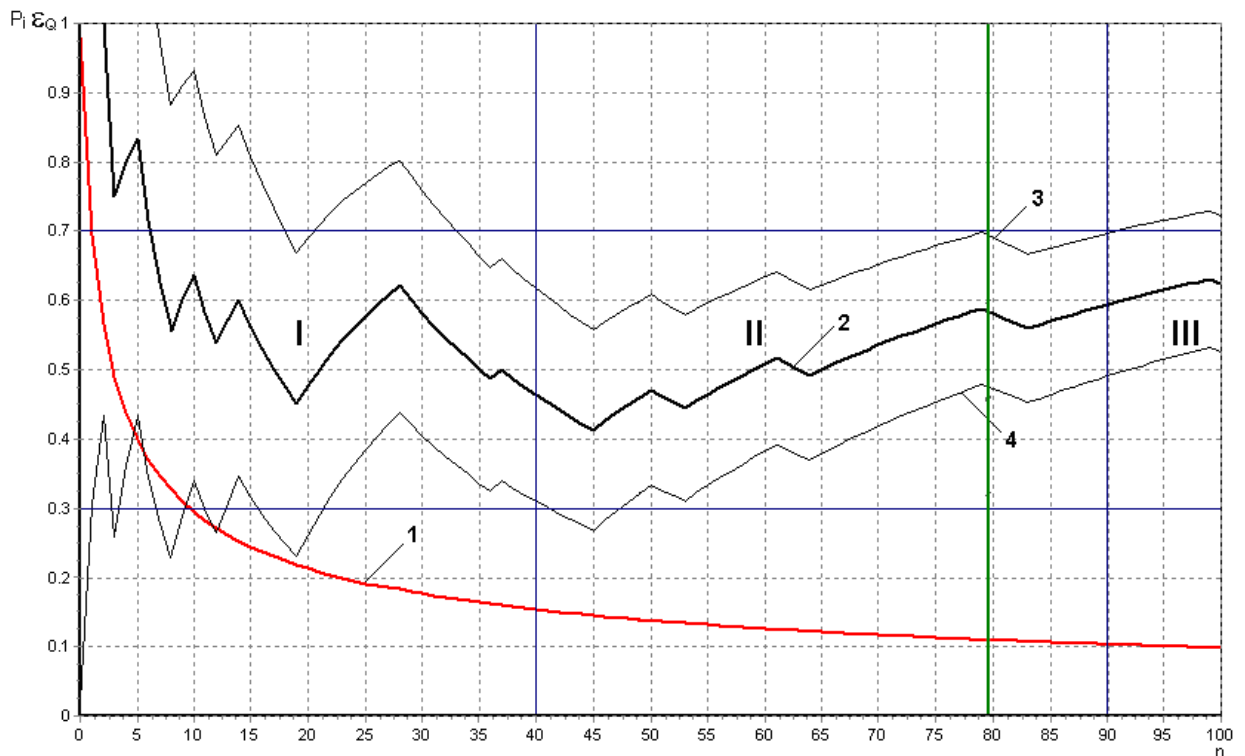


Рисунок 2.3– Визначення мінімального обсягу навчальної вибірки:

1– графік функції $\varepsilon_Q = f(n)$; 2 – графік емпіричної частоти $\tilde{p} = k_i/n$; 3 – верхня межа довірчого інтервалу; 4 – нижня межа довірчого інтервалу

Графічно довірчий інтервал можна побудувати за формулою (2.5.4), обчислюючи для кожного випробування n за виразом (2.5.5) верхню межу похибки ε_Q і відкладаючи її зверху та знизу від графіка частоти $p_i = k_i / n$ (крива 2). При цьому за законом великих чисел верхня $\text{Sup}_n \tilde{p}_i$ (крива 3) та нижня $\text{Inf}_n \tilde{p}_i$ (крива 4) межі довірчого інтервалу при збільшенні числа випробувань мають тенденцію до зближення з емпіричною частотою.

Для знаходження мінімального числа випробувань n_{\min} , яке гарантує прийнятні з практичних міркувань величину похибки і оперативність реалізації алгоритму обчислювання, необхідно задати критерій зупини випробувань. Таким моментом можна вважати випробування, при якому поточний довірчий інтервал накривається заданим інтервалом $[0,5 \pm \Delta]$, де $|\Delta| < 0,5$. Останній (правий) перетин заданого інтервалу з однією з меж довірчого інтервалу визначає випробування n_{\min} , яке гарантує з імовірністю $1 - Q$, що максимальна похибка ε_Q не перебільшує значення функції $\varepsilon_Q = f(n)$ при $n = n_{\min}$. Отже, вибір n_{\min} доцільно здійснювати в компромісній області Π (на рис. 2.3 $n_{\min} = 79$ при $\Delta = \pm 0,2$) за умови відсутності викидів значень емпіричної частоти до значень, близьких до нуля або одиниці. Для багатьох практичних задач значення Δ визначається з інтервалу $[0,2; 0,4]$ за алгоритмом, наведеним, у додатку А.

У загальному випадку, треба будувати довірчі інтервали для всіх N ознак і вибирати n_{\min} за умови

$$n_{\min} = \max_i (n_{\min 1}, \dots, n_{\min i}, \dots, n_{\min N}).$$

При відповідному виборі системи контрольних допусків на незалежні ознаки розпізнавання та забезпеченні умов статистичної сталості проведення випробувань можна вибирати n_{\min} за довірчим інтервалом, побудованим для

будь-якої однієї ознаки, що значно знижує обчислювальну трудомісткість алгоритму.

Таким чином у другому розділі в рамках ІЕІ-технології на рівні системного аналізу розроблено математичні моделі задач факторного класифікаційного аналізу, набули подальшого удосконалення задачі оцінки функціональної ефективності АСК, що навчаються, статистичної стійкості, статистичної однорідності та репрезентативності навчальної вибірки. Запропонований підхід до вирішення проблеми кластер-аналізу дозволяє з єдиних методологічних і теоретичних позицій здійснювати інформаційний синтез АСК, здатних приймати рішення як за апріорно класифікованими, так і за некласифікованими навчальними матрицями.

РОЗДІЛ 3

ОПТИМІЗАЦІЯ ПАРАМЕТРІВ ФУНКЦІОНУВАННЯ АДАПТИВНОЇ АСКТП, ЩО НАВЧАЄТЬСЯ В РЕЖИМІ ФКА

3.1 Оптимізація фенотипних параметрів функціонування СППР, що навчається

Основним завданням процесу навчання СППР є побудова для кожного класу розпізнавання із заданого алфавіту класів $\{X_m^o\}$, $m=\overline{1, M}$, оптимальної в інформаційному розумінні розподільної гіперповерхні, яка забезпечує на етапі екзамену прийняття рішень з достовірністю, близькою до максимальної асимптотичної достовірності, що визначається ефективністю навчання системи. Тому за ІЕІ-технологією процес навчання СППР розглядається як ітераційна процедура наближення глобального максимуму інформаційного КФЕ в робочій (допустимій) області визначення його функції до теоретично граничного значення. У рамках цієї процедури послідовно здійснюється оптимізація параметрів навчання, які безпосередньо впливають на достовірність рішень, що приймаються в режимі екзамену. Як фенотипні параметри навчання, які впливають на форму контейнерів класів розпізнавання, що відновлюються в радіальному базисі простору ознак розпізнавання, розглядаються геометричні параметри – еталонні вектори $\{x_m\}$, $m=\overline{1, M}$, вершини яких є геометричними центрами класів розпізнавання $\{X_m^o\}$, і радіуси роздільних гіперсфер $\{d_m\}$.

Визначення у рамках ІЕІ-технології оптимальної в інформаційному розумінні кодової відстані d_m^* – радіуса розподільної гіперповерхні, здійснюється за послідовним базовим алгоритмом навчання:

$$d_m^*(k) = \arg \max_{\{k\}} E[d(k)], \quad (3.1.2)$$

де k – число збільшень радіуса; $E[d(k)]$ – інформаційний КФЕ навчання СППР, значення якого обчислюються в робочій області визначення його функції.

Обчислення КФЕ закінчується при знаходженні глобального максимуму функції $E_m^* = \max_{G_d} E[d(k)]$, де $G_d \in [0; d(x_n \oplus x_c) - 1]$ – область допустимих значень параметра $d_m(k)$. Тут $d(x_m \oplus x_{m+1})$ – міжцентрова кодова відстань для класу X_m^o і його найближчого сусіднього класу X_c^o .

Розглянемо приклад побудови на етапі навчання оптимального в інформаційному розумінні контейнера класу X_m^o за базовим алгоритмом (3.1.2). На підготовчому етапі формуються бінарна навчальна матриця $\{x_m^{(j)}\}$, $j=1, \overline{n_{\min}}$, яка складається із векторів-реалізацій класів розпізнавання $\{X_m^o\}$, бінарний масив еталонних векторів-реалізацій $\{x_m\}$, які є статистичним усередненням реалізацій відповідних класів, і здійснюється розбиття множини $\{x_m\}$ на пари сусідніх еталонних векторів за умови мінімальної кодової відстані між ними. У процесі навчання за алгоритмом (3.1.2) будуються навколо вершини вектора x_m послідовно з кроком збільшення радіуса $h=1$ концентровані роздільні гіперсфери і на кожному k -му кроці навчання обчислюються значення точнісних характеристик.

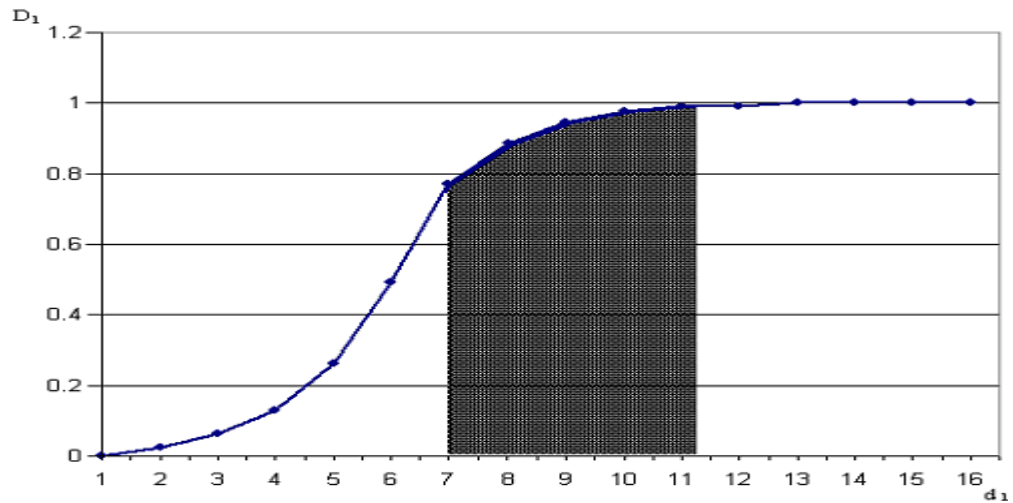
У табл. 3.1 наведено значення точнісних характеристик, обчислених за базовим алгоритмом (3.1.2) за 15 кроків навчання СППР. Тут значення K_2 дорівнює кількості реалізацій класу X_m^o , які знаходяться в побудованому на

k -му кроці навчання контейнері $K_{m,k}^o \in X_m^o$, а значення K_3 відповідно дорівнює кількості реалізацій сусіднього класу, які знаходяться так само в контейнері $K_{m,k}^o$. При цьому коефіцієнти K_1 , K_2 , K_3 і K_4 обчислюються за алгоритмом (2.4.10) за умови, що $K_1 = 1 - K_2$ і $K_4 = 1 - K_3$.

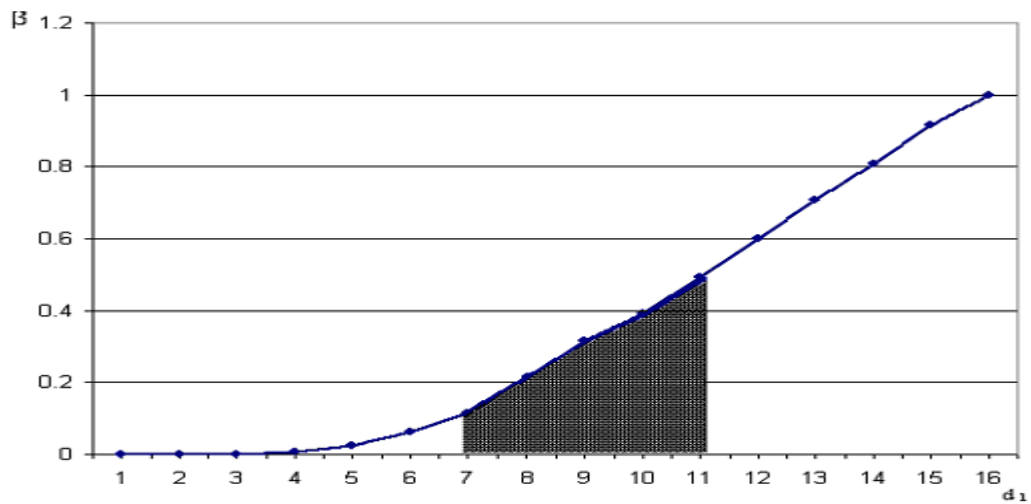
Таблиця 3.1 – Значення точнісних характеристик при оптимізації
контейнера K_m^o

k	K_2	K_3	D_1	α	D_2	β
1	0	0	0.00	1	1	0
2	3	0	0.02	0.98	1	0
3	8	0	0.06	0.94	1	0
4	17	1	0.13	0.87	0.99	0.01
5	34	3	0.26	0.74	0.98	0.02
6	64	8	0.49	0.51	0.94	0.06
7	100	15	0.77	0.23	0.88	0.12
8	115	28	0.88	0.12	0.78	0.22
9	123	41	0.95	0.05	0.68	0.32
10	127	51	0.98	0.02	0.61	0.39
11	129	64	0.99	0.01	0.51	0.49
12	129	78	0.99	0.01	0.40	0.60
13	130	92	1	0	0.29	0.71
14	130	105	1	0	0.19	0.81
15	130	119	1	0	0.08	0.92

На рис. 3.1а показано залежність першої достовірності D_1 від радіуса d_1 контейнера класу X_1^o , а на рис. 3.2б – залежність помилки другого роду β так само від радіуса d_1 , одержаних в процесі оптимізації за базовим алгоритмом навчання (3.1.2) у рамках ІЕІ-технології. На рис. 3.1 темна ділянка графіку означає робочу (допустиму) область визначення функції інформаційного КФЕ навчання системи.



а)



б)

Рисунок 3.1– Графіки залежності значення точнісних характеристик від радіуса контейнера класу X_1^o : а) перша достовірність D_1 ; б) помилка другого роду β

Аналіз рис. 3.2 показує, що перша достовірності D_1 та помилка другого роду β із збільшенням радіусу d_1 так само збільшуються до своїх граничних значень, тобто відповідно помилка першого роду α і друга достовірність D_2 зменшуються до нуля.

3.2 Оптимізація контрольних допусків на ознаки розпізнавання

На відміну від відомих методів [66] нормалізація образів у рамках ІЕІ-технології здійснюється безпосередньо в процесі навчання системи шляхом цілеспрямованої трансформації апріорного нечіткого розбиття простору ознак на класи розпізнавання в чітке розбиття еквівалентності, що дозволяє побудувати безпомилковий за навчальною матрицею класифікатор. Така трансформація здійснюється за допомогою допустимих перетворень у дискретному парацептуальному просторі, серед яких важливу роль відіграє оптимізація системи контрольних допусків на ознаки розпізнавання.

Розглянемо оптимізацію контрольних допусків як за паралельним алгоритмом, коли поля допусків для всіх ознак змінювалися одночасно, так і за послідовним алгоритмом. Якщо паралельний алгоритм оптимізації, який виправдано застосовувати, наприклад, при обробленні великих масивів відеоінформації при розпізнаванні зображень, дозволяє підвищити оперативність навчання, то він втрачає в точності обчислення глобального максимуму інформаційного КФЕ навчання системи. Послідовний алгоритм оптимізації системи контрольних допусків навпаки втрачає в оперативності навчання, але дозволяє часто побудувати, наприклад, в системах «технічного зору» для образів $2D$ безпомилковий класифікатор навіть без оптимізації інших просторово-часових параметрів функціонування. Слабким місцем існуючих послідовних алгоритмів оптимізації системи контрольних допусків є неформалізований вибір стартових допусків на ознаки розпізнавання. Невдалий вибір таких допусків може вивести систему на початку оптимізації за межі робочої (допустимої) області визначення функції КФЕ навчання, що призводить до збільшення часу оптимізації системи контрольних допусків.

За умови обґрунтування у загальному випадку нечіткої гіпотези компактності реалізацій образу оптимізацію системи контрольних допусків на ознаки розпізнавання розглянемо в рамках ІЕІ-технології постановку задачі інформаційного синтезу СППР для гіперсферичного класифікатора, в

якому контейнери класів розпізнавання відновлюються на кожному кроці навчання в радіальному базисі дискретного простору Хеммінга.

Нехай дано алфавіт класів розпізнавання $\{X_m^o \mid m = \overline{1, M}\}$, навчальна матриця типу “об’єкт властивість” $\|y_{m,i}^{(j)}\|$, $i = \overline{1, N}$, $j = \overline{1, n}$, де N , n – кількість ознак розпізнавання та векторів-реалізацій образу відповідно. Задано структурований вектор параметрів функціонування системи, що навчається, $g = \langle x_1, d_1, \delta \rangle$, який складається з еталонного вектора-реалізації x_1 найбільш бажаного (базового) класу X_1^o , геометричного параметра d_m – кодової відстані гіперповерхні контейнера K_1^o класу X_1^o від вершини еталонного вектора $x_1 \in X_1^o$. Задано допустимі області значень відповідних параметрів: $x_1 \in \Omega_B^{|N|}$, де $\Omega_B^{|N|}$ – бінарний простір ознак потужності N ; радіуси $d_1 \in [0; d(x_1 \oplus x_c) - 1]$, де x_c – еталонний вектор-реалізація сусіднього (найближчого) класу X_c^o , і параметра поля допусків $\delta \in [0; \delta_H / 2]$, де δ_H – нормоване (експлуатаційне) поле допусків на ознаки розпізнавання.

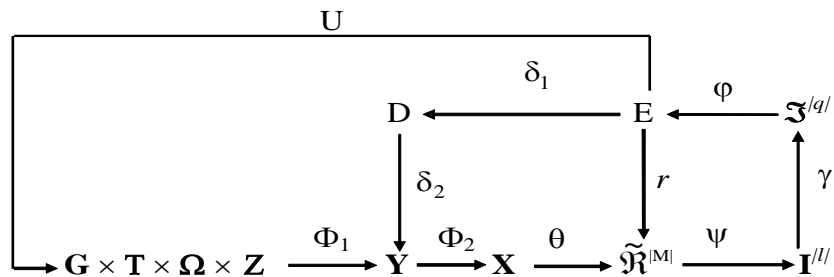
Треба на етапі навчання за апріорно класифікованими реалізаціями нечітких образів побудувати оптимальне в інформаційному розумінні чітке розбиття $\mathfrak{R}^{|M|}$ дискретного простору ознак Ω_B на M класів розпізнавання шляхом ітераційної максимізації усередненого за алфавітом $\{X_m^o\}$ значення інформаційного КФЕ навчання АСК:

$$\bar{E}^* = \frac{1}{M+1} \sum_{m=1}^{M+1} \max_{\{d\}} E_m, \quad (3.2.1)$$

де E_m – інформаційний КФЕ навчання системи розпізнавати реалізації класу X_m^o ;

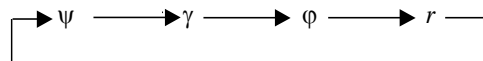
$\{d\}$ – множина кроків навчання.

Категорійна модель у вигляді діаграми відображення множин, що застосовуються при оптимізації СКД на ознаки розпізнавання в процесі навчання АСК, має такий вигляд [30]:



(3.2.2)

За ІЕІ-технологією на етапі навчання системи оператор $\theta: X \rightarrow \tilde{\mathfrak{R}}^{|M|}$ відновлює на кожному кроці навчання у загальному випадку нечітке розбиття $\tilde{\mathfrak{R}}^{|M|}$. Оператор класифікації $\Psi: \tilde{\mathfrak{R}}^{|M|} \rightarrow I^{|l|}$ перевіряє основну статистичну гіпотезу $\gamma_1: y_{m,i}^{(j)} \in X_m^o$, де $I^{|l|}$ –множина статистичних гіпотез. Оператор γ визначає множину точнісних характеристик процесу навчання $\mathfrak{Z}^{|q|}$, де $q=l^2$, а оператор ϕ обчислює терм-множину E значень інформаційного критерію оптимізації. Оператор r корегує розбиття $\tilde{\mathfrak{R}}^{|M|}$ шляхом ітераційного пошуку глобального максимуму критерію (3.2.1) у робочій області визначення його функції. У діаграмі (3.2.2) контур операторів



реалізує базовий алгоритм навчання (3.1.2), який на кожному кроці навчання обчислює значення критерію $E_m \in E$, здійснює пошук глобального максимуму його функції в робочій області її визначення і оптимізує геометричні параметри розбиття $\tilde{\mathfrak{R}}^{|M|}$. Оператор $U: E \rightarrow G \times T \times \Omega \times Z$ регламентує процес навчання і дозволяє оптимізувати параметри його плану,

які визначають, наприклад, обсяг і структуру випробовувань, черговість розгляду класів розпізнавання та інше. Контур оптимізації СКД, яка утворює терм-множину D , замикається послідовно оператором $\delta_1 : E \mapsto D$ і оператором $\delta_2 : D \rightarrow Y$, який змінює реалізації образу в процесі максимізації критерію $E_m \in E$.

Розглянемо підхід до оптимізації системи контрольних допусків на ознаки розпізнавання в рамках ІЕІ-технології. Існує декілька можливих стратегій зміни поля допусків $\delta_{k,i}$, серед яких відмітимо дві основні [30]:

симетрична стратегія $S_1(\overset{\rightarrow}{\leftarrow} \text{var } A_{HK}, \overset{\leftarrow}{\rightarrow} \text{var } A_{BK})$, яка є виправданою, наприклад, за умови підтвердження розвідувальним аналізом збігання номінального значення A_0 з теоретичним центром розсіювання значень навчальної вибірки $\{y_{m,i}^{(j)} \mid j = \overline{1, n}\}$;

асиметрична стратегія $S_2(\overset{\leftarrow}{\rightarrow} \text{var } A_{HK}, \overset{\leftarrow}{\rightarrow} \text{var } A_{BK})$, яка має місце при відхиленні значення A_0 від центра розсіювання значень вибірки $\{y_{m,i}^{(j)} \mid j = \overline{1, n}\}$.

Задача оптимізації контрольних допусків на ознаки розпізнавання є частинною задачею інформаційного синтезу, в якій необхідно визначити екстремальні значення параметра поля допусків δ :

$$\delta^* = \arg \max_{G_\delta} \{ \max_{G_\Omega} \{ \max_{G_d} \bar{E} \} \},$$

де G_δ – допустима область значень параметра поля допусків δ ;
 G_Ω – допустима область значень ознак розпізнавання; G_d – допустима область значень радіусів контейнерів класів розпізнавання $\{X_m^o\}^\Lambda$.

Алгоритм оптимізації контрольних допусків, як і інших параметрів навчання АСК, у рамках ІЕІ-технології полягає у наближенні глобального максимуму інформаційного критерію оптимізації до граничного його значення в області значень функції критерію. У праці [31] доведено збіжність алгоритму послідовної оптимізації контрольних допусків на ознаки розпізнавання, який має таку структуру ітераційної процедури оптимізації:

$$\{\delta_{K,i}^*\} = \langle \arg\{\max_{G_{\delta_i}}\{\max_{G_E} \left[\bigotimes_{l=1}^L \max_{G_{d_1}} E_1^{(l)} \right]\}\} \rangle, \quad i = \overline{1, N}, \quad (3.2.3)$$

де G_{δ_i} – область допустимих значень параметра поля контрольних допусків для i -ї ознаки; G_E – області допустимих значень критерію оптимізації; G_{d_1} – області допустимих значень кодової відстані d_1 ; \otimes – символ операції повторення; L – кількість прогонів ітераційної процедури послідовної оптимізації контрольних допусків; $E_1^{(l)}$ – значення інформаційного критерію при оптимізації контрольного поля допусків для i -ї ознаки розпізнавання на l -му прогоні процедури оптимізації.

Розглянемо реалізацію алгоритму послідовної оптимізації контрольних допусків на ознаки розпізнавання за процедурою (3.2.3). Вхідні дані: масив реалізацій образу $\{y_m^{(j)} \mid m = \overline{1, M}; j = \overline{1, n}\}$; стартова система контрольних допусків $\{\delta_{K,i}\}$ і система нормованих допусків $\{\delta_{H,i}\}$, яка визначає область значень відповідних контрольних допусків. Попередньо для кожної ознаки визначається ціна градації Δ_i , на яку змінюється i -та ознака.

Схема алгоритму послідовної оптимізації контрольних допусків така:

Крок 1. Обнуління лічильника прогонів процедури оптимізації параметрів навчання: $l:=0$.

Крок 2. Для стартової системи допусків обчислюється за базовим алгоритмом навчання (3.1.2) значення функції $E_{\max,1}^{(l)}$.

Крок 3. Формування лічильника прогонів: $l: l+1$.

Крок 4. Обнуління лічильника ознак розпізнавання: $i:=0$.

Крок 5. Формування лічильника ознак розпізнавання: $i: i+1$.

Крок 6. Визначення екстремального значення параметра поля допусків $\hat{\delta}_{K,i}^{(l)}$ за процедурою (3.2.3), в якій внутрішній цикл оптимізації реалізує базовий алгоритм навчання.

Крок 7. $\hat{\delta}_{K,i}^{(l)} := \max_{\delta_{K,i} \in \delta_{H,i}} \delta_{K,i}^{(l)}$.

Крок 8. Якщо $i \leq N$, то виконується крок 5, інакше крок 9.

Крок 9. Якщо $\left| E_{\max,1}^{(l-1)} - E_{\max,1}^{(l)} \right| \leq \varepsilon$, де ε – будь-яке мале позитивне число, то виконується крок 10, інакше крок 3.

Крок 10. $\{\delta_{K,i}^*\} := \{\hat{\delta}_{K,i}^{(l)}\}$ і “Зупин”.

Для формування стартової системи контрольних допусків застосуємо паралельний алгоритм їх оптимізації, що оптимізує контрольні допуски на всі ознаки одночасно. Вхідні дані такі самі як і для алгоритму послідовної оптимізації, але за область значень параметра δ приймається інтервал $[1; \delta_H/2]$, де δ_H – ширина нормованого поля допусків.

Розглянемо кроки реалізації алгоритму паралельної оптимізації:

Крок 11. Обнулюється лічильник кроків зміни параметра $\delta: l:=0$.

Крок 12. Запускається лічильник: $l:=l+1$ і обчислюються нижні та верхні контрольні допуски для всіх ознак: $\{A_{HK,i}[l] := y_{m,i} - \delta[l]\}$ і $\{A_{BK,i}[l] := y_{m,i} + \delta[l]\}$, $i = \overline{1, N}$, відповідно.

Крок 13. Реалізується базовий алгоритм навчання.

Крок 14. Якщо $E_1^*[l] \geq E_1^*[l-1]$, то виконується крок 15, інакше – 16.

Крок 15. Якщо $\delta \leq \delta_H/2$, то виконується крок 12, інакше крок 16.

Крок 16. $\{A_{HK,i}^* := A_{HK,i}[l-1]\}$; $\{A_{BK,i}^* := A_{BK,i}[l-1]\}$, $i = \overline{1, N}$;

$E_1^* := E_1^*[l-1]$ і “ЗУПИН”.

Як приклад розглянемо реалізацію алгоритму оптимізації системи контрольних допусків у рамках ІЕІ-технології в процесі навчання СППР, яка є складовою частиною АСКТП виробництва складного мінерального добрива *НКР* (азот – калій – фосфор) у ВАТ «Сумихімпром». Для наочності розглянемо оптимізацію апріорного нечіткого розбиття для трьох класів: X_1^o , X_2^o і X_3^o , які відрізнялися один від одного вмістом хімічних компонентів. Тут клас X_1^o – характеризує найкращий технологічний режим, що забезпечує відповідність кінцевого продукту заданому стандарту, а інші класи відповідно характеризують порушення техпроцесу, що призвело до відхилення вмісту азоту і фосфору від норми. Навчальна матриця складалась із 40 векторів-реалізацій для кожного класу. Структуровані вектори-реалізації класів мали 41 ознаку розпізнавання (із них 28 – технологічні параметри, значення яких візуально відображалися на моніторі, і 13 – результати поточного хімічного аналізу, які періодично вводилися в систему оператором).

На рис. 3.2 – 3.4 показано графіки залежності функції КФЕ (2.4.) від радіусів контейнерів відповідних класів при реалізації базового алгоритму навчання СППР.

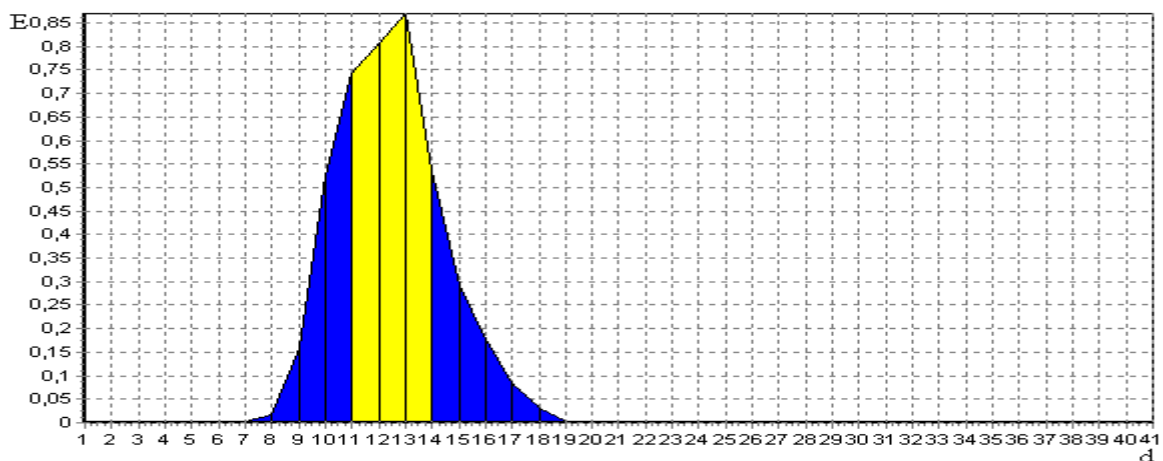


Рисунок 3.2– Оптимізація контейнера класу X_1^o за базовим алгоритмом навчання

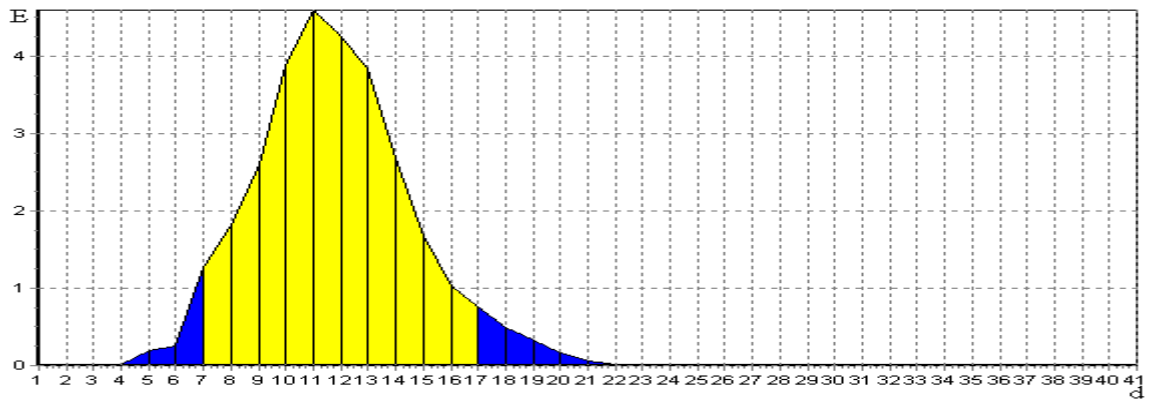


Рисунок 3.3– Оптимізація контейнера класу X_2^o за базовим алгоритмом навчання

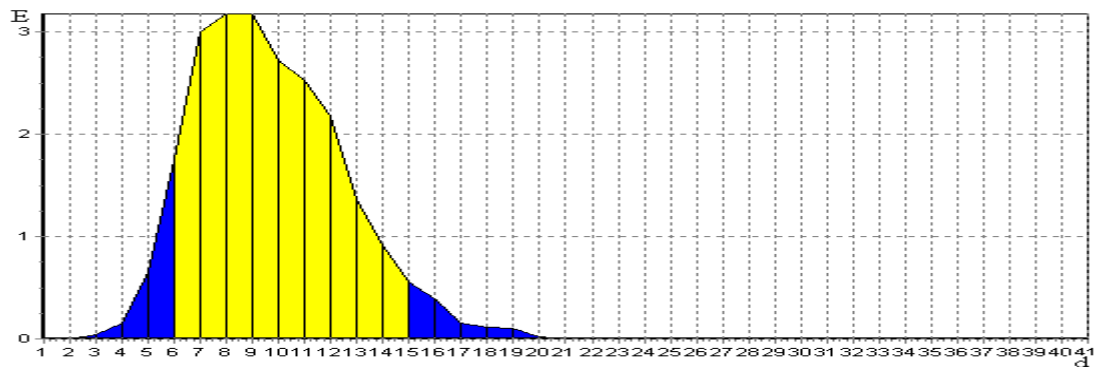


Рисунок 3.4 – Оптимізація контейнера класу X_3^o за базовим алгоритмом навчання

Тут і далі світла область графіка означає робочу область визначення інформаційного КФЕ навчання системи. Аналіз рис. 3.4-3.6 показує, що оптимальні радіуси контейнерів у кодових одиницях дорівнюють відповідно: $d_1^* = 13$, $d_2^* = 11$ і $d_3^* = 9$.

При реалізації послідовного алгоритму оптимізації для визначення стартових допусків на ознаки розпізнавання було попередньо запущено алгоритм паралельної оптимізації допусків. На рис. 3.5 показано динаміку зміни КФЕ навчання системи в процесі оптимізації параметру поля допусків δ за паралельним алгоритмом.

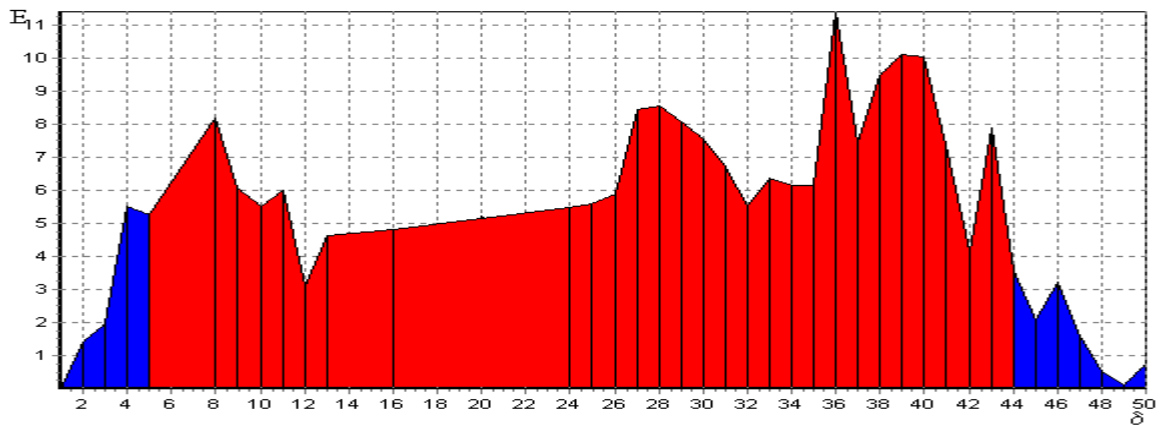


Рисунок 3.5 –Графік залежності критерію Кульбака від параметра поля контрольних допусків

Як показано на рис. 3.5 оптимальне значення параметра δ в робочій області визначення функції критерію (2.4.6) дорівнює $\delta^* = \pm 36$ відносних одиниць для всіх ознак розпізнавання. Крім того, усереднене значення КФЕ є значно більшим ($\bar{E}^* = 11,32$) у порівнянні з базовим алгоритмом навчання (рис. 3.4 – 3.6), але воно не досягає граничного значення критерію ($\bar{E}^* = 13,96$ при $n=40$ і $r = 2$). Після визначення стартової системи контрольних допусків $\{\delta^*\}$ запускався алгоритм послідовної оптимізації (3.2.3) за тим самим критерієм (2.4.6). На рис. 3.6 – 3.9 наведено графіки залежності КФЕ від радіусів контейнерів відповідних класів після оптимізації контрольних допусків за послідовним алгоритмом (3.2.3).

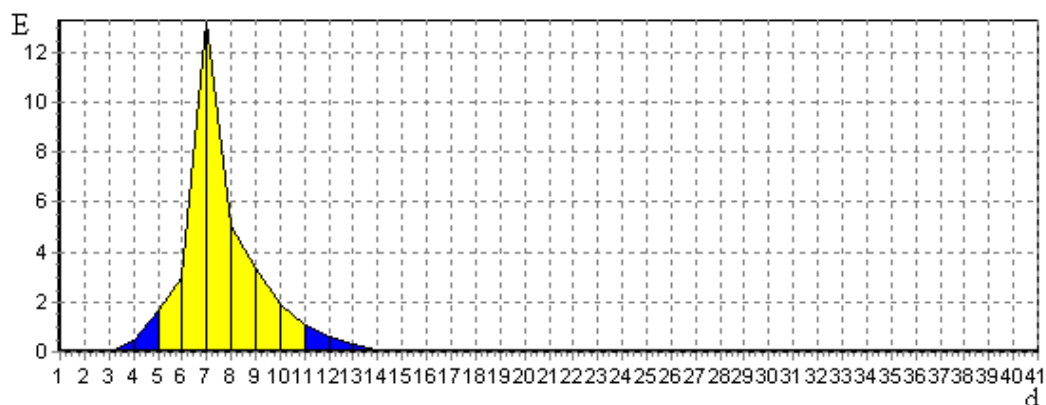


Рисунок 3.6 – Оптимізація контейнера класу X_1^o за послідовним алгоритмом

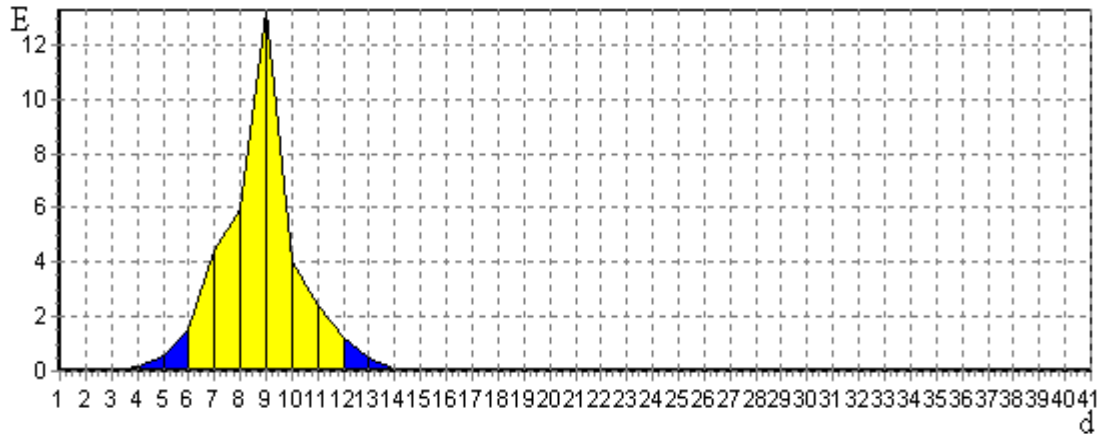


Рисунок 3.7 – Оптимізація контейнера класу X_2^o за послідовним алгоритмом

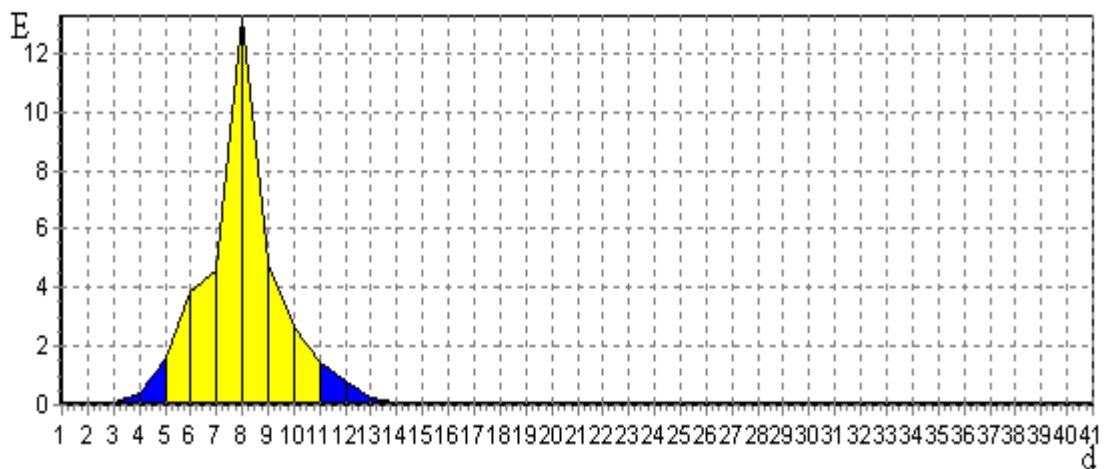


Рисунок 3.8 – Оптимізація контейнера класу X_3^o за послідовним алгоритмом

Аналіз рис. 3.6 – 3.8 показує, що оптимальні радіуси відповідних контейнерів класів розпізнавання дорівнюють: $d_1^* = 7$, $d_2^* = 9$, $d_3^* = 8$. При цьому середнє значення радіусів контейнерів дорівнює $\bar{d}^* = 5,33$, що значно менше відповідного середнього значення для базового алгоритму навчання. Це свідчить про зменшення ступеню перетину класів в процесі оптимізації контрольних допусків на ознаки розпізнавання. Аналіз зміни величини середніх значень КФЕ в процесі оптимізації СКД показав, що при реалізації

базового алгоритму навчання вона дорівнює $\bar{E} = 2,90$, при оптимізації за паралельним алгоритмом – $\bar{E} = 11,32$, а при оптимізації за послідовним алгоритмом – $\bar{E} = 13,07$. Таким чином, застосування алгоритму послідовної оптимізації системи контрольних допусків забезпечує наближення значення КФЕ до граничного ($E_{cp} = 13,96$).

Запропонований алгоритм паралельно-послідовної оптимізації контрольних допусків на ознаки розпізнавання дозволяє побудувати вирішальні правила, достовірність яких наближається до асимптотичної (граничної) достовірності. При цьому реалізація алгоритму паралельної оптимізації контрольних допусків дозволяє вибрати стартову систему допусків, що забезпечує знаходження АСК, що навчається, в робочий області визначення функції інформаційного КФЕ навчання системи.

Таким чином, поєднання паралельного і послідовного алгоритмів оптимізації контрольних допусків на ознаки розпізнавання дозволяє підвищити оперативність навчання та забезпечити наближену до асимптотичної достовірність прийняття рішень на етапі екзамену.

3.3 Максимізація міжцентрових відстаней для заданого алфавіту класів розпізнавання

Одна із переваг ІЕІ-технології аналізу та синтезу АСК слабо формалізованими процесами у порівнянні з відомими методами автоматичної класифікації [66] полягає у можливості поєднання етапів нормалізації образу і машинного навчання, на якому будуються безпосередньо вирішальні правила. У загальному випадку нормалізація математичного опису на вході класифікатора за ІЕІ-технологією полягає в реалізації багатоциклічної ітераційної процедури оптимізації за інформаційним КФЕ просторово-часових параметрів функціонування, які впливають на функціональну ефективність АСК, що навчається. Серед таких параметрів важливу роль відіграє рівень селекції (квантування) координат двійкових еталонних

векторів-реалізацій, які визначають геометричні центри контейнерів класів розпізнавання при їх цілеспрямованому відновленні в радіальному базисі простору ознак розпізнавання.

Рівнем селекції координат еталонного двійкового вектора будемо називати рівень квантування дискрет полігона емпіричних частот потрапляння значень ознак розпізнавання у свої поля контрольних допусків.

Полігон будується для кожного класу так: по осі абсцис відкладаються ранги ознак розпізнавання, які відповідають номерам ознак у векторі-кортежі $x_m^{(j)}$, а по осі ординат – відносні частоти $\omega_{m,i} = n_i / n$, де n_i – кількість випробувань, при яких значення i -ї ознаки знаходиться в своєму полі контрольних допусків. За замовчуванням на практиці приймається рівень селекції $\rho_m = 0,5$. Якщо у процесі навчання знайдено оптимальне значення $\rho_m^* < 0,5$, то це свідчить про виявлення принаймні однієї інформативної ознаки, латентність якої полягала саме в її малій частоті появи. Якщо знайдено $\rho_m^* > 0,5$, то це свідчить про негативний вплив на прийняття рішень принаймні однієї “заважаючої” ознаки.

Оптимізація таких рівнів селекції у рамках ІЕІ-технології дозволяє розв’язувати такі важливі для підвищення асимптотичної достовірності класифікатора задачі:

- центрування контейнерів класів розпізнавання, тобто наближення геометричних центрів контейнерів класів розпізнавання до їх емпіричних центрів розсіювання реалізацій образу;
- максимізація середньої міжкласової кодової відстані для заданого алфавіту у відповідності з максимально-дистанційним принципом теорії розпізнавання образів.

Параметр ρ_m непрямо пов’язаний з інформативністю ознак розпізнавання. Наприклад, якщо в процесі навчання знайдено оптимальне значення $\rho_m^* < 0,5$, то це свідчить про виявлення, у крайньому разі, однієї інформативної ознаки, латентність якої полягала саме в її малій частоті

появи. Якщо знайдено $\rho_m^* > 0,5$, то це свідчить про негативний вплив на прийняття рішень, у крайньому разі, однієї “заважаючої” ознаки.

Як показано в підрозділі 3.2 застосування паралельно-послідовного алгоритму оптимізації параметрів навчання дозволяє підвищити функціональну ефективність СППР. Тому доцільно й оптимізацію рівнів селекції координат еталонних векторів здійснювати за паралельно-послідовним алгоритмом, що дозволяє за алгоритмом паралельної оптимізації визначити стартову квазиоптимальну систему рівнів селекції, а за послідовним – оптимальні рівні селекції для всіх координат еталонних векторів-реалізацій образу.

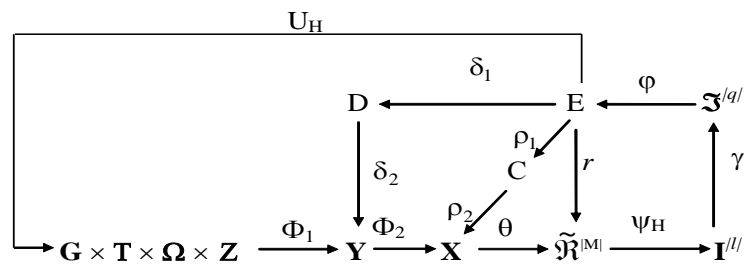
Оптимізацію рівнів селекції ρ_m за умови обґрунтування у загальному випадку нечіткої гіпотези компактності реалізацій образу будемо здійснювати на базі інформаційно-екстремального алгоритму навчання для гіперсферичного класифікатора, в якому контейнери класів розпізнавання відновлюються на кожному кроці навчання в радіальному базисі дискретного простору Хеммінга. Спочатку розглянемо алгоритм оптимізації рівнів селекції координат двійкової реалізації найбільш бажаного для ОПР першого класу X_1^o , оскільки для інших класів система контрольних допусків на ознаки розпізнавання залишається незмінною.

Нехай дано алфавіт класів розпізнавання $\{X_m^o \mid m = \overline{1, M}\}$, навчальна матриця типу “об’єкт-властивість” $\|y_{m,i}^{(j)}\|$, $i = \overline{1, N}$, $j = \overline{1, n}$, де N , n – кількість ознак розпізнавання та реалізацій образу відповідно. Задано структурований вектор параметрів функціонування системи, що навчається розпізнавати реалізації класу X_1^o : $g = \langle x_1, d_1, \delta, \rho_1 \rangle$, який складається з еталонної реалізації x_1 класу X_1^o , геометричного параметра d_1 – кодової відстані гіперповерхні контейнера K_1^o класу X_1^o від вершини еталонної реалізації $x_1 \in X_1^o$. Задано допустимі області значень відповідних параметрів: $x_1 \in \Omega_B^{|N|}$, де $\Omega_B^{|N|}$ – бінарний простір ознак потужності N ;

$d_1 \in [0; d(x_1 \oplus x_c) - 1]$, де x_c – еталонна реалізація сусіднього (найближчого до X_1^o) класу X_c^o , параметра контрольного поля допусків $\delta \in [0; \delta_H / 2]$, де δ_H – нормоване (експлуатаційне) поле допусків на ознаки розпізнавання і рівень селекції ρ_1 координат еталонної реалізації x_1 , $\rho_1 \in [0; 1]$.

Треба на етапі навчання за побудувати оптимальне в інформаційному розумінні чітке розбиття $\mathfrak{R}^{|\Lambda|}$ дискретного простору ознак Ω_B на M класів розпізнавання шляхом ітераційної оптимізації координат вектора параметрів функціонування g_1 за умови, що значення усередненого за алфавітом $\{X_m^o\}$ інформаційного КФЕ навчання АСК набуває глобального максимуму в робочій (допустимій) області визначення його функції.

Діаграма відображень множин, що застосовуються при оптимізації рівнів селекції координат еталонних векторів у процесі навчання СППР в режимі ФКА згідно з (2.2.3) має вигляд



(3.3.1)

У діаграмі (3.3.1) контур (2.3.3) безпосередньо оптимізує за інформаційним КФЕ геометричні параметри розбиття $\tilde{\mathfrak{R}}^{|\Lambda|}$, контур (2.3.4) – систему контрольних допусків на ознаки розпізнавання, а контур (2.3.5) – рівні селекції координат двійкових еталонних векторів реалізацій.

Тоді оптимальний рівень селекції ρ_1 координат еталонної реалізації $x_1 \in X_1^o$ визначається у результаті реалізації багато цикличної ітераційної процедури:

$$\rho_1^* = \arg \max_{G_\rho} \{ \max_{G_\delta} \{ \max_{G_\Omega \cup G_d} \bar{E}^* \} \}, \quad (3.3.2)$$

де G_ρ – область допустимих значень параметра ρ_1 ; G_δ – область допустимих значень параметра δ ; G_Ω – область допустимих значень ознак розпізнавання; G_d – область допустимих значень параметра d_1 .

Оптимальні рівні селекції координат інших еталонних реалізацій визначаються за процедурою максимізації інформаційного КФЕ навчання СППР розпізнавати реалізації класу X_m^o , $m = \overline{2, \Lambda}$:

$$\rho_m^* = \arg \max_{G_\rho} \{ \max_{G_\Omega \cup G_d} E_m \}, X_m^o, m = \overline{2, \Lambda}. \quad (3.3.3)$$

Таким чином, процедура (3.3.3) на відміну від процедури (3.3.2) не реалізує алгоритм оптимізації контрольних допусків на ознаки розпізнавання, оскільки вони визначаються тільки для пріоритетного для ОПР класу X_1^o . При цьому внутрішні цикли алгоритмів (3.3.2) і (3.3.3) реалізуються послідовно операторами, які утворюють базовий контур (2.3.3). Для алгоритму (3.3.2) наступний цикл реалізується послідовністю операторів контуру (2.3.4) і зовнішній цикл – послідовністю операторів контуру оптимізації рівнів селекції (2.3.5).

Розглянемо детальніше етапи алгоритму послідовної оптимізації рівнів селекції:

Крок 17. Обнулюється лічильник ознак розпізнавання: $k:=0$;

Крок 18. Встановлюємо значення для верхнього та нижнього рівнів селекції відповідно: $\rho n:=0.3, \rho v:=0.7$;

Крок 19. Встановлюємо крок зміни рівня селекції: $\rho c:=0.1$;

Крок 20. Обнулюється лічильник кроків оптимізації рівня селекції: $z:=0$;

Крок 21. Встановлюємо значення поточного рівня селекції: $\rho z := \rho n - \rho c$;

Крок 22. $\rho z := \rho n + \rho c$;

Крок 23. Знаходиться та запам'ятовується оптимальне значення КФЕ E_z^* за робочими формулами (2.4.8) або (2.4.9);

Крок 24. Якщо $\rho z < \rho v$, то виконується крок 22, інакше крок 25;

Крок 25. Знаходиться максимальне значення КФЕ серед усіх E_z^* : $E_\rho^* = \langle \max E_z^* \rangle$; Якщо маємо декілька однакових значень, то вибираємо значення з найменшим усередненим коефіцієнтом нечіткої компактності реалізації образу:

Крок 26. Обчислення коефіцієнта нечіткої компактності:

$$L_\delta = \frac{1}{\Lambda} \sum_{m=1}^{\Lambda} \frac{d_m^*}{d(x_m \oplus x_c)}. \quad (3.3.4)$$

Крок 27. Система контрольних допусків, яка відповідає E_ρ^* при мінімальному L_δ вважається оптимальною;

Крок 28. Для k -ої ознаки розпізнавання встановлюємо оптимальний рівень селекції, який відповідає E_ρ^* .

Крок 29. $k := k + 1$;

Крок 30. Якщо $k \leq N$, де N –кількості ознак розпізнавання, то виконується крок 20, інакше крок 31;

Крок 31. Зупин.

Таким чином, алгоритм оптимізації рівнів селекції координат двійкових еталонних векторів-реалізацій образу, як і інших параметрів навчання інтелектуальної СППР у рамках ІЕІ-технології полягає у наближенні глобального максимуму інформаційного критерію оптимізації до найбільшого його значення в області значень функції критерію. Тому важливого значення набуває дослідження збігу такого алгоритму. Аналогічно праці [31], де доведено збіг алгоритму оптимізації контрольних допусків, розглянемо збіг алгоритму послідовної оптимізації рівнів селекції координат

двійкових еталонних векторів-реалізацій класу X_m^o . Введемо такі позначення:

$v_m = \langle v_{m,1}, \dots, v_{m,i}, \dots, v_{m,N} \rangle$ – структурований вектор значень дискрет полігону частот “успіхів”, під якими розуміється знаходження значення ознаки розпізнавання в своєму контрольному полі допусків;

$\rho_m = \langle \rho_{m,1}, \dots, \rho_{m,i}, \dots, \rho_{m,N} \rangle$ – структурований вектор стартових рівнів селекції координат двійкових еталонних векторів-реалізацій класу X_m^o ;

l – кількість прогонів ітераційної процедури послідовної оптимізації рівнів селекції координат двійкових еталонних векторів-реалізацій образу;

$E_{\max}^{(l)}(v_m, \rho_m)$ – максимальне значення КФЕ в робочій області його визначення при l -му прогоні ітераційної процедури;

E^* – найбільший глобальний максимум функції КФЕ в області її значень;

$\rho_{m,i}^{(l)}$ – значення рівня селекції i -ої координати двійкового еталонного вектора-реалізацій класу X_m^o , яке отримано при l -му прогоні ітераційної процедури;

$\hat{\rho}_{m,i}^{(l)}$ – екстремальне значення рівня селекції i -ої координати двійкового еталонного вектора x_m , яке отримано при l -му прогоні ітераційної процедури;

$\rho_{m,i}^*$ – оптимальне значення рівня селекції i -ої координати двійкового еталонного вектора-реалізацій класу X_m^o , яке одержано при реалізації ітераційної процедури:

$$\rho_{m,i}^* = \arg \max_{G_E} E_m^*.$$

З урахуванням (3.3.2) і введених позначень структурований багато цикличний ітераційний алгоритм послідовної оптимізації рівнів селекції координат двійкових еталонних векторів-реалізацій класу X_m^o приймає вигляд

$$\{\rho_{m,i}^*\} = \langle \arg\{\max_{G_{\rho_i}} \{\max_{G_{\delta_i}} \left[\bigotimes_{l=1}^L \max_{G_E \cup G_{d_l}} E^{(l)}(v_m, \rho_m) \right]\} \} \rangle, \quad i = \overline{1, N}, \quad (3.3.5)$$

де G_{ρ_i} – область допустимих значень рівня селекції i -ої координати двійкового еталонного вектора-реалізації класу X_m^o ; G_{δ_i} – область допустимих значень поля контрольних допусків для i -ої ознаки розпізнавання; G_E – область допустимих значень критерію оптимізації; G_{d_m} – область допустимих значень кодової відстані-радіуса d_m ; \otimes – символ операції повторення.

Розглянемо послідовність $u_{m,N}^{(l)} = \langle u_{m,1}^{(l)}, \dots, u_{m,i}^{(l)}, \dots, u_{m,N}^{(l)} \rangle$, де

$$u_{m,i}^{(l)} = E_{\max}^{(l)}(v_m, \hat{\rho}_{m,1}, \dots, \hat{\rho}_{m,i}, \rho_{m,i+1}, \dots, \rho_{m,N}) - E_{\max}^{(l)}(v_m, \hat{\rho}_{m,1}, \dots, \hat{\rho}_{m,i-1}, \rho_{m,i}, \dots, \rho_{m,N}).$$

Т в е р д ж е н н я 3.3.1. Послідовність $u_{m,N}^{(l)}$ монотонно спадає і обмежена знизу.

Д о в е д е н н я. Для доведення умови твердження 3.3.1 достатньо показати, що $u_{i-1}^{(l)} \geq u_i^{(l)}$. Це легко доводиться за індукцією. Нехай при $l=1$ для першої ознаки знайдено екстремальне значення $\rho_{m,1}^{(1)}$ за умови, що значення контрольних допусків для інших ознак залишаються стартовими. Тоді має місце:

$$E_{\max}^{(1)}(v_m, \hat{\rho}_{m,1}, \dots, \rho_{m,i}, \dots, \rho_{m,N}) \geq E_{\max}^{(1)}(v_m, \rho_{m,1}, \rho_{m,2}, \dots, \rho_{m,i}, \dots, \rho_{m,N}).$$

Оскільки за властивістю інформаційного критерію має місце $E_{\max}^{(l)}(v_m, \rho_m) \geq 0$, то відношення рівності може бути тільки за умови, що стартове значення $\rho_{m,1}$ дорівнює екстремальному. Так само справедливо і для всіх N екстремальних значень $\{\hat{\rho}_{m,i}^{(l)} \mid i = \overline{1, N}\}$:

$$E_{\max}^{(1)}(v_m, \hat{\rho}_{m,1}, \dots, \hat{\rho}_{m,i}, \dots, \hat{\rho}_{m,N}) \geq E_{\max}^{(1)}(v_m, \hat{\rho}_{m,1}, \dots, \hat{\rho}_{m,i}, \dots, \hat{\rho}_{m,N-1}, \rho_{m,N}).$$

Таким чином, послідовність $\{u_i^{(l)} \mid i = \overline{1, N}\}$, є спадною і обмеженою знизу, оскільки її члени додатні. Але не ясно, чи є послідовність стаціонарною, тобто чи існує таке L , що для будь-якого $l > L$ має місце $u_{i-1}^{(l)} = u_i^{(l)}$. Відповідь на це запитання дає така теорема.

Т е о р е м а 3.3.1. Ітераційний алгоритм послідовної оптимізації рівнів селекції координат двійкових еталонних векторів-реалізацій образу (3.3.5) збігається з імовірністю одиниця.

Д о в е д е н н я. Для доведення теореми необхідно та достатньо показати, що екстремальні значення рівнів селекції координат двійкових еталонних векторів-реалізацій $\{\hat{\rho}_{m,i}^{(l)}\}$ збігаються відповідно до оптимальних $\{\rho_{m,i}^*\}$, $i = \overline{1, N}$. Припустимо, що послідовність $u_N^{(l)}$, яка за твердженням спадна і обмежена знизу, збігається на L -му прогоні ітераційної процедури і вона є стаціонарною, тобто $u_N^{(L+1)} = u_N^{(L)}$. Тоді маємо: $E_{\max}^{(L+1)}(v_m, \hat{\rho}_m) = E_{\max}^{(L)}(v_m, \hat{\rho}_m)$. Але звідси не обов'язково витікає, що $\{\hat{\rho}_{m,i}^{(L)}\} = \{\rho_{m,i}^*\}$, оскільки функція $E_m(v_m, \hat{\rho}_m)$ не є взаємно-однозначною. Покажемо концептуально, що все-таки існує $\hat{\rho}_{m,i}^{(L)} = \rho_{m,i}^*$, $i = \overline{1, N}$, для будь-

якого $l \geq L$. Для цього скористаємося методом евристичного доведення, який базується на концептуальних положеннях і категоріях методу аналізу і синтезу інтелектуальної СППР у рамках ІЕІ-технології. У силу дистанційно-максимального принципу розпізнавання образів, у процесі оптимізації розбиття $\tilde{\mathfrak{R}}^{[2]}$ для класу X_m^o і його найближчого (сусіднього) класу X_c^o повинна виконуватися умова: $d(x_m \oplus x_c) \rightarrow \max_{G_d}$ з обмеженнями (1.4.3) і (2.1.1).

Нехай існує для функції $E_{\max}^{(l)}(v_m, \rho_m)$ множина екстремальних параметрів: $\langle \hat{\rho}_{m,i,\min}^{(l)}, \hat{\rho}_{m,i,\min+1}^{(l)}, \dots, \hat{\rho}_{m,i,\max-1}^{(l)}, \hat{\rho}_{m,i,\max}^{(l)} \rangle$, тобто в робочій області визначення інформаційного КФЕ навчання СППР має місце ділянка екстремальних значень рівнів селекції типу “плато”. Оскільки збільшення рівня селекції $\rho_{m,i}^{(l)} \in [0;1]$ збільшує ймовірність переходу i -ї координати еталонного вектора $x_m \in X_m^o$ в одиницю, то має місце:

$$d[(x_m \oplus x_c) / \hat{\rho}_{m,i,\min}^{(l)}] \geq d[(x_m \oplus x_c) / \hat{\rho}_{m,i,\min+1}^{(l)}] \geq \dots \geq d[(x_m \oplus x_c) / \hat{\rho}_{m,i,\max}^{(l)}],$$

тобто міжцентрова відстань між двома сусідніми класами зменшується із збільшенням $\hat{\rho}_{m,i}^{(l)}$.

Таким чином, за оптимальне значення параметра навчання, виходячи із науково-методологічних засад ІЕІ-технології, слід прийняти $\rho_{m,i}^* = \hat{\rho}_{m,i,\min}^{(l)}$, $i = \overline{1, N}$, і можна стверджувати, що припущення про стаціонарність послідовності $u_N^{(l)}$ є справедливим, оскільки в силу максимально-дистанційного принципу розпізнавання образів за умови $E_{\max}^{(L+1)}(v_m, \hat{\rho}_m) = E_{\max}^{(L)}(v_m, \hat{\rho}_m)$ існує оптимальне значення $\rho_{m,i}^* = \hat{\rho}_{m,i,\min}^{(L)}$.

Реалізація алгоритму послідовної оптимізації рівня селекції координат еталонних векторів здійснювалася в процесі функціонування здатної

навчатися АСКТП виробництва складного мінерального добрива *NPК* у ВАТ «Сумхімпром». Як критерій оптимізації процесу навчання АСКТП розглядався модифікований критерій Кульбака (2.4.6). На рис. 3.9 наведено залежність усередненого для алфавіту класів розпізнавання інформаційного КФЕ (2.4.6) від рівня селекції координат еталонних векторів для 12-ї в структурі навчальної вибірки ознаки розпізнавання – щільність сірчаної кислоти. Аналіз рис. 3.9 показує, що при значеннях рівня селекції в інтервалі $[0,5;0,6]$ критерій оптимізації має менше значення ніж при інших значеннях, тобто маємо погіршення функціональної ефективності навчання системи.

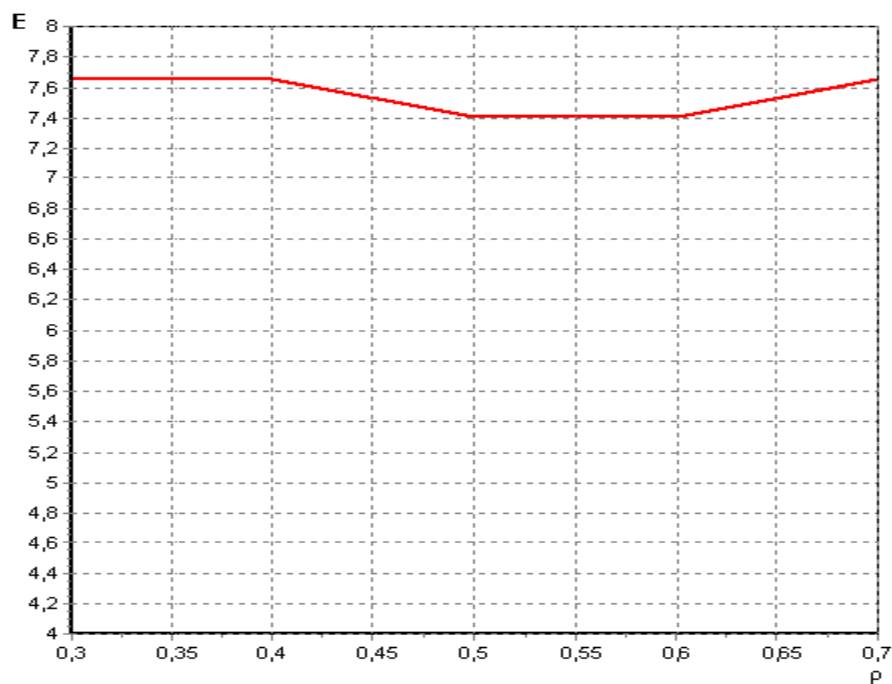


Рисунок 3.9 – Залежність усередненого інформаційного критерію Кульбака від рівня селекції координати щільності сірчаної кислоти

На рис. 3.10 наведено залежність усередненого коефіцієнту нечіткої компактності реалізацій образу (3.3.4) від рівня селекції координат еталонних векторів для 12-ї ознаки розпізнавання – щільність сірчаної кислоти.

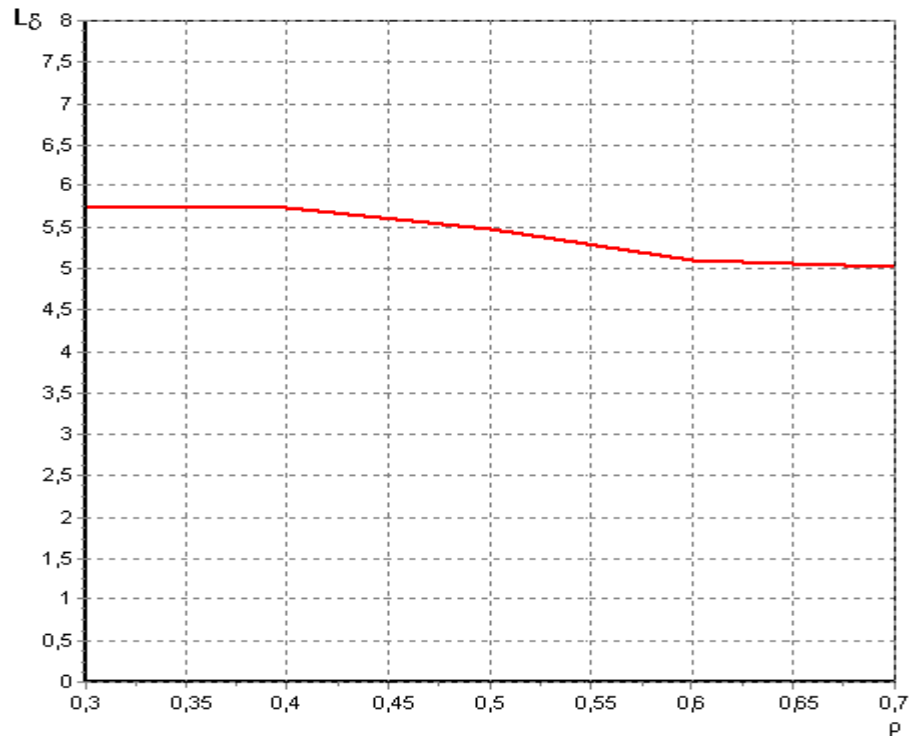


Рисунок 3.10 – Графік залежності усередненого коефіцієнта нечіткої компактності реалізацій образу від рівня селекції значень щільності сірчаної кислоти

Аналіз рис. 3.10 показує, що усереднений коефіцієнт нечіткої компактності реалізацій образу (3.3.4) має тенденцію до зниження із збільшенням ρ .

Відповідні значення оптимальних радіусів контейнерів класів розпізнавання при різних рівнях селекції значень щільності сірчаної кислоти наведено на рис. 3.11.

Аналіз рис. 3.11 показує, що критерій (2.4.6) досягає свого максимуму при значеннях рівнів селекції $\rho=0,3$, $\rho=0,6$ і $\rho=0,7$.

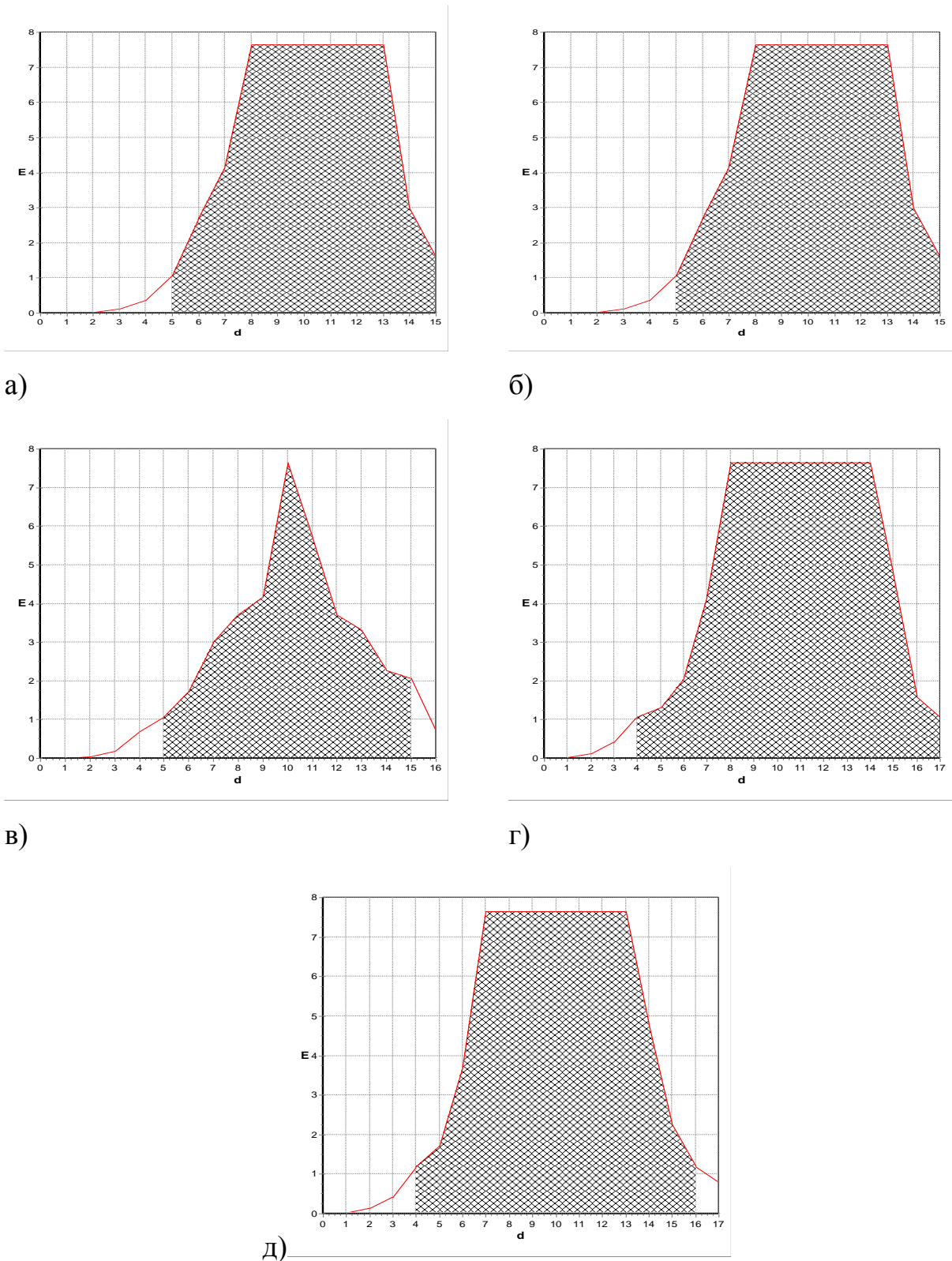


Рис. 3.11. Графік значення КФЕ (2.4.6) від радіусів контейнерів класів розпізнавання для різних рівнів селекції значень щільності сірчаної кислоти: а) $\rho=0.3$; б) $\rho=0.4$; в) $\rho=0.5$; г) $\rho=0.6$; д) $\rho=0.7$

Враховуючи те, що усереднений інформаційний КФЕ за Кульбаком на рис. 3.11 має три однакових максимальних значення при трьох рівнях селекції і те, що усереднений коефіцієнту нечіткої компактності реалізацій образу (3.12) має тенденцію до зниження із збільшенням ρ , то серед усіх значень ρ для 12-ї ознаки слід вибирати $\rho=0.7$.

3.4 Прогностичне навчання АСК виробництвом фосфорної кислоти

Створення детерміновано-статистичних методів класифікаційного прогнозування на основі самонавчання та розпізнавання образів є одним із перспективних підходів до збільшення функціональної ефективності адаптивних автоматизованих систем керування (АСК), що навчаються. Основними причинами відсутності широкого застосування класифікаційного прогнозування на практиці є:

- відомі методи мають модельний характер, спричинений нехтуванням перетину гіперсферичних контейнерів, що трапляється у практичних задачах контролю;
- відсутність алгоритмів побудови безпомилкових за багато вимірною навчальною матрицею вирішальних правил, що зумовлено методологічними причинами.

На усунення цих недоліків спрямовані методи класифікаційного прогнозування, розроблені у рамках інформаційно-екстремальної інтелектуальної технології (ІЕІТ), в основу якої покладено максимізацію кількості інформації шляхом введення додаткових інформаційних обмежень в процесі навчання АСК [31].

Концепція класифікаційного прогнозування функціонального стану АСК, що навчається, в рамках ІЕІТ така. У процесі зміни функціонального стану системи відбувається дрейф вершин еталонних векторів-реалізацій образу в просторі ознак розпізнавання, що призводить до зниження достовірності розпізнавання функціональних станів внаслідок утримування

константними в пам'яті АСК геометричних параметрів контейнерів класів розпізнавання, які сформовано на попередньому етапі навчання. Крім того, внаслідок переміщення в часі вершин векторів-реалізацій зменшується ймовірність знаходження поточних значень ознак розпізнавання у своїх контрольних допусках. При цьому також змінюються статистичні властивості реалізацій образу, що впливають на значення екстремальних порядкових статистик (ЕПС), інваріантних до широкого сімейства законів розподілу ймовірностей [68,69]. Таким чином, у процесі прогнозування система повинна слідкувати за зміною як детермінованої складової переміщення вершин еталонних векторів-реалізацій образу, так і статистичної складової, що спричиняє зміну значень ЕПС. Для підтримки належної ефективності функціонування АСК здійснюється її перенавчання. Таким чином, достатньо оцінити ступінь наближення значення поточної ЕПС до межі відповідного довірчого коридору для визначення моменту перенавчання АСК. Коли будь-яка статистика виходить за межі свого коридору, то це призводить до статистичної однаковості вибірових послідовностей, що робить необхідним перенавчання АСК.

Розглянемо інформаційно-екстремальний алгоритм прогностичного навчання інтелектуальної АСК технологічним процесом виробництва фосфорної кислоти.

Нехай дано $\{X_m^o \mid m = \overline{1, M}\}$ – алфавіт M класів розпізнавання, які у загальному випадку можуть перетинатися, і матрицю даних технологічного процесу $\|y_{m,i}^{(j)}\|$, $i = \overline{1, N}$, $j = \overline{1, n}$, де N, n – кількість ознак розпізнавання і моментів зчитування інформації відповідно. Відомо вектор-кортеж просторово-часових параметрів функціонування, що навчається, $g_m = \langle g_{m,1}, \dots, g_{m,\xi}, \dots, g_{m,\Xi} \rangle$ з відповідними обмеженнями $R_\xi(g_1, \dots, g_\Xi) \leq 0$.

Треба в процесі навчання СППР оптимізувати систему контрольних допусків D та побудувати варіаційний ряд ЕПС шляхом цілеспрямованої трансформації нечіткого розбиття $\tilde{\mathfrak{R}}^{|M|}$ в чітке розбиття еквівалентності за

умов (1.4.1-1.4.4) і (2.1.1) і при цьому оптимальні значення параметрів функціонування забезпечують максимум інформаційного критерію функціональної ефективності (КФЕ) навчання

$$E^* = \langle \max_{\{G_\xi\}} \left\{ \frac{1}{M} \sum_{m=1}^M E_m^* \right\} \rangle,$$

де E_m^* – максимум КФЕ навчання розпізнаванню реалізацій класу X_m^o ; $\{G_\xi\}$ – області допустимих значень параметрів навчання.

Таким чином, навчання інтелектуальної АСК включає визначення ЕПС в процесі оптимізації параметрів функціонування.

Математичні моделі прогностичної автоматичної класифікації в рамках ІЕІТ мають певні особливості. У цьому випадку обов'язковим етапом процесу навчання є відображення вибіркової множини \mathbf{X} на множину вільних статистик \mathbf{S} , яка обчислюється, на кожному кроці навчання:

$$\mu: \mathbf{X} \rightarrow \mathbf{S}. \quad (3.4.2)$$

Статистика $S \in \mathbf{S}$ має задовольняти таким вимогам:

- бути одномірною статистичною характеристикою вибіркової множини;
- бути інваріантною до широкого сімейства ймовірнісних мір;
- бути чутливою до зміни функціонального стану АСК, що навчається

Будемо вважати „успіхом” знаходження при випробуванні значення ознаки в своєму полі контрольних допусків. Нехай у процесі проведення випробувань виконується умова рівності ймовірностей знаходження N ознак у межах своїх контрольних допусків, тобто $p_1 = p_2 = \dots = p_N = p$. Тоді ймовірність одержання k успіхів – числа ознак, що знаходяться в допуску, визначається за біноміальним розподілом:

$$P(k, N, p) = C_N^k p^k q^{N-k},$$

де $q = 1 - p$ – ймовірність виходу значення ознаки за межі поля допусків;

$$C_N^k = \frac{N!}{k!(N-k)!} \text{ – біноміальні коефіцієнти.}$$

Введемо вільну статистику, інваріантну щодо групи всіх $N!$ перестановок координат N -вимірного вектора-реалізації $x_m^{(j)} \subset X_m^o$, $j = \overline{1, n}$, яка залежить тільки від обсягу навчальної вибірки. В процесі реалізації багатоциклічного ітераційного алгоритму оптимізації параметрів навчання за ІЕІТ формується екстремальна порядкова статистика (ЕПС), яка може розглядатися як одномірна статистична прогнозуюча функція відповідного стану АСК.

$$S_{m,n} = \sum_{j=1}^n \left(\frac{k_{m,j} - \bar{k}_{m,n}}{s_{m,n}} \right)^2, \quad m = \overline{1, M}, \quad (3.4.3)$$

де $k_{m,j}$ – число успіхів при j -му випробуванні; $\bar{k}_{m,n}$ – вибіркоче середнє значення числа успіхів після n випробувань; $s_{m,n}^2$ – вибіркоче середнє квадратичне числа успіхів після n випробувань відповідно.

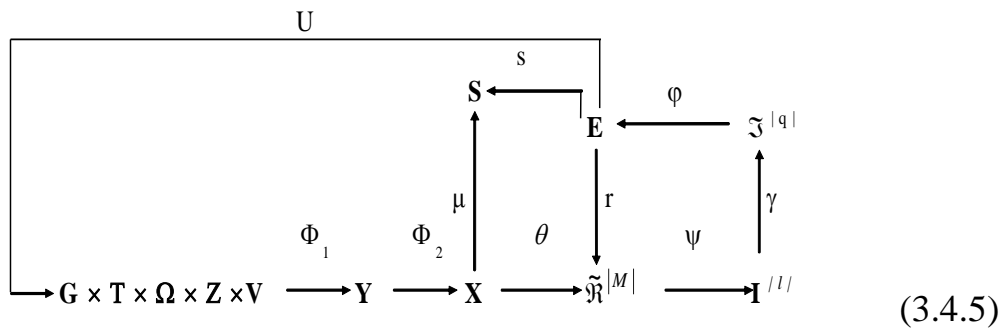
Відомо, що статистика (3) має розподіл χ^2 із ступенем свободи $k = n - 1$ і не залежить ні від математичного сподівання, ні від дисперсії, а залежить тільки від обсягу випробувань n . Як показує аналіз виразу (3), при збільшенні n вибіркоче дисперсія збігається до нуля, а функція $S_{m,n}$, маючи тенденцію до збільшення, прийме яке завгодно велике значення. Таким

чином, вільна статистика $S_{m,n}$ є членом варіаційного ряду – порядковою статистикою, ранг якої визначається номером випробування.

Позначимо через $\Pi(E_m) \subset E$ частково впорядковану підмножину значень КФЕ, обчислених у процесі навчання розпізнаванню реалізацій класу X_m^o . При цьому E_m^* – найбільший елемент підмножини $\Pi(E_m)$. Тоді множина E має впорядковану структуру $\mathbf{E} = \langle \{\Pi(E_m)\} \rangle$. Аналогічну структуру має і множина \mathbf{S} : $\mathbf{S} = \langle \{\Pi(S_m)\} \rangle$, де $\Pi(S_m)$ – підмножина статистик, обчислених при навчанні розпізнаванню реалізації класу X_m^o , найбільшим елементом якої є екстремальна статистика S_m^* . Таким чином, елементи терм-множин \mathbf{E} і \mathbf{S} знаходяться у взаємно-однозначній відповідності і визначення елементів підмножини $\Pi(S_m^*) \subset \mathbf{S}$ здійснюється в результаті бієктивного відображення:

$$s : \mathbf{E} \mapsto \mathbf{S}. \tag{3.4.4}$$

Тоді діаграма процесу базового алгоритму прогностичного навчання в рамках ІЕІ-технології, з урахуванням виразів (2) і (4), має вигляд:



Обчислення порядкових статистик $S_{m,n}$ для кожного класу здійснюється за формулою (3) на кожному кроці оптимізації параметрів навчання АСК. Статистика $S_{m,n}$ приймає екстремальне значення при випробуванні n^* , при якому КФЕ навчання системи розпізнавати реалізації

класу X_m^o набуває максимального значення в робочій області визначення його функції. Після закінчення навчання СППР формується варіаційний ряд ЕПС S_m^* за збільшенням з метою організації процедури їх співставлення з поточною статистикою S_n , що обчислюється в режимі прогностичного екзамену.

Розглянемо базовий алгоритм прогностичного навчання системи керування. Структурну схему базового алгоритму навчання LEARNING наведено на рис. 3.12. Тут показано такі вхідні дані: $\{Y[J,I,K]\}$ – масив навчальних вибірок, $J=1..NM$ – змінна кількості випробувань, де NM – мінімальний обсяг репрезентативної навчальної вибірки, $I=1..N$ – змінна кількості ознак розпізнавання, $K=1..M$ – змінна кількості класів розпізнавання; $\{NDK[I]\}$, $\{VDK[I]\}$ – масиви нижніх і верхніх контрольних допусків на ознаки відповідно. Результатом реалізації алгоритму є: $\{DOPT[K]\}$ – цілий масив оптимальних значень радіусів контейнерів класів розпізнавання у кодовій відстані Хеммінга; $\{EV[K]\}$ – масив еталонних двійкових векторів класів розпізнавання; $\{EM[K]\}$ – дійсний масив максимальних значень інформаційного КФЕ процесу навчання; $\{D1[K]\}$, $\{A[K]\}$, $\{B[K]\}$, $\{D2[K]\}$ – дійсні масиви оцінок екстремальних значень точнісних характеристик процесу навчання для відповідних класів розпізнавання: перша вірогідність, помилки першого та другого роду і друга вірогідність відповідно. Змінна D є робочою змінною кроків навчання, на яких послідовно збільшується значення радіуса контейнера класа розпізнавання. У структурній схемі алгоритму блок 3 формує масив навчальних двійкових вибірок $\{X[J,I,K]\}$ шляхом порівняння значень елементів масиву $\{Y[J,I,K]\}$ з відповідними контрольними допусками на ознаки розпізнавання за правилом

$$x_{m,i}^{(j)} = \begin{cases} 1, & \text{if } y_{m,i}^{(j)} \in \delta_{K,i}; \\ 0, & \text{if } y_{m,i}^{(j)} \notin \delta_{K,i}, \end{cases}$$

і формує масив еталонних двійкових векторів $\{EV[K]\}$ шляхом статистичного усереднення стовпців масиву $\{X[J,I,K]\}$ за правилом

$$x_{m,i} = \begin{cases} 1, & \text{if } \frac{1}{n} \sum_{j=1}^n x_{m,i}^{(j)} > \rho_m; \\ 0, & \text{if } \frac{1}{n} \sum_{j=1}^n x_{m,i}^{(j)} \leq \rho_m. \end{cases}$$

при відповідному рівні селекції, який за умовчанням дорівнює $\rho_m = 0,5$.

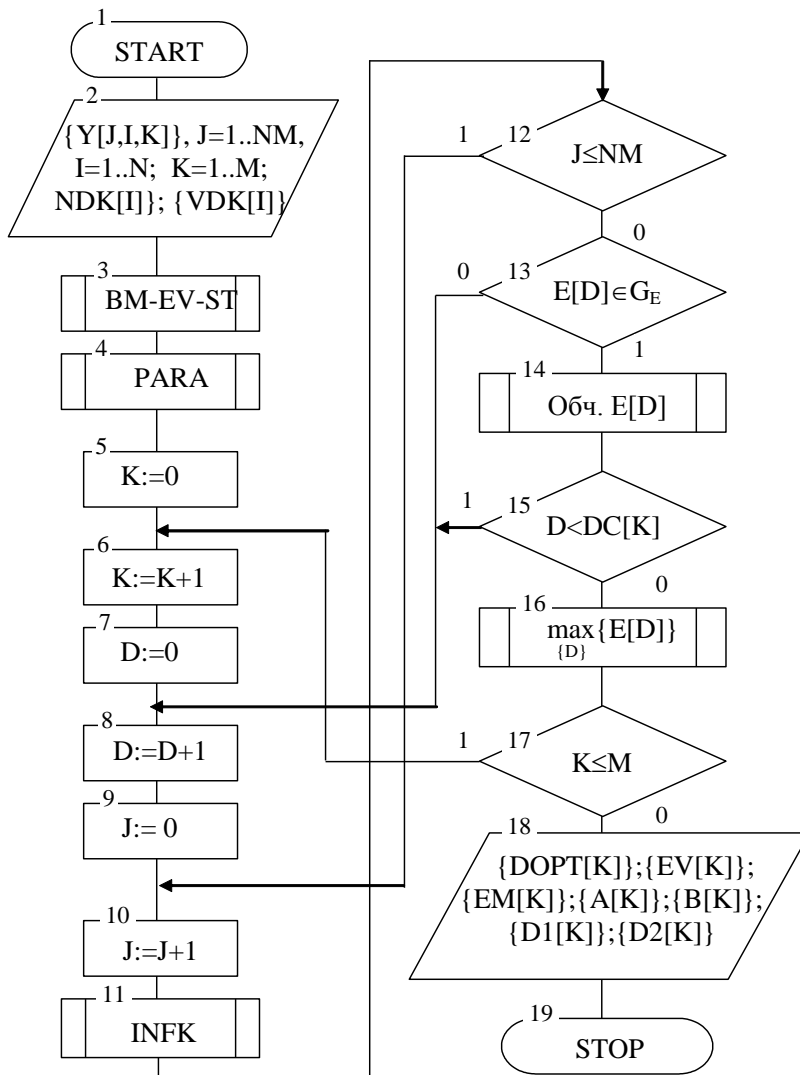


Рисунок 3.12 – Структурна схема базового алгоритму прогностичного навчання

Крім того, блок 3 на кожному кроці навчання для відповідної бінарної навчальної матриці обчислює порядкову статистику $S_{m,n}$ за формулою (3.4.3).

Блок 4 здійснює розбиття множини еталонних векторів на пари “найближчих сусідів”. Блок 11 обчислює на кожному кроці навчання значення інформаційного КФЕ і оцінки точніших характеристик процесу навчання. При невиконанні умови блоку порівняння 12 блок 13 оцінює належність поточного значення критерію $E[D]$ робочій області G_E визначення його функції і при позитивному рішенні блоку 13 це значення запам’ятовується блоком 14.

Інформаційно-екстремальний прогностичний алгоритм навчання було реалізовано для прогнозуванні функціональної ефективності автоматизованої системи керування технологічним процесом виробництва фосфорної кислоти у ВАТ „Сумхімпром”. На рис. 3. 13 показано одержані на етапі навчання порядкові статистики (3.4.3) для чотирьох класів, сформованих залежно від змісту концентрації P_2O_5 у сировині, при контрольних допусках, за яких КФЕ навчання СППР досягає глобального максимуму в робочій області визначення його функції. Як КФЕ навчання СППР розглядалася інформаційна міра Кульбака (2.).

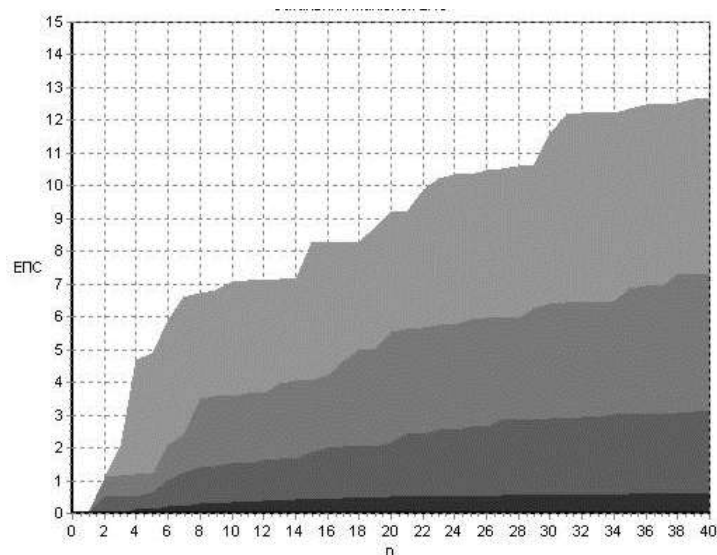


Рисунок 3.13 – Графік залежності ЕПС від кількості випробувань

Аналіз рис. 3.13 показує відповідність всім вищенаведеним вимогам до ЕПС, а також чітку залежність між трендом ЕПС і еталонним вектором відповідного класу, зокрема клас X_1^o із найбільшою ймовірністю „успіхів” має найменше значення ЕПС, для інших класів кінцеві значення варіаційних рядів зростають обернено пропорційно ймовірності „успіху”.

ВИСНОВКИ

Одержано такі наукові результати:

1. Запропоновано у рамках ІЕІ-технології новий інформаційно-екстремальний метод аналізу і синтезу СППР, що здатна навчатися в режимі факторного кластер-аналізу за умов нечіткої компактності реалізацій образу та обмеження обсягу навчальної вибірки.

2. Запропоновано комплекс логічно пов'язаних у рамках ІЕІ-технології категорійних математичних моделей, що дозволяє здійснювати аналіз і синтез алгоритмів функціонування здатних навчатися АСК в режимі ФКА.

3. Розроблено категорійну модель і алгоритм функціонування СППР, що дозволяє на практиці в режимі ФКА забезпечувати статистичну стійкість і однорідність багато вимірної навчальної вибірки.

4. Удосконалено робочі модифікації інформаційних критеріїв за Шенноном і Кульбаком для оцінки функціональної ефективності АСК, що навчаються в режимі ФКА. Показано, що у рамках ІЕІ-технології інформаційні критерії є загальними критеріями оцінки функціональної ефективності, оскільки характеризують як точнісні, так і дистанційні характеристики СППР, що відновлює в процесі навчання контейнери в радіальному базисі простору ознак розпізнавання.

5. Досліджено і розроблено алгоритм визначення мінімального обсягу репрезентативної навчальної вибірки (числа випробувань) n_{\min} , який гарантує прийнятні з практичних міркувань величину статистичної похибки і оперативність реалізації алгоритму навчання СППР.

6. Запропоновано алгоритм паралельно-послідовної оптимізації контрольних допусків на ознаки розпізнавання, що дозволяє побудувати вирішальні правила, достовірність яких наближається до асимптотичної, та розглянута реалізація алгоритму паралельнопослідовної оптимізації контрольних допусків, що дозволяє підвищити оперативність послідовного алгоритму оптимізації через попередній вибір за допомогою паралельного

алгоритму квазіоптимальної стартової системи допусків на ознаки розпізнавання.

7. Розроблено і досліджено алгоритм статистичної корекції центрів контейнерів класів розпізнавання шляхом цілеспрямованої зміни рівня селекції координат бінарних еталонних векторів класів розпізнавання в рамках інформаційно-екстремального алгоритму навчання СППР. Це дозволяє розв'язувати такі важливі для підвищення достовірності класифікації задачі:

- центрування контейнерів класів розпізнавання;
- максимізація середньої міжцентрової кодової відстані для заданого алфавіту класів розпізнавання;
- оцінка інформативності ознак розпізнавання.

8. Запропоновано як прогностичну функцію використовувати в режимі навчання СППР інваріантну до широкого класу розподілів векторів-реалізацій образів екстремальну порядкову статистику (3.4.3), яка є чутливою до зміни функціональних станів АСК технологічним процесом.

ПЕРЕЛІК ПОСИЛАНЬ

1. Довбиш А.С. Основи проектування інтелектуальних систем: Навчальний посібник.– Суми: Видавництво Сум ДУ, 2009.– 171 с.
2. Шелехов І.В., Дзюба О.О. Дослідження впливу параметрів гібридного алгоритму на функціональну ефективність навчання системи підтримки прийняття рішень // Радіоелектронні і комп'ютерні системи.–2009.–№3(37).– С. 88-91.
3. Довбиш А.С., Чекалов О.П., Мартиненко С.С. Інтелектуальна система діагностування онкопатологій // Радіоелектронні і комп'ютерні системи.– 2009.–№3(37).–С.92-96.
4. Довбиш А.С., Мартиненко С.С. Інформаційно-екстремальний метод розпізнавання електронограм // Вісник СумДУ. Серія: Технічні науки.–2009.– №2.– С. 85-91.
5. Довбиш А.С., Тронь В.А. Прогностичне навчання автоматизованої системи керування технологічним процесом виробництва фосфорної кислоти // Вісник СумДУ. Серія: Технічні науки.–2009.–№2.– С. 85-91.
6. Довбиш А.С., Жаловага В.О. Інтелектуальна система підтримки прийняття рішень для діагностування онкопатологій // Вісник СумДУ. Серія: Технічні науки.–2009.–№3.– С.
7. Довбиш А.С., Барило Р.Б., Барило О.Б. Визначення та корекція індивідуальних траєкторій навчання із застосуванням індивідуально орієнтованих електронних освітніх ресурсів // Вісник СумДУ. Серія: Технічні науки.–2009.–№3.– С.
8. Довбиш А.С., Якушев О.А. Оптимізація кроку квантування акустичних сигналів при розпізнаванні голосових команд // Вісник СумДУ. Серія: Технічні науки.–2009.–№3.– С.
9. Довбиш А.С. Стан та тенденція розвитку інтелектуальних технологій в інформаційному суспільстві // Тези доповідей Другої

міжнародної науково-технічної конференції «Інтелектуальні системи в промисловості та освіті», 3-5 грудня 2009 р.– Суми, 2009.– С. 52-56.

10. Дзюба О.О. Оцінка оперативності гібридного алгоритму навчання інтелектуальної системи підтримки прийняття рішень // Тези доповідей Другої міжнародної науково-технічної конференції «Інтелектуальні системи в промисловості та освіті», 3-5 грудня 2009 р.– Суми, 2009.– С. 23-24.

11. Котенко С.М. Навчання автоматизованої системи керування слабоформалізованим процесом без учителя // Тези доповідей Другої міжнародної науково-технічної конференції «Інтелектуальні системи в промисловості та освіті», 3-5 грудня 2009 р.– Суми, 2009.– С. 35-36.

12. Востоцький В.О. Алгоритм автоматичної класифікації // Тези доповідей Другої міжнародної науково-технічної конференції «Інтелектуальні системи в промисловості та освіті», 3-5 грудня 2009 р.– Суми, 2009.– С. 37-38.

13. Слабко М.А. Спосіб мінімізації вхідної інформації у задачах діагностики // Тези доповідей Другої міжнародної науково-технічної конференції «Інтелектуальні системи в промисловості та освіті», 3-5 грудня 2009 р.– Суми, 2009.– С. 51.

14. Шелехов І.В., Барило О.Б. Інформаційно-екстремальна корекція генотипних параметрів навчання // Тези доповідей Другої міжнародної науково-технічної конференції «Інтелектуальні системи в промисловості та освіті», 3-5 грудня 2009 р.– Суми, 2009.– С. 18-19.

15. Андрієнко Н.І., Тиркусова Н.В. Інтелектуальна система підтримки прийняття рішень для автоматизації виробництва композиційних матеріалів// Тези доповідей Другої міжнародної науково-технічної конференції «Інтелектуальні системи в промисловості та освіті», 3-5 грудня 2009 р.– Суми, 2009.– С. 30-31.

16. Мартиненко С.С., Наджафіан Мохаммед. Алгоритм формування навчальної матриці при розпізнаванні магніто кардіограм // Тези доповідей

Другої міжнародної науково-технічної конференції «Інтелектуальні системи в промисловості та освіті», 3-5 грудня 2009 р.– Суми, 2009.– С. 29.

17.Петров С.А., Руденко А.Г. Моделирование процессов, меняющихся во времени // Тези доповідей Другої міжнародної науково-технічної конференції «Інтелектуальні системи в промисловості та освіті», 3-5 грудня 2009 р.– Суми, 2009.– С. 66-67.

18.Троне В.А. Прогностична система підтримки прийняття рішень для хімічного технологічного процесу // Тези доповідей Другої міжнародної науково-технічної конференції «Інтелектуальні системи в промисловості та освіті», 3-5 грудня 2009 р.– Суми, 2009.– С.84-86.

19. Барило Р.Б. Оптимізація контрольних допусків на ознаки розпізнавання системи e-Learning при адаптивному навчанні // Тези доповідей Другої міжнародної науково-технічної конференції «Інтелектуальні системи в промисловості та освіті», 3-5 грудня 2009 р.– Суми, 2009.– С.134-135.

20.Промышленные роботы: Конструирование, управление, эксплуатация / [Костюк В.И., Гаврик А.П., Ямпольский Л.С., Карлов А.Г.]. – К.:Вища шк. Головное изд-во, 1985.– 359 с.

21. Кириченко Н.Ф., Куц Р., Лепеха Н. П. Множества принадлежности в задачах классификации сигналов // Проблемы управления и информатики. – 2001. – № 5. – С: 71–84.

22. Zadeh L. A. Fuzzi sets // Information and Control. – 1965. –№8. – Р. р. 338–353.

23. Дмитриев А. К., Мальцев П. А. Основы теории построения и контроля сложных систем.– Л.: Энергоатомиздат. Ленингр. отд-ние, 1988. –192 с.

24.Петров Е.Г., Новожилова М.В., Гребеннік І.В. Методи і засоби прийняття рішень у соціально-економічних системах: Навч. посібн./ За ред. Е.Г. Петрова.– К.: Техніка, 2004.– 256 с.

25. Ямпольский Л.С., Лавров О.А. Штучний інтелект у плануванні та управлінні виробництвом: підруч.– К. : Вища шк., 1995. – 255 с.
26. Кузьмин И. В. Теоретические основы информационной техники. – Харьков: ХВАКИУ, 1969. – 162 с.
27. Солодов А. В. Теория информации и её применение к задачам автоматического управления и контроля. – М.: Наука, 1967. – 325 с.
28. Синдеев И. М. О выборе параметров, определяющих состояние технического устройства при автоматическом контроле // Труды ВВИА им. проф. Н.Е. Жуковского, вып. 1020, 1963.– С. 11 – 20.
29. Кузьмин И. В. Оценка эффективности и оптимизация автоматизированных систем контроля и управления. – М.: Сов. радио, 1971. – 296 с.
30. Основы моделирования сложных систем / [Л.И. Дыхненко, И.В. Кузьмин, С.Г. Петров и др.]: под ред. И. В. Кузьмина.– К.: Вища школа, 1981. – 360 с.
31. Краснополюсовський А.С. Інформаційний синтез інтелектуальних систем керування: Підхід, що ґрунтується на методі функціонально-статистичних випробувань. – Суми: Видавництво Сумського державного університету, 2004. – 261с.
32. Экспертные системы. Принципы работы и примеры: пер. с англ. / [А. Брунинг, П. Джонс, Ф. Кокс и др.]: под ред. Р. Форсайта. – М.: Радио и связь, 1987. – 224 с.
33. Основы системного анализа и проектирования АСК: Учеб. пособие / [Павлов А.А., Гриша С.Н., Томашевский В.Н. и др.] – К. : Вища шк.; 1991. – 367 с.
34. Рідкокаша А.А., Голдер К.К. Основи систем штучного інтелекту: Навчальний посібник Голдер К.К. – Черкаси: Відлуння–Плюс, 2002. – 240 с.

35. Соколов А. Ю. Алгебраическое моделирование лингвистических динамических систем // Проблемы управления и информатики. – 2000. – С. 141–147.
36. Стенін О.А., Лапковський С.В., Солдатова М.О. Системно-структурні принципи комплексної технологічної підготовки виробництва при проектуванні гнучких виробничих систем // Системні дослідження та інформаційні технології. – 2004. – №1. – С. 18-31.
37. Васильев В. И. Распознающие системы. Справочник. 2-е изд., перераб. и доп. – Киев: Наукова думка, 1983. – 422 с.
38. Харкевич А. А. Теория информации. Опознание образов: Избранные труды: В 3-х т. – М.: Наука, 1973. – 3 т. – 524 с.
39. Duda R. O., Hart P. E., Stork D. G. Pattern Classification / Second ed. – John Wiley & Sons, New York, 2001. – 738 p.
40. Андерсон Т. Введение в многомерный статистический анализ: пер. с англ./ Под ред. Б. В. Гнеденко. – М.: Физматгиз, 1963. – 400 с.
41. Прикладная статистика: Классификация и снижение размерности: справ. изд./ [С.А. Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин] : под ред. С.А. Айвазяна. – М.: Финансы и статистика, 1989. – 607 с.
42. Фомин Я. А., Тарловский Г. Р. Статистическая теория распознавания образов. – М.: Радио и связь, 1986. – 264 с.
43. Jensen F. V. Bayesian networks and decision graphs. Springer-Verlag, Berlin, 2001. – 354 p.
44. Ковалевский В. А. Метод оптимизации решений в распознавании изображений. – М.: Наука, 1976. – 328 с.
45. Довбиш А.С., Козинець М.В., Котенко С.М. Оптимізація контрольних допусків на ознаки розпізнавання в інформаційно-екстремальних методах автоматичної класифікації // Вісник Сумського державного університету. Серія «Техніка», №1, 2007. – С. 169-178.

- 46.Ивахненко А. Г. О принципах построения обучающихся систем управления сложными процессами. – М.: Наука, 1970. – 252 с.
- 47.Цыпкин Я. З. Основы теории обучающихся систем. – М.: Наука, 1970. – 251 с.
- 48.Турбович И. Т., Гитис В. Г., Маслов В. К. Опознавание образов. Детерминир.-статист. подход. –М.: Наука, 1971.– 246 с.
49. Анисимов Б. В., Курганов В. Д., Злобин В. К. Распознавание и цифровая обработка изображений. – М.: Высшая школа, 1983.–256 с.
50. Соколов О. Ю. Алгебраїчні моделі та методи аналізу і синтезу систем керування слабо формалізованими процесами: автореф. дис. на здобуття наук. ступеня д-ра техн. наук: спец. 05.13.03 «Системи та процеси керування» / О.Ю. Соколов.–Харків, 2002.– 32 с.
- 51.Васильев В. И. Интеллектуальные системы управления с использованием генетических алгоритмов / Васильев В. И., Ильясов Б. Г.// Информационные технологии. – 2000.– №12.– Приложение.
- 52.Hybrid Methods in Pattern Recognition / Edited by H.Runke & A.Kandel. – Verlag Shpringer. – 2002. – 336 p.
- 53.Lihua Zhang, Wengli Xu, Cheng Chang. Genetik algorithm for affinne point pattern matching // Pattern Recognition letters. –Vol. 24.–2003.– P. 9 – 19.
54. Васильев В.И. Теория редукции в проблемах экстраполяции // Проблемы управления и информатики. – 1996.– №1,2.–С 239-251.
- 55.Li C., Goldgof D. B., Hall L. O. Automatic segmentation and tissue labeling of MR brain images // IEEE TMI.–Vol.12.–№ 4–1993.– P. 740-750.
56. Методы анализа данных: Подход, основанный на методе динамических сгущений: пер. с фр. / Кол. авт. под рук. Э. Дидэ: под ред. и с предисл. С. А. Айвазяна и В. М. Бухштабера.– М.: Финансы и статистика, 1985.–375 с.
- 57.Васильев В. И. Проблема обучения распознаванию образов – К.: Вища школа. Головное издательство, 1989 – 64 с.

58. Сироджа И. Б. Квантовые модели и методы искусственного интеллекта для принятия решений и управления. – Киев: Наукова думка, 2002.– 490 с.

59. Браверман Э. М., Мучник И. Б. Структурные методы обработки эмпирических данных. – М.: Наука, 1983. – 464 с.

60. Ивахненко А. Г. Моделирование сложных систем: (информационный подход). – К.: Вища школа. Головное изд-во, 1987.– 63 с.

61. Тутубалин В. Н. Теория вероятностей.–М.: Изд-во МГУ, 1972.–287 с

62. Колмогоров А. Н. Три подхода к определению понятия «количество информации» // Проблемы передачи информации, 1965.–Т.1. – Вып.1. – С. 3–11.

63. Шеннон К. Работы по теории информации и кибернетике. –М.: Издат. иностр. лит., 1963. – 878 с.

64. Цымбал В. П. Основы теории информации и кодирования.– Киев: Вища школа, 1977. – 288 с.

65. Кульбак С. Теория информации и статистика: Пер. с англ.– М.: Наука, 1967.– 408 с.

66. Путятин Е.П., Аверин С.И. Обработка изображений в робототехнике. – М: Машиностроение, 1990.– 320 с.

67. Силин В.Б., Заковряшин А.И. Автоматическое прогнозирование состояния аппаратуры управления и наблюдения. – М., «Энергия», 1973. – 336 с.

68. Барра Д. Математическая статистика.– М.: Мир, 1983.– 342 с.

69. Соле Ж. Л. Основные структуры математической статистики.– М.: Мир, 1972. – 127 с.