

РЕФЕРАТ

Пояснювальна записка до дипломного проекту «Система інтелектуального аналізу даних на базі нечітких темпоральних описів»: 112 с., 29 рис., 8 табл., 15 використаних джерел.

ІНФОРМАЦІЯ, ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ, ТЕМПОРАЛЬНА ЛОГІКА, ТЕХНІЧНЕ ДІАГНОСТУВАННЯ, АВТОМАТИЗАЦІЯ, КОДУВАННЯ, БАЗА ДАНИХ.

Об'єкт дослідження: система інтелектуального аналізу даних для технічного діагностування і моніторингу пристроїв автоматики та телемеханіки.

Предмет дослідження: принципи побудови, методи і алгоритми синтезу моделей залежностей між діагностичними станами пристроїв автоматики та телемеханіки.

Мета проекту: вдосконалення технічного діагностування на базі методів синтезу моделей залежностей, як основи побудови систем підтримки прийняття рішень для пристроїв автоматики та телемеханіки.

Методи дослідження: Для вирішення поставлених завдань використовувалися методи теорії інформації, теорії кодування, теорії графів, теорії ймовірності, теорії нечітких множин і елементів темпоральної логіки.

Матеріали дипломного проекту рекомендується використовувати у навчальному процесі та в практичній діяльності фахівців у сфері автоматизації та комп'ютерно-інтегрованих технологій.

ЗМІСТ

ВСТУП.....	6
..	
РОЗДІЛ 1. ТЕХНІЧНЕ ДІАГНОСТУВАННЯ ТА ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ.....	10
1.1 Стан та актуальні проблеми технічного діагностування пристроїв автоматики та телемеханіки.....	10
1.2. Сучасні системи діагностування технічних пристроїв.....	14
1.3. Системи зберігання та інтелектуального аналізу даних.....	21
1.4. Висновки до розділу 1.....	33
РОЗДІЛ 2. ТЕОРІЯ І МЕТОДОЛОГІЯ ПОБУДОВИ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ АНАЛІЗУ ДАНИХ.....	34
2.1. Представлення знань в інтелектуальних системах.....	34
2.2. Методи інтелектуального аналізу даних.....	45
2.3. Методи формування міркувань за неповним даними.....	54
2.4. Висновки до розділу 2.....	63
РОЗДІЛ 3. РОЗРОБКА МОДЕЛЕЙ ПРЕДСТАВЛЕННЯ ЗНАНЬ, МЕТОДІВ І АЛГОРИТМІВ АНАЛІЗУ ДІАГНОСТИЧНИХ ДАНИХ.....	64
3.1. Мережева модель модель представлення залежностей.....	65
3.2. Метод і алгоритми синтезу мережевих моделей залежностей.....	71
...	
3.2.1 Синтез моделі залежностей на основі критерію мінімальної довжини опису.....	72
3.2.2 Синтез моделі залежностей на основі логарифмічної метрики.....	83
.	
3.2.3 Аналіз якості роботи алгоритмів за навчальною вибіркою.....	89
3.3. Метод формування нечітко-темпоральних описів залежностей.....	90
...	
3.4. Метод нечітко-темпорального опису змін параметрів контрольованих	95

пристроїв.....	
3.5. Висновки до розділу	10
3.....	6
ВИСНОВКИ.....	10
..	7
СПИСОК БІБЛОГРАФІЧНИХ ПОСИЛАНЬ ТА ВИКОРИСТАНИХ ДЖЕРЕЛ.....	11
.	1

ВСТУП

Актуальність роботи. Стратегія розвитку транспорту передбачає першочергове вирішення ключових завдань, в числі яких забезпечення зростання ефективності, стійкості та безпеки перевезень, розвиток новітніх систем керування технологічними процесами на основі інноваційних технологій, зниження негативного впливу «людського фактора» за рахунок впровадження сучасних інформаційних та комп'ютерних технологій. У зв'язку з цим пріоритетним напрямком робіт є розробка і впровадження систем технічного діагностування та моніторингу (ТДМ) пристроїв автоматичної та телемеханіки, які дозволять мінімізувати час надходження інформації про зміни стану пристроїв і виниклих збоїв, підвищити ефективність дій обслуговуючого персоналу, оптимізувати процес виконання робіт з технічного обслуговування і ремонту пристроїв і перейти до стратегії технічного обслуговування «по фактичному стану».

В даний час широкого поширення набули такі системи ТДМ пристроїв, як АПК-ДК, АСДК і АДК-СЦБ. На основі цих систем організуються Центри технічного діагностування та моніторингу (ЦТДМ), в яких акумулюється і обробляється інформація про поточний стан пристроїв. Значний обсяг що надходить і накопиченої інформації, в поєднанні з перерахованими вище завданнями, робить актуальною розробку

інтелектуальних систем, що включають базу знань, підсистеми аналізу даних, підтримки прийняття рішень і взаємодії з персоналом ЦТДМ. Подібні інтелектуальні системи дозволяють оптимізувати процес технічного обслуговування, підвищити ефективність роботи персоналу ЦТДМ, і за рахунок цього знизити число виникаючих збоїв пристроїв.

Об'єктом дослідження є система інтелектуального аналізу даних для технічного діагностування та моніторингу пристроїв автоматики та телемеханіки.

Предмет дослідження - принципи побудови, методи та алгоритми синтезу моделей залежностей між діагностичними станами пристроїв автоматики та телемеханіки.

Мета роботи. Метою дипломного проекту є дослідження і розробка нових методів синтезу моделей залежностей, як основи побудови систем підтримки прийняття рішень для технічного діагностування та моніторингу пристроїв.

Основні задачі, які визначаються метою дослідження:

- розробка принципів побудови моделей залежностей між діагностичними станами пристроїв;
- розробка базових інтелектуальних методів синтезу моделей на основі аналізу даних діагностування пристроїв;
- розробка інтелектуальних методів формування нечітко-темпоральних описів залежностей, а також описів зміни параметрів пристроїв при зміні діагностичного стану;
- розробка структури програмної реалізації підсистем обробки даних і структури бази знань;
- розробка методики автоматизації адаптації програмного забезпечення системи інтелектуального аналізу даних.

Рішення поставлених задач дозволить підвищити ефективність застосування програмних засобів СТДМ, за рахунок автоматизації та

інтелектуалізації процесів взаємодії системи і персоналу Центрів технічного діагностування та моніторингу.

Методи дослідження. Дослідження проводилися з використанням методів теорії інформації, теорії кодування, теорії графів, теорії ймовірності, теорії нечітких множин та елементів темпоральної логіки.

Наукова новизна дипломного проекту полягає в наступному:

- запропоновано новий підхід до подання залежностей між діагностичними станами пристроїв на основі мережевої моделі з ймовірними і нечітко-темпоральними атрибутами для підвищення ефективності організації та ведення бази знань інтелектуальної системи;

- розроблено комплексний метод асоціативного аналізу даних і побудови мережевої моделі залежностей між подіями, розроблені алгоритми побудови моделі на основі принципу мінімальної довжини опису і модифікованої логарифмічною метрики моделі;

- розроблений метод нечітко-темпорального опису залежностей між діагностичними станами пристроїв і алгоритм формування НТО залежності по діагностичним даними;

- розроблений метод і алгоритми формування нечітко-темпорального опису зміни параметрів контрольованого об'єкта на основі аналізу протоколів вимірювань.

Практична цінність роботи полягає в застосуванні результатів дослідження для розробки системи інтелектуального аналізу даних (ІАД) на базі системи технічного діагностування та моніторингу (ТДМ), для Центру технічного діагностування та моніторингу пристроїв, а також для розробки систем автоматизації адаптації програмного забезпечення і автоматизації проектування діагностичних комплексів. Зокрема, в роботі досягнуті наступні практичні результати:

- розроблено серверне програмне забезпечення, що реалізує методи і алгоритми виявлення залежностей між діагностичними станами пристроїв,

функціональність виявлення фактів проведення робіт з технічного обслуговування і ремонту пристроїв і передачі інформації;

- розроблена клієнтська частина програмного забезпечення, що реалізує взаємодію системи ІАД з АРМ ШДМ (АРМ технолога ЦТДМ);

- розроблена структура бази діагностичних даних і бази знань, що містить моделі виявлених залежностей, на основі засобів СУБД MS SQL Server;

- розроблені програмні засоби автоматизації адаптації (в частині конфігурацій і складу дистрибутива) програмного забезпечення системи ІАД і сервера уніфікованого інформаційного взаємодії (СУІВ), автоматизації процесу проектування діагностичних комплексів (інформаційне забезпечення, проектна документація, схеми ліній зв'язку).

Практичне використання результатів дослідження для побудови системи інтелектуального аналізу даних технічного діагностування пристроїв АТ дозволить **автоматизувати:** виявлення залежностей між змінами діагностичних станів; контроль ефективності та своєчасності виконання робіт з технічного обслуговування і ремонту (ТОіР) пристроїв; процес формування інформаційного та адаптації програмного забезпечення системи ІАД і СУІВ.

Достовірність і обґрунтованість. Достовірність і обґрунтованість наукових положень, висновків і результатів, сформульованих в дипломному проекті, підтверджується результатами обчислювальних експериментів на практичних і модельних задачах.

РОЗДІЛ 1

ТЕХНІЧНЕ ДІАГНОСТУВАННЯ ТА ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ

Сучасні системи діагностування та моніторингу не тільки повинні забезпечувати високий рівень автоматизації діяльності експлуатаційного персоналу, а й володіти високою гнучкістю, адаптивністю до мінливих умов експлуатації, здатністю до обробки складно формалізованої інформації та підтримки користувачів при аналізі даних і прийнятті рішень. Ці якості можуть бути реалізовані тільки за рахунок застосування особливих методологій побудови програмних систем і використання інтелектуальних методів аналізу даних, виділення знань з даних і прийнятті рішень в умовах невизначеності.

Результати, досягнуті в області інтелектуальних експертних систем і систем підтримки прийняття рішень, широко використовуються в різних галузях промисловості і економіки, в енергетиці, соціології, медицині і т.д. Однак, як показує проведений аналіз, в області діагностування і моніторингу пристроїв автоматики і телемеханіки інтелектуальні методи аналізу даних поки практично не застосовуються.

Завдання створення програмних засобів синтезу параметрів систем автоматичного

1.1. Стан та актуальні проблеми технічного діагностування пристроїв автоматики та телемеханіки.

Як відомо, основне завдання системи діагностування пристроїв автоматики та телемеханіки - формування точного діагнозу стану функціональних пристроїв.

Кафедра АЕМ				НАУ 21 03 86 000 ПЗ			
Виконав	Кузюра А.А.			ТЕХНІЧНЕ ДІАГНОСТУВАННЯ ТА ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ДАНИХ	Літера	Аркуш	Аркушів
Керівник	Товкач С.С.					1	24
Консульт.							
Н-контр.	Товкач С.С.				АТ-412Б	15	19
Зав. каф.	Захарченко В.П.						

Вирішенню цього завдання перешкоджає ряд проблем, наприклад, складна електромагнітна обстановка, що впливає на точність виконуваних вимірювань, труднощі з формалізацією опису станів пристроїв, а також зазначене в [1] властивість раптовості виникнення значної частини відмов пристроїв сигналізації, централізації та блокування (СЦБ).

Характерними прикладами складно формалізованих станів пристроїв є множинне функціонування приводів, розрив або пошкодження вузлів, порушення показань дешифраторної комірки.



Рис. 1.1. Приклад графіку споживання струму при множинному функціонуванні приводів

Аналіз ситуації множинного функціонування приводів потребує діагностування електродвигунів (ЕД). Складність викликана тим, що напруга і струм живлення ЕД знімаються з одного або двох шунтів, а значить, для формування діагнозу однієї окремо взятої ділянки доводиться аналізувати сумарний струм, що споживаються всіма ЕД (див. Рис. 1.1) [1,2].

Розрив або пошкодження вузлів кола діагностується за результатами аналізу змін рівня і фази напруги на вході приймача []. В даних змінах проявляється також вплив на кола різних перешкод, що викликають небезпечні електрорушійні сили або струми внаслідок гальванічного зв'язку або електричного і магнітного впливу. Крім того, своєчасний аналіз напруг

кіл може попередити пробої ізолюючих стиків, які призводять до таких відмов як помилкова вільність і помилкова зайнятість кіл.

Порушення роботи і відхилення показань дешифраторної комірки діагностується за наявністю перешкод, накладених на кодові посилки кодovаних ланцюгів. Перешкоди найчастіше пов'язані з обривами міжстикових з'єднань, намагніченістю і елементів, впливом зворотних тягових струмів на коротких секціях, впливом розташованих в безпосередній близькості високовольтних кабелів.

Крім індивідуального діагностування пристроїв є завдання, що вимагають аналізу стану групи пристроїв. Приклад такого завдання - виявлення проведення робіт з технічного обслуговування і ремонту (ТО і Р) пристроїв АТ []. Одним із способів її рішення є аналіз послідовності змін діагностичного стану пристроїв на предмет виявлення патернів (однотипних послідовностей ситуацій на різних пристроях), які відповідають певним видам робіт з ТО і Р [1,2].

Особливо слід підкреслити, що цінність діагностики полягає, перш за все в упереджувальному виявленні станів, що передують відмові пристрою (передвідмов), в своєчасному передбаченні виникнення відмови. Після того, як сталася серйозна відмова, її виявлення корисне тільки з точки зору протоколювання і якнайшвидшого усунення. Для виключення можливості несення матеріальних і часових втрат внаслідок виникнення відмови слід будувати систему діагностування на принципі раннього, попереджувального виявлення ознак несправності пристрою. Отже, стає актуальною задача застосування в системі діагностування сучасних методів інтелектуального аналізу даних, наділення її функціями експертної системи.

Крім вимог до функціональних можливостей, до сучасних систем моніторингу пред'являються підвищені вимоги до захищеності від зовнішніх впливів (від механічної захищеності і здатності до роботи в екстремальних умовах апаратного забезпечення, до інформаційної захищеності і стійкості до

спроб отримання несанкціонованого доступу програмного забезпечення), до стабільності роботи системи в цілому при відмові одного з її модулів.

Системи моніторингу та діагностики пристроїв автоматики та телемеханіки є розподіленими і повинні забезпечувати персонал постів і центрів діагностування повною і достовірною інформацією в режимі реального часу.

Ще однією критично важливою вимогою до системи діагностування, зазначеною в [2,3] є можливість всебічної інформаційної взаємодії паралельно функціонуючих і взаємодоповнюючих систем між собою. Особливо актуальною на сьогоднішній день є можливість взаємодії систем діагностування пристроїв АТ з наступними системами:

- системою контролю та планування технологічного обслуговування; підсистеми обліку відмов і виявлення усунення відхилення від норм утримання.

- системами диспетчерської централізації;
- мікропроцесорними системами керування станціями (ділянками);
- системою керування сортувальними станціями.

Взаємодія з цими системами вирішує такі основні завдання:

- автоматизація контролю і діагностування стану пристроїв АТ;
- централізація даних для організації технічного обслуговування засобів АТ в дистанції сигналізації та зв'язку;
- оптимізація програмно-апаратних засобів систем в частині підключення до дискретних і аналогових сигналів пристроїв;
- розвиток інформаційного середовища та інфраструктури диспетчерського контролю пристроїв і систем АТ (віддалений моніторинг);
- розвиток технічного і сервісного обслуговування на основі віддаленого моніторингу результатів діагностування в дистанціях і дорожніх центрах;

- впровадження малолюдних і ресурсозберігаючих технологій, зменшення експлуатаційних витрат на технічне обслуговування, заснованих на автоматизації технічного діагностування та моніторингу засобів АТ.

Широкий набір вимог безумовно веде до ускладнення архітектури і функціонування системи діагностування та моніторингу. При цьому необхідно забезпечити наявність інтуїтивно зрозумілого інтерфейсу користувача автоматизованого робочого місця (ПО АРМ системи діагностування), а також реалізувати функції помічника-експерта, зводячи до мінімуму дії оператора, і надаючи йому підтримку в прийнятті рішень з питань запобігання та усунення відмов пристроїв [3].

Таким чином, в області діагностування пристроїв АТ є потреба в застосуванні моделей і методів, що володіють низьким ступенем помилкових спрацьовувань і допускають деякі некритичні відхилення сигналу від заданих норм. Критичними для системи діагностування є такі параметри, як наявність розподіленої структури, властивість розширюваності і функціональність всебічної інформаційного взаємодії з суміжними системами.

1.2. Сучасні системи діагностування технічних пристроїв.

На сьогоднішній день лідерами серед систем діагностування та віддаленого моніторингу є три системи: система диспетчерського контролю та діагностики пристроїв АТ АПК-ДК, автоматизована система диспетчерського контролю АСДК і система автоматизації контролю та діагностування пристроїв СЦБ АДК-СЦБ. Також, широке застосування має система технічної діагностики МПК в складі ЕЦ-МПК.

Система АПК-ДК [3] призначена для централізованого контролю і діагностування технічного стану пристроїв АТ, автоматизації технічного обслуговування і процесу усунення відмов, а також для організації керування рухом. До складу АПК-ДК входять такі апаратні і технологічні комплекси:

- апаратний комплекс контролю та діагностики об'єктів АТ;

- технологічний комплекс контролю, діагностики та усунення відмов;
- технологічний комплекс центру керування перевезеннями;
- підсистема обміну інформацією.

Автоматизовані робочі місця системи АПК-ДК забезпечують:

1) моніторинг місця знаходження:

- контроль місця знаходження і стану пристроїв;
- моніторинг відмов і передвідмовних станів;
- аналіз станів і місця знаходження за останній рік;
- контроль за рухом на основі моделі руху з виявленням логічних відмов;

2) моніторинг діагностичних станів:

- контроль виявлених ситуацій: з відмовами, передвідмовними станами, технологічними ситуаціями, проведенням ТО, діагностикою АПК-ДК;

- аналіз всіх порушень за кількісним критерієм у фільтрах і угрупованнях;

- автоматизована фіксація порушень;

3) моніторинг вимірюваних параметрів:

- контроль параметрів пристроїв: напруг, струмів, опорів і часових значень;

- аналіз параметрів за кількісними та якісними критеріями на часових графіках за останній рік;

- виявлення критичних параметрів і сигналізація про них;

- прогноз часу виходу параметрів за норму і інші статистичні показники.

Відмінною особливістю системи АПК-ДК є широка взаємодія з системою контролю та планування технологічного обслуговування [3]. Застосування в дистанції цих двох систем одночасно дозволяє автоматизувати облік відмов пристроїв і коригувати план техобслуговування за фактичним станом пристроїв.

Система АСДК [3] призначена для контролю стану пристроїв АТ, виявлення несправностей і передвідмовних станів, ведення протоколів подій з можливістю динамічного відновлення ситуації, ідентифікації рухомих одиниць, ведення моделі на ділянці контролю, формування подій для автоматизованого ведення графіка виконаного руху.

АСДК складається з двох підсистем: верхнього і нижнього рівнів. Нижній - це електричні датчики стану технічних засобів: ламп і реле, вимірювальних панелей, пристроїв збору цифрової та аналогової інформації, її первинної обробки та передачі в мережу системи. В якості пристроїв збору інформації використовуються КДК, апаратура ДК-М [3,4]. На верхньому рівні здійснюється прийом і маршрутизація потоків інформації, її обробка та відображення на автоматизованих робочих місцях - абонентах мережі АСДК, а також зв'язок із зовнішніми інформаційно-обчислювальними системами.

В даний час фахівцями ведуться розробки нових алгоритмів підсистеми логічного виявлення залежностей електричної централізації і автоблокування [3,4].

Система СТД МПК призначена для:

- фіксування, зберігання, відображення в реальному часі і в режимі «чорного ящика» стану дискретних сигналів пристроїв ЕЦ;
- вимірювання часових характеристик роботи САТ;
- вимірювання аналогових характеристик сигнального струму фазочутливого і тональних кіл, що живлять установки і інші пристрої ЕЦ;
- вимірювання опору ізоляції кабельних мереж щодо землі без відключення монтажу;
- реалізація діагностичних алгоритмів для визначення передвідмовного стану САТ, пошуку місця і визначення причин відмов;
- експертна оцінка роботи пристроїв електричної централізації.

Система АДК-СЦБ виконує наступні основні функції:

- надання обслуговуючому персоналу діагностичної інформації для автоматизації аналізу та пошуку причин взаємовпливу при роботі пристроїв у

випадках складних несправностей за динамічними протоколами комплексної діагностики;

- автоматизація технічного обслуговування пристроїв СЦБ;
- формування і зберігання архівів відмов, статистичної діагностики, динамічних протоколів та інших спеціалізованих баз даних;
- отримання діагностичної інформації від існуючих, а також модернізованих і вдосконалюваних систем типу АСДК, АБТ за рахунок інтеграції інформаційних потоків;
- інтеграція з системами АТ і формування контрольних-діагностичних комплексів (КДК), що вирішують завдання комплексного підходу за контролем та обслуговуванням систем і пристроїв;
- автоматизація контролю роботи з пристроями, що впливають на безпеку руху (при техобслуговуванні і ремонтно-відновлювальних роботах);
- забезпечення повної і достовірної інформації оперативного та обслуговуючого персоналу систем АТ на об'єктах для своєчасного прийняття рішень при обслуговуванні та керуванні технологічним процесом, в тому числі і щодо попередження відмов за даними передвідмовної діагностики.

Таким чином, розглянуті системи діагностування пристроїв АТ дуже схожі між собою, хоча у кожній з них був свій шлях розвитку. Це обумовлено наявністю в експлуатаційно-технічних вимогах до систем технічної діагностики і моніторингу пристроїв АТ [5] визначень всіх основних функцій. Вимоги пред'являються як до архітектури системи і часу доставки інформації на центральний пост діагностування та моніторингу, так і до складу завдань, які виконуються системою на рівні лінійного пункту. Цей документ був основним при виконанні даного дослідження.

Проте, між розглянутими системами є і відмінності. В першу чергу вони пов'язані з пріоритетами завдань, поставлених перед кожною з систем. АПК-ДК спочатку планувалася як підсистема системи контролю та планування технічного обслуговування (СКтаПТО), і перший час вони розвивалися паралельно. Згодом вони розвивалися окремо, і тільки в даний

час планована взаємодія стала здійснюватися в повній мірі. Широка взаємодія СКтаПТО і АПК-ДК, очевидно, є основним плюсом цієї системи []. Система АСДК орієнтована на диспетчерський контроль, і тому основними завданнями є ув'язки з різними системами диспетчерської централізації [3,4], системою ведення графіка виконаного руху об'єктів. Система АДК-СЦБ в першу чергу орієнтована на діагностування і віддалений моніторинг пристроїв АТ. АДК-СЦБ має найширший перелік вимірюваних параметрів, а значить, має найбільш повну вихідну інформацію для виявлення відмов і, що особливо актуально, передвідмовних станів пристроїв СЦБ.

На додаток до розгляду систем діагностування, проведемо аналіз найбільш представницьких систем в інших областях. Всі розглянуті нижче системи мають властивості сучасних систем моніторингу, а саме:

- діагностування процесів в реальному часі;
- підтримка сучасних методів інтелектуального аналізу даних;
- експертна оцінка ситуації.

Широко відомі системи, розроблені НВЦ «Інтелеком». Вони застосовуються як на підприємствах космічної галузі, так і в медицині, екології та транспорті. У системах НВЦ «Інтелеком» традиційні методи і засоби обробки інформації доповнені механізмом використання баз знань і нейромережових технологій. Це дозволило застосовувати при аналізі важко формалізовані поняття, такі як експертні оцінки, отримувати нові знання, виявляти приховані тенденції та закономірності, а також швидко і з високою ймовірністю прогнозувати поведінку об'єктів моніторингу і вирішувати завдання класифікації.

Існує ряд систем, розроблених для вирішення завдань вібродіагностики: ВАСТ, КОМПАКС, VDR і ін. Ці системи дозволяють проводити всебічний аналіз аналогових і дискретних сигналів і виконують функції статистичного аналізу, побудови залежностей, обчислення спектральних характеристик, візуалізації.

Подібні системи мають підтримку способів і алгоритмів обробки сигналів, що містять детерміновані і випадкові складові, постановки діагнозу і формування прогнозу стану вузлів роторних машин, містять експертні підсистеми визначення несправностей, попередження аварій і видачі рекомендацій обслуговуючому персоналу, вбудовані підсистеми самодіагностики технічного стану апаратури.

Розвиваючим напрямком є розробка механізмів подання та діагностування складних технологічних процесів. Лідуючою системою в цій галузі є система «Когра» [3], що застосовується в галузях енергетики, хімічної і нафтохімічної промисловості, а також в системах транспорту нафти і газу. Система «Когра» використовується для представлення стану складних технологічних об'єктів, керованих оперативно, у вигляді т.зв. когнітивних образів. Когнітивний образ - це графічна динамічна модель поведінки складного керованого технологічного об'єкта. Головне призначення когнітивних динамічних зображень полягає в можливості інтегрування великої кількості регульованих змінних в один образ з однозначною інтерпретацією цього способу оператором.

На підприємствах газової промисловості успішно застосовується система автоматизації компресорних цехів (СА КЦ) [3,4], адаптована компанією «ГазЕксперт». СА КЦ є комплексом програмно-технічних засобів, що використовує сучасні методи математичного моделювання та оптимізації з метою підвищення ефективності неперервних процесів, а також збільшення продуктивності і терміну служби встановленого обладнання. СА КЦ виконує функції діагностування, аналізуючи змінні як функції часу, для контролю можливої деградації обладнання. Отримана інформація може бути використана при прийнятті рішень про технічне обслуговування устаткування. Залежно від прогнозованого режиму роботи система передбачає досягнення встановлених нормативних показників, на основі чого можливе планування технічного обслуговування.

Досить широке поширення за кордоном отримала універсальна система MoniTeqIII [5]. MoniTeq являє собою систему, засновану на знаннях і забезпечує детальний аналіз отриманих даних, який дозволяє отримати природу відхилення, його серйозність (якісну або кількісну оцінку), можливі причини, наслідки при бездіяльності, а також рекомендовані дії для зміни ситуації.

Однією з найбільш високонадійних зарубіжних систем, що застосовуються на атомних електростанціях, є система MSET (The Multivariate State Estimation Technique) [5].

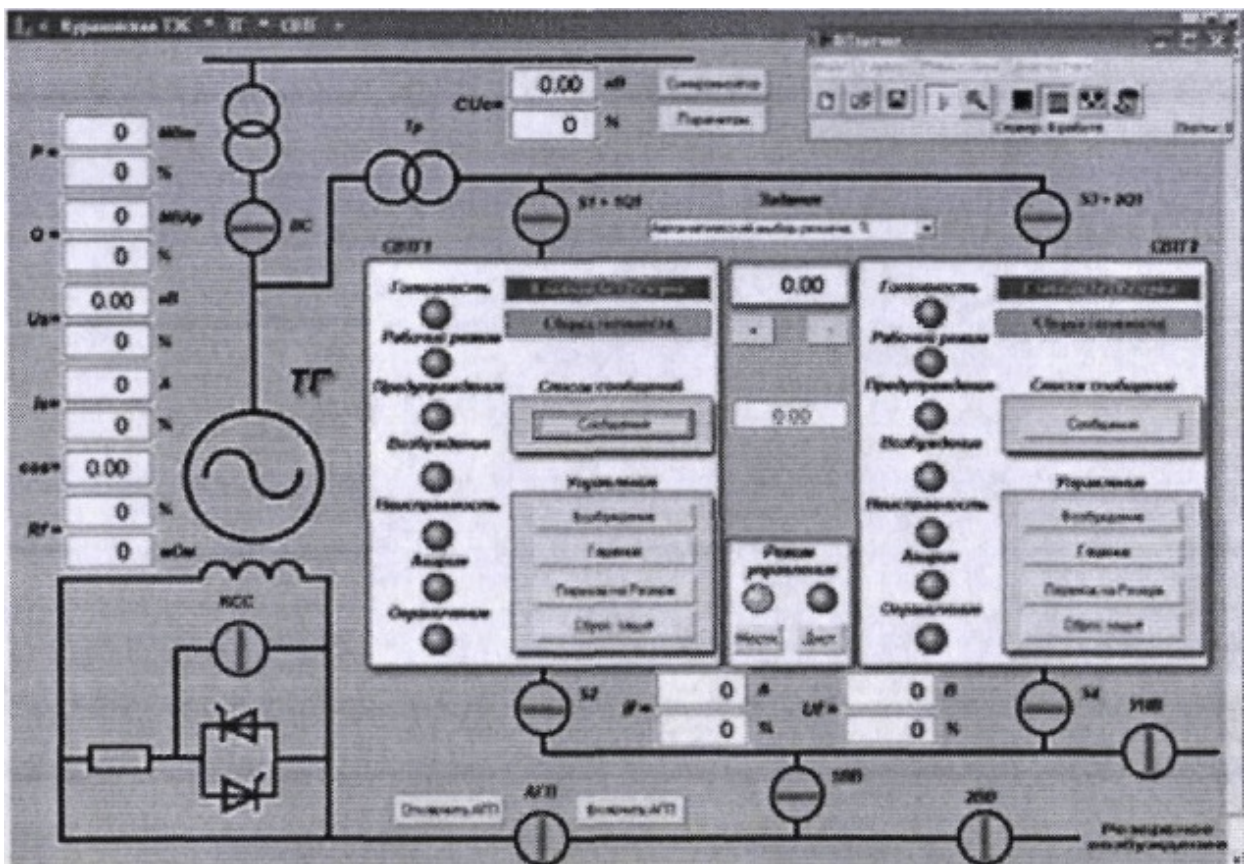


Рис. 1.2. Елемент інтерфейсу АРМ системи моніторингу обладнання MSET

Це система призначена для моніторингу пристроїв в режимі реального часу. Вона постачає операторів як своєчасною та надійною інформацією про стан процесу, отриманої зі стану сенсорів, так і його очікуваної поведінки, заснованої на минулих спостереженнях. Система використовує високоефективні методи: формування аналітичної оцінки сигналів на основі їх фактичного стану і попередньо збережених залежностей між ними і аналіз статистичних характеристик часових рядів, отриманих різницею між кожним

вимірним сигналом і його чисельно розрахованим значенням для найшвидшого визначення часу, коли процес поведе себе аномально.

Також можна відзначити базову інструментальну систему G2 фірми Gensym для підтримки проектування інтелектуальних експертних систем реального часу. Платформа G2 використовує технології логічного виведення в реальному часі, включаючи правила, процедури, об'єктну модель, імітаційне моделювання та графічне представлення в єдиному середовищі розробки.

Відмінними рисами системи G2 є стандартизованість, простота інтеграції і вбудовування в інші системи завдяки підтримці технологій ActiveX/COM, .NET, Java RMI, SQL, ODBC, JMS, OPC, XML, HTTP. Система має інтуїтивним призначенням для користувача інтерфейсом і забезпечує простоту прототипування, розробки та впровадження проектів.

Крім того, система G2 відрізняється високою стабільністю і надійністю, застосовується в таких областях як нафтохімія, енергетика, авіація, транспорт і телекомунікації.

Таким чином, у порівнянні з системами діагностування перерахованих вище галузей, що забезпечують формування експертних оцінок стану технічних пристроїв, виявлення прихованих закономірностей і тенденцій за допомогою методів інтелектуального аналізу даних, реалізацію функцій прогнозування стану пристроїв, системи діагностування пристроїв АТ поступаються за рівнем інтелектуалізації та реалізації експертних функцій, забезпечуючи лише класифікацію діагностичних станів, розрахунок і візуалізацію статистичних показників.

1.3. Системи зберігання та інтелектуального аналізу даних.

Функціональність і ефективність сучасної інформаційної системи в значній мірі залежить від повноти, релевантності та актуальності бази знань, використовуваної для формування експертних оцінок, вироблення рішень і

рекомендацій. Перераховані ключові характеристики баз знань в свою чергу в значній мірі залежать від якості джерел і технологій отримання знань. Питання розробки технологій вилучення знань для практичного використання теоретичних досягнень в галузі інтелектуальних моделей розглядалось в ряді джерел, присвячених проблемам штучного інтелекту [5].

Дослідження в області розробки моделей, заснованих на знаннях, починалися в 1960-1970гг. [8,10]. Протягом тривалого періоду часу в якості основного джерела знань розглядалися експерти предметної області, і тим не менш, спроби створення баз знань для реальних додатків виключно на основі узагальнення їх досвіду зазнали невдачі.

До початку ХХІ століття погляди на пріоритети в області отримання знань змінилися, і в якості інформаційної основи в багатьох додатках стали використовуватися бази і сховища даних, а задача видобування знань перейшла в центр уваги дослідників і розробників промислових систем обробки даних.

Увага сучасної інформаційної індустрії до завдань вилучення знань викликано збільшенням темпів розвитку інформаційних технологій та накопиченням все більших обсягів даних, значна частина яких зберігаються без активного практичного застосування. Яскравим прикладом є великі масиви даних, що містять цінну інформацію про фінансові операції, характеристиках продукції, що виробляється, результати медичної і технічної діагностики, дослідженнях земної поверхні (NASA) і т.д. [9].

Для отримання безпосередньої практичної або комерційної віддачі від даних, потрібна їх серйозна обробка. Найчастіше однорідні за змістом відомості виходять з різних джерел і зберігаються в різних БД, в різних форматах і виміряні з різною точністю. З іншої сторони, користувачі прикладних систем пред'являють все більш серйозні вимоги до результатів дослідження даних. Кінцевим користувачам потрібні висновки, які допомагають у вивченні комплексних питань і проблем, у прийнятті важливих, відповідальних рішень.

Незважаючи на досить довгу історію теоретичних і прикладних досліджень в області видобування корисної інформації з баз даних, саме усвідомлення потенційної комерційної цінності накопичених даних привело до зростання числа програмних розробок на початку 1990-х років, які настільки ж активно продовжують з'являтися і зараз. У всьому світі користуються комерційним успіхом розроблені методи і проблемно-орієнтовані стандартні програмні засоби обробки даних у фінансовій сфері, в управлінні виробництвом і бізнес-процесами, в соціальній сфері, а також в наукових дослідженнях.

З моменту початку свого розвитку інформаційні технології в сфері зберігання та обробки даних значно змінилися, пройшовши шлях від файлових сховищ до сучасних інтегрованих систем управління базами даних, що володіють гнучкістю налаштувань, масштабованістю і відмовостійкістю, підсистемами авторизації доступу та контролю цілісності даних.

Файлові системи зберігання даних, поширені в 60-х роках, мали масу недоліків, таких як високий мережевий трафік, мізерний набір операцій маніпулювання даними (файлами), відсутність підсистеми забезпечення безпеки доступу. За допомогою файлових систем можна було створити лише слабо-структуровані сховища, як правило, малого обсягу, не призначені для одночасної роботи великої кількості користувачів.

Необхідність підтримки складних структур даних в інформаційних системах, привела до розробки індивідуальних засобів управління даними, які практично повторювалися від однієї системи до іншої як в частині принципів взаємодії програмних компонентів, так і в частині схожості реалізованих алгоритмів. Прагнення виділити загальну частину, відповідальну за управління складноструктурованих даними, а також «відокремити» самі дані від програм для забезпечення можливості їх використання в різних додатках стало причиною створення систем управління базами даних (СУБД).

Крім підтримки складних структур даних, СУБД були наділені також безліччю інших корисних функцій. Вони в змозі самостійно здійснювати контроль цілісності даних, що зберігаються, забезпечувати багатокористувацький режим - паралельність виконання запитів від декількох користувачів, здійснювати модифікацію і швидкий пошук і т.д. Серед структур організації зберігання даних, з якими працювали перші СУБД, слід зазначити ієрархічні і мережні.

Ієрархічні БД складаються з упорядкованого набору кількох примірників одного типу дерева, який включає один «кореневий» тип запису і впорядковану структуру з деякого числа типів піддерев (кожне з них є певним типом дерева). Тип дерева являє собою ієрархічно організований набір типів записів, наприклад, на рис. 1.3 показана структура ієрархічної БД електронних компонентів [6].

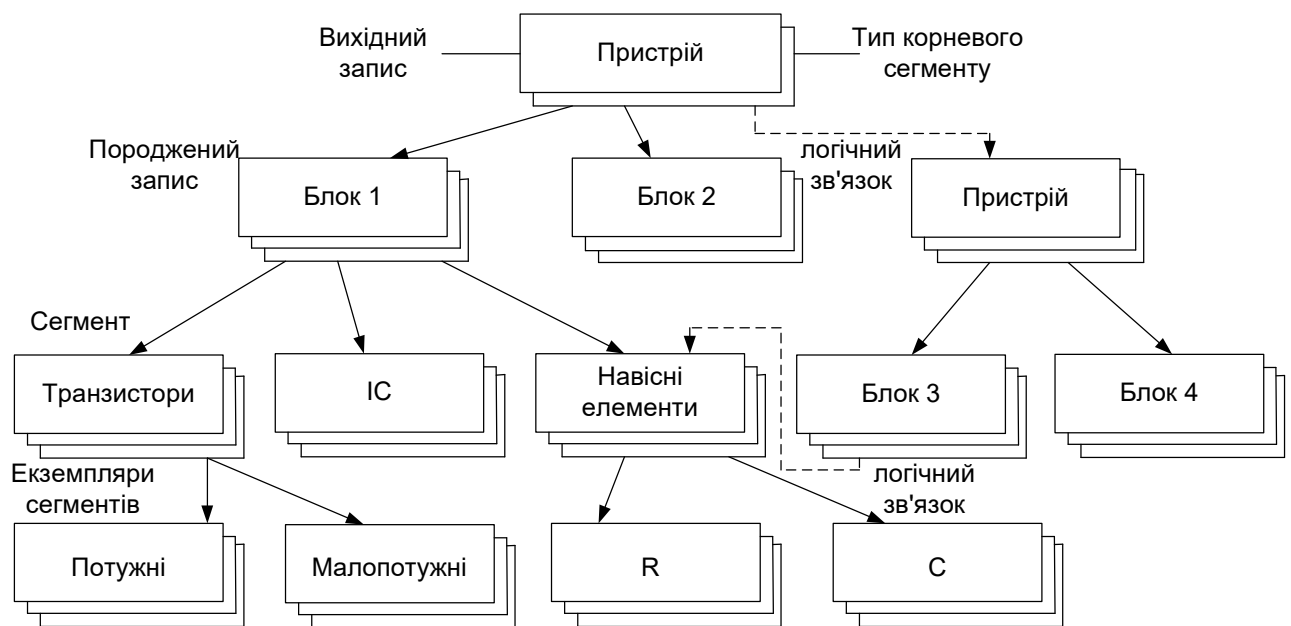


Рис. 1.3. Приклад ієрархічної організації даних

Основними інформаційними одиницями в ієрархічній моделі даних є сегмент і поле. Поле даних визначається як найменша неподільна одиниця даних, доступна користувачу. Для сегмента визначені такі параметри, як тип сегмента - іменованій перелік типів полів даних - і примірник сегмента. Примірник сегмента утворений набором конкретних значень полів даних.

Мережевий підхід до організації даних є узагальнюючим розширенням ієрархічного підходу. В ієрархічних структурах запис-нащадок повинен мати тільки одного предка; в мережевий же структурі у нащадка може бути довільне число предків. Мережева БД складається з наборів записів і зв'язків між ними, тобто з наборів примірників типів записів і типів зв'язків, заданих в схемі БД.

На рис. 1.4 показаний приклад структури мережевої БД системи обробки замовлень [6].

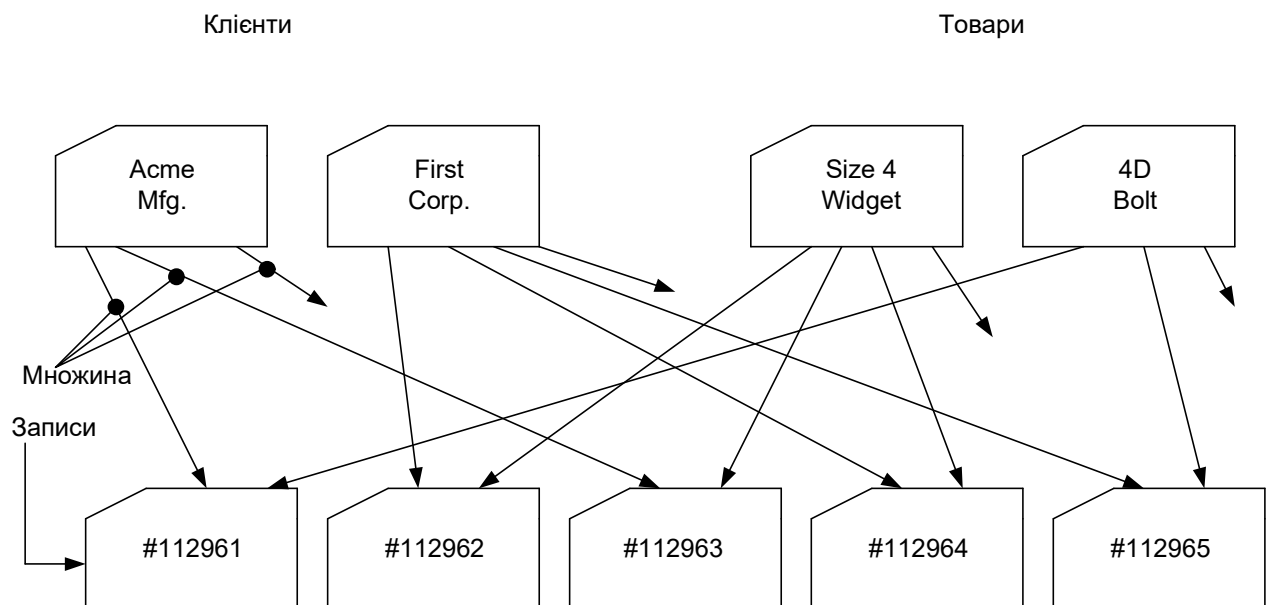


Рис. 1.4. Приклад мережевої організації даних

Незважаючи на те, що мережевий підхід вирішує деякі проблеми, пов'язані з ієрархічною моделлю, виконання простих запитів залишається досить складним процесом. Також, оскільки алгоритми процедур вибірки даних залежать від фізичної організації цих даних, мережева модель не є повністю незалежною від програми. Іншими словами, якщо необхідно змінити структуру даних, то доведеться міняти і додаток.

У більшості сучасних СУБД використовується реляційна модель даних у вигляді сукупності таблиць, пов'язаних між собою за ключовими полями (див. рис. 1.5). Маніпулювання даними базується на правилах реляційної

- незалежність контролю цілісності;
- дистрибутивна незалежність;
- узгодження мовних рівнів.

Подальша еволюція технології обробки даних принесла ряд різних засобів оптимізації доступу.

З метою підвищення продуктивності виконання запитів стали застосовуватися індекси - об'єкти БД, які формуються із значень одного або декількох стовпців таблиці і покажчиків на відповідні рядки таблиці, і таким чином, дозволяють знаходити потрібний рядок по заданому значенню. Прискорення роботи з використанням індексів досягається в першу чергу за рахунок того, що їх структура, як наприклад, збалансоване дерево (В-дерево, див. рис. 1.6) оптимізована для пошуку даних.

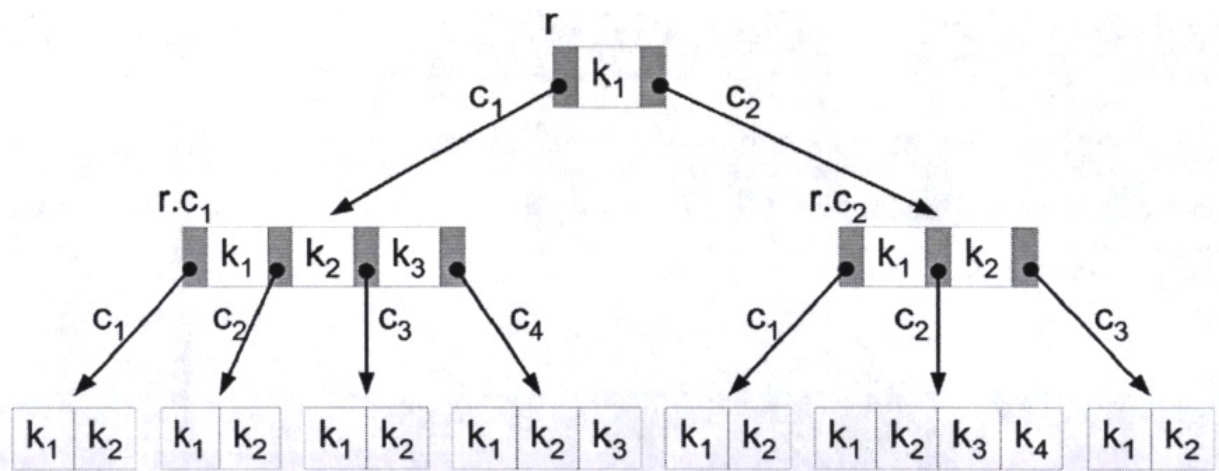


Рис. 1.6. Приклад В-дерева

Контроль цілісності даних забезпечується використанням механізму транзакцій (груп послідовних, логічно пов'язаних операцій роботи з даними). Був розроблена і стандартизована мова структурованих запитів SQL (structured query language), яка стала для розробників настільних додатків універсальним засобом доступу до даних і керування структурою БД.

Інформаційні системи, побудовані з залученням описаних механізмів і орієнтовані головним чином на введення і зберігання інформації, отримали

назву OLTP-систем (On-Line Transaction Processing, оперативна транзакційна обробка даних).

Функціональним розширенням реляційних БД стали об'єктно-орієнтовані БД, а також багатовимірні сховища і вітрини даних.

Об'єктно-орієнтовані БД (ООБД) дозволяють застосувати принципи об'єктно-орієнтованого програмування (ООП) на рівні зберігання даних. У ООБД дані оформлені у вигляді моделей об'єктів, що включають прикладні програми, керовані зовнішніми подіями. Можливості ООБД розширює підтримка класів, успадкування, інкапсуляції, перевантаження функцій, управління паралелізмом, асоційованих запитів і ряду інших можливостей.

Поява ООБД обумовлено тим, що реляційні підходи погано застосовні для роботи з даними складних типів і складної структури, так як їх важко моделювати за допомогою наявних в реляційній моделі абстракцій. ООБД дозволяє представляти об'єкти предметної області в природному для них і семантично значущу вигляді, що суттєво спрощує розробку і модифікацію схеми даних.

Спроби будувати системи підтримки прийняття рішень (СППР), які зверталися б безпосередньо до баз даних систем оперативної обробки транзакцій (OLTP-систем), часто стикаються з низкою труднощів, таких, як: установка блокувань даних при виконанні транзакцій аналітичних запитів; значні часові затримки при генерації складних уявлень на основі реляційних зв'язків між таблицями оперативних і рідко змінюваних даних; гетерогенність інформаційних систем, які виступають в ролі джерел даних для СППР.

З метою забезпечення можливості аналізу накопичених даних стали створюватися сховища даних, що представляють собою інтегровані колекції даних, зібраних з різних оперативних систем. Сховища стають в даний час основою побудови СППР.

Для назви наукового напрямку, метою якого є розробка методів і засобів отримання корисної інформації та знань з баз і сховищ даних, в англійській літературі використовуються два терміни:

- «Data Mining» (DM), сукупність методів вилучення корисної інформації, окремих патернів, фрагментів знань з баз даних;
- «Knowledge Discovery in Databases» (KDD), сукупність методів формалізації знань на основі патернів, «заготовок», що виникли в процесі DM, зокрема - формування асоціативних, логічних правил і інших моделей подання знань, з яких і формуються бази знань прикладних систем [3].

DM включає в себе статистичні методи і технології виявлення патернів в необроблених даних за допомогою пошуку за зразком (pattern matching). Термін найчастіше використовується статистиками і аналітиками. Програмне забезпечення відповідної функціональності задіяний зазвичай в інформаційних системах керівника, OLAP-системах і т.д.

На відміну від DM, KDD відноситься до широкого спектру завдань пошуку знань і використовується в додатках вищого рівня. Термін вживається в області штучного інтелекту і, зокрема, в машинному навчанні [], якщо навчальні та тестові дані представлені в БД. Більш формальні визначення та детальні пояснення термінів DM і KDD розглядаються далі.

Зазвичай процес KDD є інтерактивним і ітеративним, задіює ряд методів та технологій DM, інтерпретує і формалізує результати його роботи. Процес KDD створює високорівневу модель для бази знань, використовуюваної далі в системі підтримки прийняття рішень.

Концептуально KDD є «нетривіальний процес виявлення потенційно корисних, обґрунтованих, раніше невідомих і змістовно зрозумілих закономірностей в базах даних». Наведемо визначення окремих понять, використаних в даному визначенні.

Дані - значення полів запису в базі або сховище даних, іноді також називаються фактами.

Патерн - закономірність, формальна специфікація характерних властивостей набору фактів. Прикладами патернів є часто виникаючі поєднання атрибутів, відповідаючі їм предикати, асоціативні правила, послідовності в потоці подій.

Процес - послідовність фаз KDD, наприклад: відбір даних, трансформація, очищення, інтеграція, пошук закономірностей, отримання правил, структурування знань, визначення якості результату.

Обґрунтованість - оцінка якості закономірностей і отриманої бази знань в цілому на новому наборі даних, задана в термінах деякого множини метрик, показників якості (ймовірність виникнення технічних збоїв в системі, оцінка похибки моделі і т.д.).

Потенційна корисність - оцінка впливу використання результатів процесу KDD на рішення конкретних завдань, наприклад, економія ресурсів (технічних, енергетичних, тимчасових, фінансових).

Зрозумілість - наявність змістовної інтерпретації результатів процесу KDD в термінах конкретної предметної області.

Як узагальнююче поняття для процесів DM і KDD служить термін «інтелектуальний аналіз даних» (ІАД).

Розглянемо основні фази процесу здобуття знань з даних.

Очищення даних. На цьому етапі проводиться попередня обробка, алгоритмічно заповнюються пропущені значення, ідентифікуються і фільтруються помилкові дані.

Інтеграція даних. Якщо дані отримані з різних джерел, то для виконання аналізу їх необхідно завантажити в єдину базу даних або сховище, вирішити завдання зіставлення ідентифікаторів, типів і форматів даних, усунення дублікатів і т.д. Інтеграція даних вимагає розробки спеціальних алгоритмів і програмних засобів, які зазвичай є складовою частиною прикладних систем.

Відбір і трансформація даних. На цьому етапі виконуються нормалізація даних, агрегування (за критерієм часу, інтервалу значень

числового атрибута), узагальнення. Результатом трансформації даних має бути їх нове уявлення, що має якомога менше атрибутів різного типу, і при цьому перехід до такого представлення не повинен призводити до великих втрат інформації в контексті цільової задачі.

Пошук патернів знань. При пошуку закономірностей мова може йти про виявлення патернів, що дозволяють відрізнити нормально розвиваючі ситуації від аварійних та небезпечних. Далі патерни служать вихідним матеріалом для побудови баз знань, і подальшого використання в додатках для аналізу даних і прийняття рішень.

Оцінка патернів. Мета даного етапу - виділення з множини знайдених патернів, придатних для практичного використання і формування бази знань. Для оцінки патернів створюється набір метрик, відповідний розв'язуваній задачі і предметної області.

Представлення отриманих патернів. Раціональна форма подання патернів залежить від цільової задачі. Кінцевою формою можуть бути база знань, модель результатів, візуалізація результатів у вигляді графіків, діаграм і т.д.

На сьогоднішній день ІАД знайшов своє застосування в багатьох областях науки і техніки, в тому числі і в діагностуванні технічних пристроїв і процесів. У процесі діагностування можна виділити дві основні складові: збір інформації та формування експертного висновку про стан об'єкта, що діагностується. З розвитком технологій і ускладненням діагностованих об'єктів і систем, як до першої, так і до другої складової пред'являються все більш високі вимоги по рівню оперативності, продуктивності, достовірності, якості та автоматизації.

Остання вимога має на увазі повну або часткову заміну операторів і експертів програмно-апаратними комплексами. Проблема автоматичного збору інформації в більшості прикладних областей технічного діагностування вирішена практично повністю, але такі проблеми, як формування автоматичного висновку про стан об'єктів діагностування, і тим

більше оцінка можливості деградації значень вимірюваних параметрів і виникнення на пристрої серйозних відмов, до сих пір залишаються відкритими.

Основними причинами, що призводять до потреби в створенні систем аналізу даних, виявлення патернів і синтезу експертних моделей залежностей між діагностичними станами, систем підтримки прийняття рішень про необхідність проведення технічного обслуговування і ремонту пристроїв, є наступні:

- наявність великих БД «сирий» діагностичної інформації, аналіз якої силами експерта є трудомістким і не завжди ефективним на увазі великих витрат часу і інших ресурсів;

- високі темпи і обсяги надходження нових даних в режимі реального часу, що вимагає поліпшення реакції системи при виявленні передумов виникнення несправностей і аномалій, і підтримки прийняття оперативних заходів щодо усунення несправностей;

- необхідність надання підтримки операторам при роботі в критичних за часом реагування ситуаціях, зниження кількості (в ідеалі - повне виключення) помилок інтерпретації ситуацій.

Всі перераховані вище ознаки проявляються при централізації технічного діагностування пристроїв АТ. На сьогоднішній день в цій галузі діагностування поки не застосовуються ні системи аналізу даних, ні системи підтримки прийняття рішень. Як зазначено в розділі 1.2, така ситуація характерна практично для всіх компаній-розробників систем ТДМ, в цьому напрямку тільки починають вестися дослідження.

В цілому ІАД застосовується до завдань, де застосування формалізованої евристики не припускає повноти знань і вихідних даних. У ряді предметних областей діагностики, в тому числі і на транспорті, виникають ситуації, коли висновки експертів про стан об'єкта при одних і тих же даних можуть відрізнятися.

Важливим позитивним фактором впровадження ІАД в процес технічного діагностування є можливість автоматизованого виявлення деяких закономірностей, що не формалізовані експертами раніше. Приховані в діагностичних даних знання можуть відображати деталізовані, або ключові моменти експертної інтерпретації і принести суттєву користь як при оцінці поточного стану об'єкта, так і при оцінці перспектив (прогнозуванні) його подальшої зміни.

1.4. Висновки до розділу 1

1. Проведений аналіз стану та актуальних проблем технічного діагностування пристроїв автоматики та телемеханіки дозволив виявити потребу в застосуванні моделей і методів, що володіють низьким ступенем помилкових спрацьовувань і допускають некритичне відхилення вимірюваного параметра від заданих норм, а також важливість для системи діагностування таких параметрів, як наявність розподіленої структури, властивість розширюваності і функцій інформаційної взаємодії з суміжними системами.

2. Розгляд сучасних систем діагностування технічних пристроїв, застосовуваних на транспорті і в інших областях, дозволило виявити відставання систем діагностування пристроїв АТ від систем схожого призначення інших галузей за рівнем інтелектуалізації та реалізації експертних функцій, а саме, виявити відсутність функціональності формування експертних оцінок стану технічних пристроїв, виявлення прихованих закономірностей і тенденцій за допомогою методів інтелектуального аналізу даних, прогнозування стану контрольованих пристроїв.

3. Виконаний аналіз стану, проблем та принципів реалізації сучасних систем зберігання і аналізу даних дозволив зробити висновок про актуальність розробки нових моделей уявлення і методів аналізу даних.

4. Визначено необхідність застосування інтелектуальних методів аналізу даних і експертних систем, що дозволить усунути відставання в рівні

автоматизації і підвищить ефективність роботи персоналу центрів діагностування та моніторингу, а також забезпечить підтримку прийняття рішень при обробці позаштатних діагностичних станів контрольованих об'єктів.

РОЗДІЛ 2

ТЕОРІЯ І МЕТОДОЛОГІЯ ПОБУДОВИ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ АНАЛІЗУ ДАНИХ

Аналіз актуальних проблем технічного діагностування пристроїв автоматики і телемеханіки, а також функціональності сучасних систем діагностування, визначені в Розділі 1, показав актуальність вирішення задач автоматизації, інтелектуалізації та підтримки експертних функцій.

До найбільш важливих етапів розробки інтелектуальної системи відноситься вибір форми подання знань в системі, а також вибір методів аналізу даних і виявлення знань про закономірності і залежності.

Розділ 2 присвячений аналізу існуючих моделей представлення знань, методів і алгоритмів інтелектуального аналізу даних, а також методів формування міркувань в умовах невизначеності і неповноти даних. У підрозділі 2.1 розглянуті логічні, мережеві, продукційні, фреймові моделі подання знань. Пункт 2.2 присвячений методам інтелектуального аналізу даних. У підрозділі 2.3 розглянуті методи формування міркувань в умовах неповних даних.

2.1. Представлення знань в інтелектуальних системах.

Тематика подання знань (Knowledge Representation, KR) вважається одним з основних напрямів робіт в області штучного інтелекту, оскільки вибір способу представлення знань, відповідного предметній області і розв'язуваної задачі, є не менш значущим чинником, від якого залежить успішне створення системи, ніж використовувані методи аналізу даних.

Кафедра АЕМ				НАУ 21 03 86 000 ПЗ			
Виконав	Кузюра А.А.			ТЕХНОЛОГІЯ І МЕТОДОЛОГІЯ ПОБУДОВИ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ АНАЛІЗУ ДАНИХ	Літера	Аркуш	Аркушів
Керівник	Товкач С.С.					1	30
Консульт.							
Н-контр.	Товкач С.С.						
Зав. каф.	Захарченко В.П.						
					АТ-412Б	15	Б

Від правильного вибору способу такого уявлення залежить весь хід розробки, а також ефективність, швидкодію і зручність супроводу інтелектуальної системи.

Характер, структура і походження знань розглядаються в рамках епістемології. Виділяють апріорні і апостеріорні знання. Апріорні знання передують знанням, отриманим за допомогою органів почуттів, і не залежать від них. Апріорні знання розглядаються як універсально справжні, і ці знання неможливо спростувати, не впадаючи в протиріччя. Апостеріорні знання протилежні по відношенню до апріорним, виходять за допомогою органів почуттів. Істинність або хибність апостеріорного знань може бути перевірена на основі чуттєвого досвіду, експерименту.

Знання можуть додатково поділятися на процедурні, декларативні і неявні. Процедурними часто називають знаннями про те, як за допомогою деякого набору дій досягти тієї чи іншої мети. Декларативними називають знання про те, чи є деяке твердження істинним або хибним. Неявні знання іноді називають підсвідомими знаннями, оскільки вони не можуть бути виражені за допомогою мови. Як приклад можна вказати знання ваг синаптичних зв'язків, представлені в штучній нейронній мережі, що не мають безпосередньої інтерпретації.

Співвідношення між рівнями інформаційної цінності даних і знань можна представити у вигляді ієрархії (див. рис.2.1) []. На нижньому рівні цієї ієрархії знаходиться шум, що складається з інформаційних елементів, які не представляють інтересу і можуть лише ускладнити сприйняття та інтерпретацію даних. На більш високому рівні знаходяться безформатні (тобто «сирі») дані. На наступному рівні перебуває інформація, тобто оброблені і відформатовані дані, явно що представляють інтерес для користувачів. Далі слідує рівень знань, на якому представлена узагальнена, винятково важлива інформація.



Рис. 2.1. Ієрархія видів представлення інформації

Відповідно, інтелектуальні системи повинні бути здатні до виконання наступних дій:

- відділення даних від шуму;
- перетворення даних в інформацію;
- витяг з інформації знань.

Існують різні моделі подання знань. До них відносяться логічні, мережеві, продукційні, фреймові моделі.

В основі логічних моделей лежить формальна система, що задається четвіркою виду: $M = \langle T, P, A, B \rangle$. T множина базових елементів різної природи, наприклад слів з деякого обмеженого словника, деталей дитячого конструктора, що входять до складу деякого набору і т.д. Важливо, що для множини T існує деякий спосіб визначення належності або неналежності довільного елемента до цієї множини. Процедура такої перевірки може бути будь-якою, але за кінцеве число кроків вона повинна давати позитивну або негативну відповідь на питання, чи є x елементом множини T . Позначимо цю процедуру $P(T)$.

Множина P є множина синтаксичних правил. З їх допомогою з елементів T утворюють синтаксично правильні сукупності. Декларується

існування процедури $P(P)$, за допомогою якої за кінцеве число кроків можна отримати відповідь на питання, чи є сукупність X синтаксично правильною.

У множині синтаксично правильних сукупностей виділяється деяка підмножина A . Елементи A називаються аксіомами. Як і для інших складових формальної системи, повинна існувати процедура $P(A)$, за допомогою якої для будь-якої синтаксично правильної сукупності можна отримати відповідь на питання про приналежність її до множини A .

Множини B є множиною правил виведення. Застосовуючи їх до елементів A , можна отримувати нові синтаксично правильні сукупності, до яких знову можна застосовувати правила з B . Так формується множина виведених в даній формальній системі сукупностей. Якщо є процедура $P(B)$, за допомогою якої можна визначити для будь-якої синтаксично правильної сукупності, чи є вона виводимою, то відповідна формальна система називається вирішуваною. Це показує, що саме правило виведення є найбільш складною складовою формальної системи.

Для знань, що входять в базу знань, можна вважати, що множина A утворює всі інформаційні одиниці, які введені в базу знань ззовні, а за допомогою правил виведення з них виводяться нові похідні знання. Іншими словами, формальна система являє собою генератор породження нових знань, що утворюють множину виведених в даній системі знань. Це властивість логічних моделей сприяє їх використанню в базах знань, тому що дозволяє зберігати в базі лише ті знання, які утворюють множину A , а всі інші знання отримувати з них за правилами виведення.

В основі мережевих моделей лежить конструкція, яка називається семантичною мережею. Мережеві моделі формально можна задати у вигляді $H = \langle I, C_1, C_2, \dots, C_n, I \rangle$. Тут I є множина інформаційних одиниць; C_1, C_2, \dots, C_n - множина типів зв'язків між інформаційними одиницями. Відображення Γ задає між інформаційними одиницями, що входять в I , зв'язки із заданого набору типів зв'язків.

Залежно від типів зв'язків, які використовуються в моделі, розрізняють класифікуючі мережі, функціональні мережі і сценарії. У класифікуючих мережах використовуються відносини структуризації. Такі мережі дозволяють в базах знань вводити різні ієрархічні відносини між інформаційними одиницями. Функціональні мережі характеризуються наявністю функціональних відносин. Їх також називають обчислювальними моделями, тому що вони дозволяють описувати процедури "обчислень" одних інформаційних одиниць через інші. У сценаріях використовуються каузальні відносини, а також відносини типів "засіб - результат", "знаряддя-дія" і т.д. Якщо в мережевій моделі допускаються зв'язку різного типу, то її зазвичай називають семантичною мережею.

Семантичні мережі - це класичний спосіб представлення пропозиціональної інформації, що використовується в штучному інтелекті, тому семантичні мережі іноді називають пропозиціональними мережами. Як відомо, пропозиціональним твердженням (або висловлюванням) називається пропозиція, яке може бути або істинним, або хибним. Висловлювання мають форму декларативних знань, оскільки в них затверджуються факти. Формально семантична мережа являє собою позначений орієнтований граф. Висловлення завжди є або істинним, або хибним і називається атомарним (неподільним), так як його істинне значення не підлягає подальшій декомпозиції. На відміну від цього, нечіткі твердження, не повинні обов'язково бути тільки істинними або тільки хибними.

Структура семантичної мережі відображається графічно за допомогою вузлів і з'єднуючих їх дуг. Вузли іноді іменуються об'єктами а дуги - зв'язками, або ребрами. Зв'язки в семантичної мережі застосовуються для подання відносин, а вузли, як правило, служать для подання фізичних об'єктів, концепцій або ситуацій.

Семантичні мережі іноді називають асоціативними мережами, оскільки одні вузли в таких мережах асоційовані, або пов'язані, з іншими.

Особливо слід виділити такі широко використовувані типи зв'язків, як IS-A (є екземпляром деякого класу об'єктів) та АКО (A-KIND-OF, відноситься до типу деякого класу). На рис. 2.2 показана семантична мережа уніфікованих типів пристроїв АТ, в якій показано застосування таких зв'язків. В даному контексті поняття класу близько до математичного поняття множини, оскільки служить для позначення групи об'єктів, що володіють ідентичним набором властивостей. В даному випадку зв'язок АКО визначає відношення між окремим класом і його класом-предком (або класами).

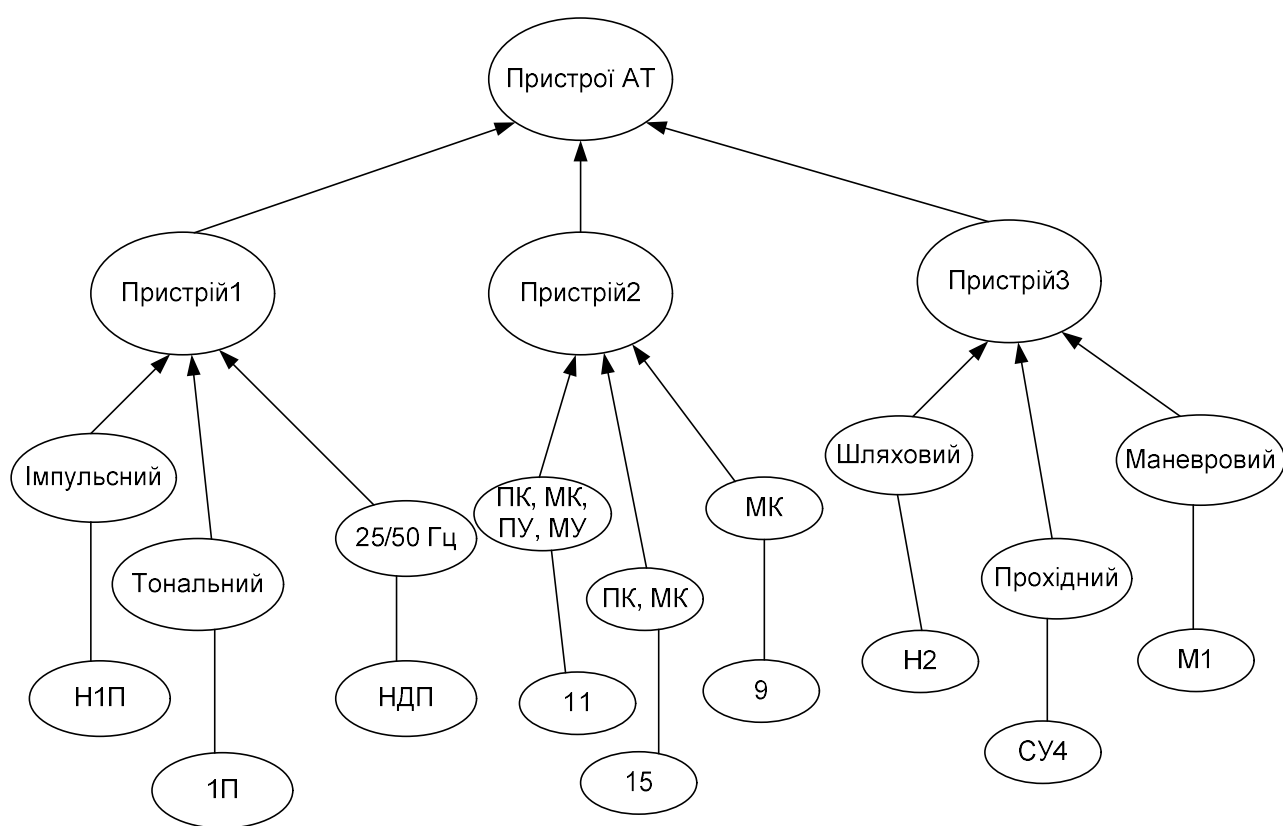


Рис. 2.2. Семантична мережа зі зв'язками IS-A і A-KIND-OF

Слід зазначити, що на рис. 2.2 більш загальні класи знаходяться у верхній частині, більш конкретизовані класи - в середній, а екземпляри класів - в нижній частині. Самий загальний клас, на який вказують стрілки АКО, називається суперкласом.

Всі об'єкти класу повинні мати один або кілька загальних атрибутів. Кожен атрибут має значення. Комбінація атрибута і значення називається

властивістю: Наприклад, уніфікований тип «Пристрій1» (П1) має такі атрибути, як тип (імпульсний, тональний і т.д.), кількість відгалужень, набір дискретних і аналогових сигналів.

При відсутності певного свідчення, що дозволяє стверджувати зворотне, передбачається, що всі елементи деякого класу успадковують всі властивості суперкласів цього класу. Спадкування є корисний засіб в поданні знань, оскільки дозволяє позбутися від дублювання загальних характеристик. Зв'язки і спадкування є основою ефективних способів подання знань, оскільки дають можливість показувати багато складних відносин за допомогою декількох вузлів і зв'язків.

Ще однією широко застосовуваним зв'язком є HAS-A, який встановлює відношення між класом і підкласом, або між двома екземплярами різних класів - екземпляром-контейнером і вкладеним в нього екземпляром. Зв'язок HAS-A спрямований в бік, протилежний по відношенню до зв'язку АКО, і часто використовується для позначення відношення між одним об'єктом і частиною іншого об'єкта.

Такі поняття, як об'єкт, атрибут і значення, зустрічаються разом настільки часто, що іноді є можливість створити спрощену семантичну мережу з використанням тільки цих понять. Щоб охарактеризувати всі знання, представлені в семантичній мережі, можна скористатися трійкою "об'єкт-атрибут-значення" (object-attribute-value, OAV). Такий підхід використовувався в експертній системі MYCIN, призначеної для діагностування інфекційних захворювань. Подання у вигляді трійки OAV є зручним способом оформлення списку наявних знань у формі таблиць, а також дозволяє легко перетворити отриману таблицю в комп'ютерний код за методом індукції правил.

Трійки "об'єкт-атрибут-значення" лежать в основі корисного способу представлення фактів і шаблонів, що застосовуються для узгодження з фактами в антецеденте правила. Семантична мережа для подібної системи складається з вузлів, що представляють об'єкти, атрибути і значення,

з'єднаних зв'язками HAS-A і IS-A. Якщо повинен бути представлений тільки один об'єкт і застосовувати відносини спадкування не потрібно, то може виявитися прийнятним ще більш просте уявлення, зване парою "атрибут-значення" (attribute-value, AV).

Безумовно, семантичні мережі можуть виявитися вельми корисним засобом представлення знань, але вони мають деякі обмеження, пов'язані, зокрема, з подвійністю призначення зв'язків. Як було зазначено вище зв'язок може ставитися як до об'єктів, так і до їх типів. Семантична мережа може застосовуватися для подання категоричних (визначення однозначно) знань тільки за умови, суворого визначення типів використовуваних зв'язків і вузлів.

Ще одна проблема, пов'язана з використанням семантичних мереж, полягає в експоненційному зростанні часу пошуку вузлів, в разі, якщо специфіка завдання вимагає обробки моделей великої розмірності, і немає можливості використовувати різні евристичні методи.

З іншого боку, семантичні мережі, за умови однотипності зв'язків, відповідають вимогам логіки і дозволяють застосовувати для визначення знань механізми логічного висновку, а також використовувати евристичні методи, що допомагають знайти оптимальне, або близьке до оптимального рішення. У штучному інтелекті можливість застосування евристик дуже важлива, оскільки типові завдання штучного інтелекту настільки складні, що алгоритмічне рішення для них не існує або малоефективно для практичного використання.

Ще одним способом удосконалення семантичних мереж, є застосування механізму закріплення за вузлами мережі спеціальних процедур узгодження значень трійок "об'єкт-атрибут-значення". Такі процедури виконуються після активації вузла, наприклад при переході контрольованого пристрою в діагностичний стан, представлений даним вузлом.

У продукційних моделях використовуються деякі елементи логічних і мережевих моделей. З логічних моделей запозичена ідея правил виведення,

які тут називаються продукціями, а з мережевих моделей - опис знань у вигляді семантичної мережі. В результаті застосування правил виведення до фрагментів мережевого опису відбувається трансформація семантичної мережі за рахунок зміни її фрагментів, нарощування мережі або виключення фрагментів. Таким чином, в продукційних моделях процедурна інформація явно виділена і описується іншими засобами, ніж декларативна інформація. Замість логічного висновку, характерного для логічних моделей, в продукційних моделях з'являється висновок на знаннях.

Однією з формальних систем позначень, що застосовуються для визначення продукцій (наборів правил), є нормальна форма Бекуса-Наура (Backus-Naur form, BNF). Ця система позначень є метамова, яка може застосовуватись для визначення синтаксису будь-якої мови. Синтаксис визначає форму, а семантика позначає сенс. Нижче наведено продукційне правило, в якому система позначень на основі BNF використовується для формулювання простого правила природної мови, згідно з яким пропозиція (sentence) складається з підмета (subject), присудка (verb), за якими слідує знак пунктуації, що позначає кінець пропозиції (end- mark).

< sentence > ::= < subject > < verb > < end - mark >

У цьому правилі кутові дужки $\langle \rangle$ і символ $:: =$ являють собою символи метамови. Символ $:: =$ означає "визначено як і являє собою еквівалент стрілки ($=>$), що застосовується в нормальній формі Бекуса-Наура, вертикальна риса означає "або". Повний набір продукційних правил, який однозначно визначає мову, називають граматикою.

На відміну від моделей розглянутих типів, у фреймових моделях фіксується жорстка структура інформаційних одиниць, яка називається протофрейма. Фрейми запропоновані для використання в якості методу розуміння особливостей зору, природних мов та інших областей і надають зручну структуру для опису об'єктів, типових для якоїсь конкретної ситуації, зокрема, стереотипів об'єктів.

Якщо семантичні мережі по суті можна розглядати як двовимірне уявлення знань, то фрейми додають третій вимір, оскільки дозволяють використовувати вузли, що мають внутрішню структуру. В якості таких структур можуть застосовуватися прості значення або інші фрейми.

Основна характерна особливість фрейма полягає в тому, що він представляє взаємопов'язані знання з вузької теми, значна частина яких - це знання, задані за замовчуванням. Система фреймів може виявитися вельми відповідною для опису технічних пристроїв. Відмінність фрейму від семантичної мережі полягає в тому, що мережа, як правило, використовується для загального уявлення знань, а фрейм дозволяє їх конкретизувати. Проте, як і в випадку семантичних мереж, стандарти визначення систем на основі фреймів відсутні.

Фрейм аналогічний структурі записи на мові високого рівня, такому як C ++ або C #. Фрейм по суті являє собою групу слотів (полів) і наповнювачів (значень), що характеризують об'єкт.

У загальному випадку фрейми складні і виявляють свою значимість в складі ієрархічних систем і в умовах застосування успадкування. Використання фреймів у вигляді наповнювачів слотів і введення в дію відносини спадкування дозволяє створювати дуже потужні системи подання знань, зокрема, дозволяючи організувати подання причинно-наслідкових знань, на відміну від продукційних систем, заснованих на правилах, і спираються на неорганізовані знання, які не є причинними.

Заповнювачами слотів фреймів можуть бути значення, або посилання на процедури, закріплені за слотами і звані процедурними вкладеннями. У фреймах значення, передбачені за замовчуванням, є виключно важливими, оскільки дозволяють моделювати деякі аспекти функціонування природного інтелекту. Застосовувані за замовчуванням значення відповідають очікуванням людини щодо деякої ситуації, які складаються на підставі досвіду. Після виявлення нової ситуації здійснюється модифікація відповідного фрейму, що дозволяє простіше пристосуватися до ситуації.

Задані за замовчуванням значення часто використовуються для представлення знань, які характеризуються як базові, або засновані на здоровому глузді.

Системи фреймів проектуються таким чином, щоб більш універсальні фрейми знаходилися ближче до вершини ієрархії. Передбачається, що спеціалізація фреймів стосовно конкретних випадків може здійснюватися шляхом модифікації застосовуваних за замовчуванням значень і створення більш конкретних фреймів. За допомогою фреймів може бути зроблена спроба моделювати об'єкти реального світу, використовуючи універсальні знання для опису більшості атрибутів того чи іншого об'єкта і більш конкретні знання - для опису окремих випадків.

Іншими словами, верхні рівні в ієрархії фреймів представляють характеристики, більшою мірою віднесені до усіх об'єктів предметної області, а нижні рівні позначають нечіткі межі між об'єктами реального світу. Об'єкт, що володіє всіма типовими характеристиками, прийнято називати прототипом.

Проте, не дивлячись на ряд переваг, при використанні систем фреймів виявляються суттєві недоліки, пов'язані з тим, що в таких системах допускається необмежена модифікація або знищення слотів. Це може призводити до розсинхронізації значень слотів подібних фреймів, а також до небажаних узагальнень (ідентичності слотів у різних фреймів).

У більшості систем фреймів не передбачені способи визначення незмінних слотів. А оскільки в зв'язку з цим може зазнати змін будь-який слот, то властивості, успадковані одним фреймом від іншого, можуть бути змінені або виключені на будь-якому рівні ієрархії. У такій необмеженій системі важко формулювати загальнозначущі твердження, що стосуються всіх представників типу. Крім того, немає можливості досить надійно створювати складні об'єкти на основі простіших визначень.

На основі виконаного аналізу існуючих підходів до подання знань в інтелектуальних системах, і з урахуванням вимог, що пред'являються до

моделі знань про залежності між діагностичними станами, таких як наочність, компактна форма візуалізації, можливість застосування процедур логічного висновку, можна зробити висновок про доцільність застосування мережевого підходу до подання знань при побудові моделі.

2.2. Методи інтелектуального аналізу даних

Випереджаючи розгляд методів інтелектуального аналізу даних, можна привести цитату Дж.Маккарті, одного з основоположників штучного інтелекту, що пояснює застосування в даному контексті терміну «інтелектуальність»: «... Ми розуміємо деякі механізми інтелекту і не розуміємо інші. Тому під інтелектом ... розуміється тільки обчислювальна складова здатності приймати рішення і досягати цілей ... ».

Як було відзначено в п. 1.4, метою інтелектуального аналізу даних є отримання корисної інформації, окремих патернів і фрагментів знань з баз даних, формування моделей уявлення знань.

Найбільш часто за допомогою методів ІАД вирішуються такі прикладні задачі []:

- визначення значущих (домінантних) чинників;
- аналіз і класифікація значущих чинників;
- скорочення або збільшення їх числа;
- виявлення залежностей, асоціацій (пошук асоційованих груп значень факторів, тобто значень, майже завжди з'являються разом);
- виявлення закономірностей і винятків, в тому числі для зменшення розмірності факторного простору;
- моделювання та прогнозування процесів;
- ранжування (упорядкування факторів за ступенем їх впливу на цільовий індикатор);
- сегментація (розподіл області значень фактора на підобласті для проведення подальшого спадного аналізу);

- профілювання найкращих досягнень (виявлення основних характеристик найбільш вірних рішень, прогнозів і т.д.) [].

Для вирішення цих завдань в інтелектуальному аналізі даних застосовуються методи дерев рішень, нейронних мереж, математичної статистики, експертного аналізу і нечіткої логіки, візуалізації, генетичних алгоритмів, еволюційного програмування [], асоціативного аналізу, а також інтегровані методи і технології.

Можна виділити два види моделей даних, що застосовуються в ІАД:

- прогнозувальні (predictive) моделі, що включають в себе моделі класифікації і моделі послідовностей;

- описові (descriptive) моделі, що включають в себе моделі регресії, кластеризації, винятків, асоціацій і т.д.

Розглянемо особливості і сфери застосування основних методів ІАД.

1. Древа рішень

Древа рішень - це метод, що дозволяє передбачати приналежність спостережень або об'єктів до того чи іншого класу категоріальної залежної змінної в залежності від відповідних значень однієї або декількох прогнозованих змінних [].

Даний метод придатний тільки для вирішення задач класифікації.

В результаті застосування цього методу до навчальної вибірки даних створюється ієрархічна структура класифікуючих правил типу "ЯКЩО ... ТО ...", що має вигляд дерева.

Для того щоб вирішити, до якого класу віднести деякий об'єкт або ситуацію, потрібно перевірити істинність логічних виразів, поставлених у відповідність вузлів цього дерева, починаючи з його кореня.

Вирази мають вигляд "значення параметра A більше x ". Якщо цей вислів вірний, переходимо до правого вузла наступного рівня, інакше – до лівого вузла; потім перевіряємо істинність виразу, пов'язаного з цим вузлом. Таким чином, ми доходимо до одного з кінцевих вузлів (листя), де стоїть вказівник, до якого класу треба віднести даний об'єкт.

Цей метод хороший тим, що таке уявлення правил наочно і його легко зрозуміти. Але для дерев рішень дуже гостро стоїть проблема значущості. Справа в тому, що окремим вузлам на кожному новому побудованому рівні дерева відповідає все менше і менше число записів даних - дерево дробить дані на велику кількість окремих випадків. Чим більше цих окремих випадків, чим менше навчальних прикладів потрапляє в кожен такий окремий випадок, тим менш упевненою стає їх класифікація.

Якщо побудоване дерево занадто "рунисте" - складається з невиправдано великого числа дрібних гілочок - воно не буде давати статистично обґрунтованих відповідей. Як показує практика, в більшості систем, що використовують дерева рішень, ця проблема не знаходить задовільного рішення.

2. Нейронні мережі

Нейронні мережі - клас аналітичних методів, побудованих на (гіпотетичних) принципах навчання мислячих істот і функціонування мозку і дозволяють прогнозувати значення деяких змінних в нових спостереженнях за даними інших спостережень (для цих же або інших змінних) після проходження етапу так званого навчання на наявних даних [].

При застосуванні нейронних методів, перш за все, постає питання вибору конкретної архітектури мережі (числа "шарів" і кількості "нейронів" в кожному з них). Розмір і структура мережі повинні відповідати (наприклад, в сенсі формальної обчислювальної складності) суті досліджуваного явища.

Після того як мережа побудована, починається процес «навчання» мережі. На цьому етапі нейрони мережі ітеративно обробляють вхідні данні і коректують свої ваги таким чином, щоб мережа найкращим чином прогнозувала (в традиційних термінах - "здійснювала підгонку") дані, на яких виконується "навчання". Після навчання на наявних даних мережа готова до роботи і може використовуватися для побудови прогнозів.

"Мережа", отримана в результаті "навчання", висловлює закономірності, присутні в даних. При такому підході вона виявляється

функціональним еквівалентом певної моделі залежностей між змінними, подібної тим, які будуються в традиційному моделюванні.

Однак, на відміну від традиційних моделей, в разі "мереж" ці залежності не можуть бути записані в явному вигляді, подібно до того, як це робиться в статистиці (наприклад, "А позитивно корелювали з В для спостережень, у яких величина С мала, а D - велика").

Іноді нейронні мережі видають прогноз дуже високої якості; однак, вони представляють собою типовий приклад нетеоретичного підходу до дослідження (іноді це називають "чорним ящиком").

При такому підході дослідник націлений виключно на практичний результат (в даному випадку - на точність прогнозів і їх прикладну цінність), а не на вивчення суті механізмів, що лежать в основі явища, або відповідності отриманих результатів будь-якої наявної "теорії".

Однак, методи нейронних мереж можуть застосовуватися і в таких дослідженнях, де метою є побудова пояснюючої моделі явища, оскільки нейронні мережі допомагають вивчати дані на предмет пошуку значущих змінних або груп таких змінних, і отримані результати можуть полегшити процес подальшої побудови моделі. Зокрема, спеціальні нейромережі за допомогою складних алгоритмів можуть знаходити найбільш важливі вхідні змінні, що вже безпосередньо допомагає будувати модель.

Одне з головних переваг нейронних мереж полягає в тому, що вони, принаймні теоретично, можуть апроксимувати будь-яку неперервну функцію, і тому досліднику немає необхідності заздалегідь приймати будь-які гіпотези щодо моделі, і навіть - в ряді випадків - про те, які змінні дійсно важливі.

Істотним недоліком нейронних мереж є та обставина, що остаточне рішення залежить від початкових установок мережі та, як уже говорилося вище, його практично неможливо "інтерпретувати" в традиційних аналітичних термінах, які зазвичай застосовуються при побудові теорії явища.

Деякі автори відзначають той факт, що нейронні мережі використовують або, точніше, припускають використання обчислювальних систем з масовим паралелізмом. Наприклад, в [] нейронна мережа визначається, як "процесор з масивним розпаралелюванням операцій, що володіє природною здатністю зберігати експериментальні знання і робити їх доступними для подальшого використання. Він схожий на мозок в двох відносинах: мережа набуває знання в результаті процесу навчання; для зберігання інформації використовуються величини інтенсивності міжнейронних з'єднань, які називаються синоптичними вагами".

Однак, як зазначено в [], більшість існуючих нейромережевих програм працюють на однопроцесорних комп'ютерах. Істотне прискорення роботи може бути досягнуто не тільки за рахунок розробки програмного забезпечення, що використовує переваги багатопроцесорних систем, але також шляхом розробки більш ефективних алгоритмів навчання.

3. Методи візуалізації

Методи візуалізації служать для багатовимірного графічного представлення інформації з метою знаходження залежностей, трендів і зсувів, прихованих в неструктурованих наборах даних.

До методів візуалізації відносяться:

- представлення даних у вигляді стовпчастих, лінійних діаграм в багатовимірному просторі;
- накладення і об'єднання декількох зображень;
- ідентифікація та маркування підгруп даних, які відповідають певним умовам;
- розщеплення або злиття підгруп даних на графіку;
- агрегування даних;
- згладжування даних;
- побудова піктографіків;
- створення мозаїчних структур, спектральних площин, карт ліній рівня;

- методи динамічного обертання і динамічного розшарування тривимірних зображень;

- виділення певних наборів і блоків даних і т.д.

4. Генетичні алгоритми

Генетичний алгоритм - алгоритм пошуку оптимальної бітового рядка, який випадковим чином вибирає початкову популяцію таких рядків і потім піддає їх процесу штучних мутацій, схрещування і відбору за аналогією з природним відбором.

Цей метод названий так тому, що в якійсь мірі імітує процес природного відбору в природі. Нехай нам треба знайти рішення задачі, оптимальне з точки зору деякого критерію. Нехай кожне рішення повністю описується деяким набором чисел або величин нечислової природи.

Наприклад, необхідно вибрати сукупність параметрів ринку, що впливають на динаміку ринку, тут рішенням буде набір імен цих параметрів. Про це наборі можна говорити як про сукупність хромосом, що визначають якості індивіда - даного рішення поставленого завдання.

Значення параметрів, що визначають рішення, будуть тоді називатися генами. Пошук оптимального рішення при цьому схожий на еволюцію популяції індивідів, представлених їх наборами хромосом.

У цій еволюції діють три механізми:

- відбір найсильніших - наборів хромосом, яким відповідають найбільш раціональні рішення;

- схрещування - виробництво нових індивідів за допомогою змішування хромосомних наборів відібраних індивідів;

- мутації, - випадкові зміни генів у деяких індивідів популяції.

В результаті зміни поколінь, врешті-решт, виробляється таке рішення поставленого завдання, яке вже не може бути далі покращено. Приклади розв'язання задач за допомогою генетичних алгоритмів наведені в роботі [].

Генетичні алгоритми мають два слабких місця. По-перше, сама постановка завдання в їх термінах не дає можливості проаналізувати

статистичну значущість одержуваного з їх допомогою рішення і, по-друге, ефективно сформулювати завдання, визначити критерій відбору хромосом під силу тільки фахівцеві.

5. Еволюційний програмування

Сьогодні це наймолодша і найбільш перспективна галузь ІАД, реалізована, наприклад, в системі Poly Analyst.

Суть методу в тому, що гіпотези про вид залежності цільової змінної від інших змінних формуються системою у вигляді програм на деякому внутрішньому мові програмування.

Якщо це універсальна мова, то теоретично на ньому можна висловити залежність будь-якого виду. Процес побудови цих програм будується як еволюція в світі програм (цим метод трохи схожий на генетичні алгоритми).

Коли система знаходить програму, досить точно виражаючу шукану залежність, вона починає вносити в неї невеликі модифікації і відбирає серед побудованих таким чином дочірніх програм ті, які підвищують точність.

Таким чином система "вирощує" кілька генетичних ліній програм, які конкурують між собою в точності вираження шуканої залежності.

Спеціальний транслюючий модуль системи переводить знайдені залежності з внутрішньої мови системи на зрозумілу користувачеві мову (математичні формули, таблиці та ін.). Можливе застосування різних засобів візуалізації виявлених залежностей. Для контролю статистичної значущості виведених залежностей застосовується вибіркоче модульне тестування програм, що створюються системою.

6. Асоціативний аналіз

Ідея методу асоціативного аналізу (міркування на основі аналогічних випадків, case based reasoning, CBR) полягає в наступному: для того щоб зробити прогноз подальшого розвитку ситуації або вибрати правильне рішення, виконується пошук аналогій в базі даних і вибирається те ж рішення, яке було визнано для аналогічній ситуації правильним раніше [].

До даної групи методів належать алгоритми типу Lazy-Learning: метод "найближчого сусіда" (nearest neighbour); метод k-найближчого сусіда (k-nearest neighbour).

Методи CBR показують хороші результати в найрізноманітніших задачах. Основний їх недолік полягає в тому, що вони взагалі не створюють будь-яких моделей або правил, узагальнюють попередній досвід: у виборі рішення вони ґрунтуються на всьому масиві доступних історичних (протокольованих) даних, тому складно виділити фактори, на основі яких CBR-методи отримують ті чи інші рішення для поставлених завдань.

Поряд з перерахованими методами, в ІАД застосовуються і класичні статистичні методи. До них відносяться: кореляційний і регресійний аналіз; методи ідентифікації та прогнозування часових рядів; кластерний, факторний, дискримінантний, спектральний аналіз і т.д. Істотним недоліком методів даної групи є орієнтованість на усереднені характеристики статистичної вибірки, що не дозволяє розраховувати на отримання змістовної інтерпретації складних феноменів, що зустрічаються в реальній практиці.

Порівняльний аналіз перерахованих вище методів ІАД (табл.3.5) показує, що для задач кластеризації та класифікації даних найбільш обґрунтованим є застосування методу дерев рішень, для задач виявлення залежностей і формування прогнозів - методів еволюційного програмування та асоціативного аналізу, а для взаємодії з користувачем-експертом - методи візуалізації (як знайдених залежностей, так і самих даних).

Таблиця 2.1 Основні переваги і недоліки методів ІАД

Переваги	Недоліки
Дерева рішень (Decision Trees)	
Наочність графічного представлення результатів. Простота інтерпретації отриманих результатів	Використовуються тільки для вирішення задач класифікації. Проблема статистичної значущості.
Нейронні мережі (Neural Nets)	
Не вимагають апріорних припущень про природу досліджуваних даних. Зручні при роботі з нелінійними залежностями, зашумленими і неповними даними	Модель типу "чорний ящик" не може пояснити виявлені знання (не піддається інтерпретації).

Методи візуалізації (Vizualization methods)	
Наочність (багатовимірне графічне представлення даних, за яким аналітик сам виявляє закономірності і відношення між ними)	Высокая доля субъективности в интерпретации результатов. Отсутствие аналитических моделей.
Генетичні алгоритми (Genetic algorithms)	
Висока швидкість розв'язання задач великої розмірності. Подібність методу з природним механізмом (Імітація процесу природного відбору в природі).	Неможливість оцінки статистичної значущості результату. Складність використання методу (складність постановки задачі, складність визначення критерію відбору хромосом).
Еволюційне програмування (Evolutionary programming)	
Високий ступінь автоматизації (автоматичне виявлення в масивах даних кластерів, випадкових викидів, прихованих закономірностей, фільтрація шумів; візуалізація виявлених залежностей, оцінка статистичної значущості результатів і т.д.).	Складність приведення (трансляції) знайдених залежностей до виду, зрозумілому для користувача
Асоціативний аналіз (Case-based Reasoning)	
Простота (для формування прогнозу або вибору рішення знаходяться аналоги наявної ситуації, і вибирається та відповідь, яка була для них правильною в минулому).	У процесі рішення не створюються моделі і правила, узагальнюючи попередній досвід.

З урахуванням основних завдань діагностики технічних пристроїв, таких як раннє виявлення відхилень у роботі пристроїв, фіксація і аналіз передвідмовних станів і своєчасне попередження аварійних ситуацій і відмов, можна зробити висновок про необхідність застосування в розроблюваній системі інтелектуального аналізу даних таких методів, які дозволили б виявляти стійкі залежно змін станів, що призводять до відмови пристрою, і формувати на їх основі базу знань.

Порівняння методів еволюційного програмування та асоціативного аналізу показує, що з точки зору простоти представлення знайдених залежностей в зрозумілому для користувача системи вигляді, перевагу має остання група методів. Недолік методів асоціативного аналізу, що виражається в безпосередньому використанні даних зі сховища при кожному запиті, може бути усунутий шляхом розробки модифікованого комплексного методу, що відноситься до цієї групи, але більш орієнтованого на створення

моделей залежностей між діагностичними станами пристроїв для формування бази знань.

Таким чином, загальна ідея комплексного методу асоціативного аналізу полягає в наступному: на етапі аналізу даних виявляються і поміщаються в базу знань стійкі послідовності подій, формується модель залежностей, а для формування прогнозу подальшої зміни стану пристрою, або надання підтримки в прийнятті рішення, в основі знань виконується пошук підходящої послідовності за критерієм відповідності поточному і попереднім станам пристрою, і в залежності від завдання, робиться висновок про найбільш ймовірний варіант розвитку подій, або вибирається рішення, яке було визнано раніше правильним в аналогічній ситуації.

2.3. Методи формування міркувань за неповними даними

Методи формування міркувань в умовах невизначеності з використанням теорії ймовірностей і нечіткої логіки грають важливу роль в інтелектуальних системах, оскільки однією з найсильніших сторін подібної системи є її здатність вирішувати невизначені ситуації за образом дій справжнього експерта. Здатність справлятися з невизначеністю є одним з основних переваг інтелектуальної системи над простим деревом рішень, в якому всі факти повинні бути відомі заздалегідь, щоб можна було досягти результату.

Основою деяких теорій невизначеності є теорія ймовірностей, інші з них ґрунтуються на наближених міркуваннях. Кожна з цих теорій має свої переваги і недоліки. Вибір підходящої теорії, і відповідно методів оперування знаннями, має багато спільного з вибором алгоритмів і структур даних при розробці комп'ютерної програми. Розуміння переваг і недоліків кожного підходу до обробки невизначеності дозволяє створювати інтелектуальні системи, які найбільше підходять для моделювання експертних знань конкретної предметної області.

Невизначеність може розглядатися як нестача адекватної інформації для прийняття рішення. Невизначеність стає проблемою, оскільки може перешкодити виробленню найкращого рішення і навіть стати причиною того, що буде прийнято неякісне рішення. У технічному діагностуванні через некоректну обробку невизначеності можуть бути зроблені помилки при виявленні стані пристрою, упущена можливість попередження збоїв і відмов.

Розроблено цілий ряд теорій, що дозволяють успішно діяти в умовах невизначеності. До них відносяться:

- теорії, засновані на класичному визначенні ймовірностей і на Байєсовському підході до умовної ймовірності;
- теорія Хартлі, заснована на класичному визначенні множин;
- теорія Шеннона, заснована на понятті ймовірності;
- теорія нечітких множин Л. Заде.

Зокрема, теорія байєсовских ймовірностей і теорія нечітких множин виявилися вельми широко застосовними у багатьох областях, таких як фізика, біологія, інженерні науки, психологія. Нечітка логіка застосовується також у багатьох споживчих пристроях і приладах, електроніці, техніці і т.д.

Коли мова йде про невизначені факти, строгі міркування стають непридатними, тому що через невизначеність зростає кількість можливих результатів, тому може виявитися не просто складно, а й взагалі неможливо знайти найкраще рішення. Запропоновано цілий ряд різних способів організації дій в умовах невизначеності і засобів, що сприяють вибору найкращого ув'язнення. Багато з них з самого початку орієнтовані на пошук досить якісного, а не найкращого рішення. Причина такого підходу полягає в тому, що якісне рішення, досяжне на 99% в реальному часі, в більшості випадків виявляється більш прийнятним, ніж оптимальне рішення, для обчислення якого будуть потрібні десятки годин, днів, а можливо і років.

Саме на етапі проектування інтелектуальної системи необхідно виконати вибір методу, найбільш підходящого для даного конкретного додатку.

Ймовірнісні міркування і теорема Байєса успішно використовуються в якості альтернативи підходу на основі правил в прикладних експертних системах, де кількість етапів логічного виведення, необхідних для обґрунтування гіпотези або для досягнення укладення велика, а основна частина або навіть всі ці етапи логічного виведення здійснюються з урахуванням невизначеності свідоцтв і самих правил.

Підходящою архітектурою для таких експертних систем можна вважати мережу логічного виведення. Модель, призначена для використання в системі, представляється у вигляді мережі зі зв'язками (або відношеннями), що з'єднують свідоцтва і гіпотези. Таким чином, мережа логічного виведення є різновидом семантичної мережі. Спостережувані факти складають свідчення, що застосовується для обґрунтування проміжних гіпотез. Потім групи проміжних гіпотез використовуються для обґрунтування гіпотези верхнього рівня, яку потрібно підтвердити або відкинути. Якщо відмінність між свідченням і гіпотезою не важлива, то для позначення того чи іншого застосовується термін твердження.

У випадку, коли використовується експертний метод створення моделі, а експерти стикаються з труднощами при заданні значень апріорних ймовірностей, можуть також використовуватися коефіцієнти достовірності тверджень. Також, для комбінування свідоцтв, можуть використовуватися елементи нечіткої логіки.

Мережі логічного виведення мають статичну структуру знань. Це означає, що вузли і зв'язки між вузлами залишаються постійними. В цьому полягає принципова відмінність мереж логічного виведення від систем, заснованих на правилах, оскільки в таких системах правила, шаблони яких узгоджуються з фактами, поміщаються в робочий список, після чого здійснюється вирішення конфліктів, а потім застосовується правило з найвищим пріоритетом.

З іншого боку, хоча структура мережі логічного виведення залишається незмінною, ймовірно, пов'язані з кожним вузлом гіпотези, у міру отримання

свідоцтв змінюються. Основною характерною особливістю мережі логічного виведення є те, що в них у міру накопичення свідоцтв відбувається зміна ймовірностей, як апіорних, так і апостеріорних. Вказана зміна ймовірностей відбувається в напрямку до верхньої (гіпотетичної) частини мережі, що дозволяє в кінцевому підсумку обґрунтувати або спростувати перевіряючу гіпотезу.

Слід зазначити, що для кожної окремої області застосування потрібні свої моделі логічного виведення, коефіцієнти достовірності та формули нечіткої логіки. Саме тому необхідно мати можливість використовувати не тільки знання експертів, а й системи автоматизованої побудови моделей на основі вилучення знань про залежності в прикладній області з експериментальних даних.

Ще одним важливим методом формування міркувань в інтелектуальних системах при наявності невизначеності є наближені міркування, які також називають м'якими (на протипагу суворой двозначній логіці) обчисленнями, що ґрунтуються на нечіткій логіці. Ця теорія в основному присвячена проблемам оцінки кількості і формування міркувань з використанням природної мови, в якому багато слів, такі як "багато", "мало", "досить" і т.д., мають неоднозначний зміст.

Нечітка логіка - це надмножина звичайної (булевої) логіки, яка була доповнена з урахуванням поняття часткової істинності - істиннісних значень, що лежать між "повністю істинними" і "повністю помилковими". Визначення поняття нечіткої логіки не слід сприймати буквально, оскільки нечітка логіка - це не туманний, розпливчастий, невизначений спосіб мислення. Насправді даний підхід до формування міркувань є абсолютно протилежним невизначеному. Невизначеним стає те, що сприймається як надто складне для повного розуміння. Чим складніше виявляється деяке поняття, тим більше неточним або "нечітким" воно стає. Нечітка логіка надає точний підхід до врахування невизначеності, обумовленої складністю людської поведінки.

Поняття нечіткої множини було вперше сформульовано Заде і стало теоретичною основою для нечітких комп'ютерних систем і апаратних засобів. Теорія, сформульована Заде, привела до появи принципово нових гілок математики та інженерних наук. Увійшов в експлуатацію термін м'які обчислення, що охоплює нечітку логіку, нейронні мережі і імовірнісні міркування.

Звичайна чітка множина являє собою окремий випадок нечіткої множини з функцією приналежності $\{0,1\}$. У межі до чітких множин повинні бути застосовними всі визначення, докази і теореми нечітких множин, після того як нечіткість наближається до нуля, і нечітку множина стає чіткою. Таким чином, теорія нечітких множин має більш широкий спектр додатків, і зокрема, дозволяє справлятися з ситуаціями, в яких повинні враховуватися суб'єктивні думки. Основний задум, що лежить в основі теорії нечітких множин, полягає в можливості визначати нечіткі концепції реального світу, не вимагаючи визначення жорсткої бінарної межі.

Традиційний спосіб представлення інформації про те, які об'єкти є елементами множини, полягає у використанні характеристичної функції, іноді званої відрізняючою функцією. Якщо деякий об'єкт є елементом множини, то його характеристична функція дорівнює 1, а якщо об'єкт не є елементом множини, то - 0:

$$\mu_A(x) = \begin{cases} 1, & \text{якщо } x \text{ - елемент множини } A; \\ 0, & \text{якщо } x \text{ не є елементом множини } A. \end{cases}$$

У цьому визначенні виражається класичне поняття множини, згідно з яким деякий об'єкт або належить, або не належить до множини. Множини, до яких застосовується це поняття, називаються чіткими, на відміну від нечітких множин. Зазначений підхід виходить з аристотелівських уявлень про двозначну логіку, в якій можливі тільки справжні і несправжні значення.

На відміну від класичної логіки, в якій робляться спроби класифікувати всю інформацію за допомогою бінарних шаблонів, таких як в нечіткій логіці приділяється увага "забороненому третього" і робиться спроба врахувати

наявність ситуацій, в яких доводиться стикатися з частково істинними і частково помилковими твердженнями, що лежать в основі більшої частини міркувань людини в реальному житті.

Основний недолік двозначної логіки обумовлений тим, що люди живуть в неперервному, а не в дискретному світі. У реальній дійсності будь-які предмети зазвичай не знаходяться лише в одному певному стані.

Особливістю нечітких множин є те, що будь-який об'єкт може належати до множини лише частково. Ступінь приналежності до нечіткої множини вимірюється за допомогою функції приналежності, або функції сумісності, яка визначається наступним чином:

$$\mu_A(x) : X \rightarrow [0,1]$$

Це визначення схоже з визначенням характеристичної функції, але фактично значно від нього відрізняється. Характеристична функція відображає всі елементи універсуму X на один з елементів множини. На відміну від цього функція приналежності відображає універсум X на область значень дійсних чисел, певну в інтервалі від 0 до 1 включно. Таким чином, функція приналежності є дійсне число, і в разі рівності 0 свідчить про неналежність до множини, при 1 ж позначає повну приналежність до множини. Конкретне значення функції приналежності, наприклад 0.5, називають ступенем приналежності.

Нечітке істинне значення називається нечітким специфікатором і може використовуватися в якості нечіткої множини або модифікувати його. На відміну від чітких висловлювань, нечіткі твердження можуть мати квантори, такі як "більшість", "багато", "звичайно" і т.д., тому по відношенню до них не проводяться такі відмінності, як в класичній логіці, тобто відмінності між твердженнями і висловлюваннями.

Також існує уявлення про функції приналежності як про функції сумісності. Термін сумісність показує, наскільки повно деякий об'єкт узгоджується з певним атрибутом, тому дійсно дозволяє ясніше описувати нечіткі множини. Елементи нечітких множин часто представляють за

допомогою трійки "об'єкт-атрибут-значення", тоді як стосовно чітким множинам можуть розглядатися тільки об'єкт і атрибут, оскільки передбачається, що значення дорівнює або 0, або 1.

При описі нечітких множин в якості функції належності часто використовується математична функція, яка називається S-функцією. Ця функція визначається наступним чином:

$$S(x; \alpha, \beta, \gamma) = \begin{cases} 0 & \text{для } x \leq \alpha; \\ 2\left(\frac{x - \alpha}{\gamma - \alpha}\right)^2 & \text{для } \alpha \leq x \leq \beta; \\ 1 - 2\left(\frac{x - \gamma}{\gamma - \alpha}\right)^2 & \text{для } \beta \leq x \leq \gamma; \\ 1 & \text{для } x \geq \gamma. \end{cases}$$

S-функція показала себе цінним інструментальним засобом при визначенні нечітких функцій, і дозволила відмовитися від супроводу таблиці з даними, що визначають функцію приналежності, і компактно представити всі дані за допомогою однієї формули. У наведеному вище визначенні символи α , β , і γ являють собою параметри, які можуть бути відкоректовані для досягнення відповідності з даними, що описують приналежність до множини. Приклад графіка S-функції показаний на рис. 2.3.

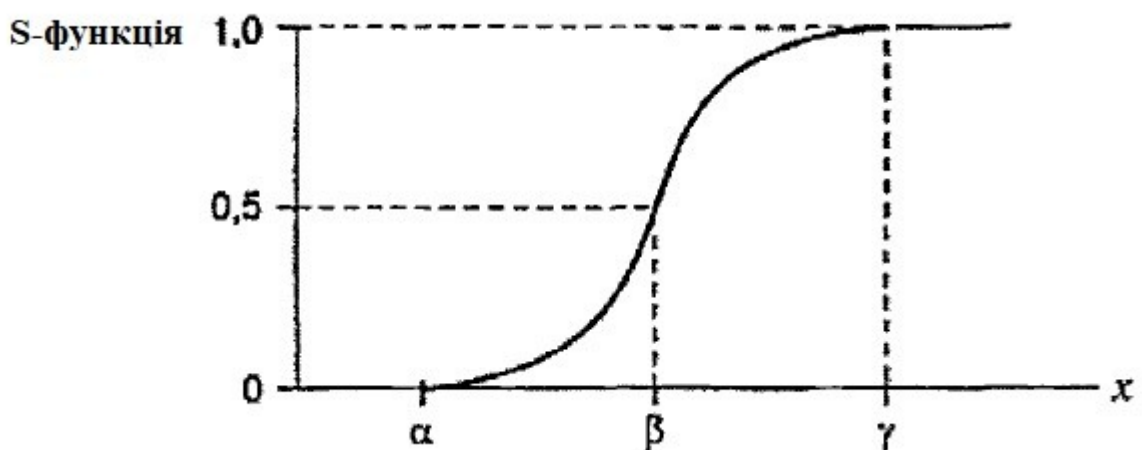


Рис. 2.3. S-функція приналежності

Нижче приведена функція, яка також дозволяє отримати подібну криву, але досягає нуля в зазначених точках.

$$P(x; \beta, \gamma) = \begin{cases} S(x; \gamma - \beta, \gamma - \beta/2, \gamma) & \text{для } x \leq \gamma \\ 1 - S(x; \gamma, \gamma + \beta/2, \gamma + \beta) & \text{для } x \geq \gamma \end{cases}$$

Графік P -функції показаний на рис. 2.4, параметр β в точках перетину визначає ширину інтервалу.

Функція приналежності не тільки може виражатися неперервною функцією, але і представляти собою кінцеву множину елементів, наприклад:

$$U = \{5, 5.5, 6, 6.5, 7, 7.5, 8\}$$

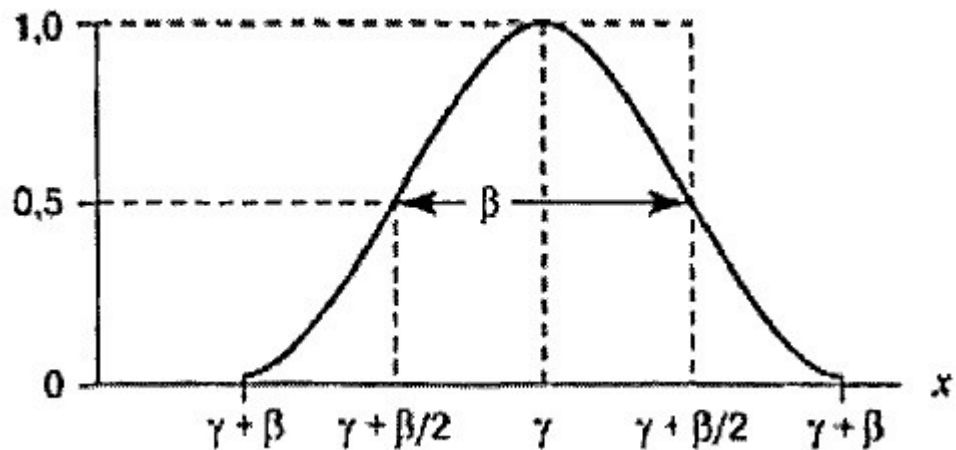


Рис. 2.4. P -функція приналежності

Нечіткі множини дозволяють моделювати ще одну важливу концепцію - відношення, тобто певні асоціації між елементами.

Нечітке відношення для N множин визначається як розширення чіткого відношення, яке включає позначення ступенів належності. Таким чином, має місце наступне співвідношення, яке визначає асоціацію зі ступенем приналежності кожного N -елементного кортежу

$$R = \{\mu_R(x_1, x_2, \dots, x_N) / (x_1, x_2, \dots, x_N) \mid x_i \in X_i, i = 1, \dots, N\}$$

Для кількісної оцінки сенсу пропозицій природної мови застосовуються нечіткі множини і лінгвістичні змінні, після чого з'являється можливість маніпулювати цими пропозиціями. Лінгвістичним змінним присвоюються значення, що представляють собою такі вирази, як слова, фрази або пропозиції природної або штучної мови.

Лінгвістичні змінні звичайно використовуються в евристичних правилах. Але значення змінних можуть встановлюватися за допомогою

логічного виведення на базі заданих правил. Множиною термів $T(L)$ лінгвістичної змінної L називається множина значень, які може приймати ця змінна. Такі множини можуть бути об'єднанням інших множин, що складаються з підмножин.

В ході пошуку альтернатив класичній логіці, поряд з нечіткою логікою був сформульований цілий ряд різних логічних теорій, які були засновані на множинних значеннях істинності; в їх число входять теорії Лукашевича, Бочвара, Хейтінга і Рейхенбаха. В N -значній логіці передбачається, що множина T_N істинісних значень розподілено рівномірно за замкненим інтервалом $[0,1]$:

$$T_N = \left\{ \frac{i}{N-1} \right\} \text{ для } 0 \leq i < N$$

В якості прикладу можна навести такі визначення:

$$T_2 = \{0,1\} \quad T_3 = \left\{ 0, \frac{1}{2}, 1 \right\}$$

Нечітка логіка може розглядатися як розширення багатозначної логіки. Але призначення і область застосування нечіткої логіки є іншими, оскільки нечітка логіка - це логіка наближених міркувань, а не точних багатозначних міркувань. Теорія нечітких множин в цілому є найбільш загальною з усіх сформульованих теорій невизначеності і має дуже широку сферу застосування.

На основі проведеного аналізу та з урахуванням специфіки складноформалізованих процесів функціонування груп контрольованих пристроїв, для підвищення описових властивостей мережевої моделі залежностей має сенс використовувати при її побудові засоби нечітко-логічного і непевного підходів шляхом додавання відповідних атрибутів кожному з вузлів моделі, наприклад, введення нечіткої лінгвістичної змінної дозволяє оцінити тривалість перебування пристрою в даному (відповідному елементу моделі) стані, а за допомогою таблиці умовних ймовірностей можна оцінити ймовірність переходу пристрою в одне з пов'язаних з поточним станів.

2.4. Висновки до розділу 2

1. На основі виконаного аналізу існуючих підходів до подання знань в інтелектуальних системах, і з урахуванням вимог, що пред'являються до моделі знань про залежності між діагностичними станами, зроблено висновок про доцільність застосування мережевого підходу до подання знань.

2. Розгляд методів інтелектуального аналізу даних з урахуванням основних задач діагностики технічних пристроїв дозволило визначити групу методів для вирішення задачі виявлення залежностей, запропонувати спосіб поліпшення характеристик даної групи і сформулювати загальну ідею комплексного методу асоціативного аналізу даних.

3. Проведений аналіз методів формування міркувань на основі неповних даних з урахуванням специфіки складноформалізованих процесів функціонування пристроїв АТ показав корисність застосування засобів імовірнісного і нечітко-логічного підходів на етапі використання виявлених залежностей для підтримки прийняття рішень персоналу центру діагностування та моніторингу, і як наслідок - необхідність доповнення моделі залежностей діагностичних станів відповідними атрибутами.

РОЗДІЛ 3

РОЗРОБКА МОДЕЛЕЙ ПРЕДСТАВЛЕННЯ ЗНАНЬ, МЕТОДІВ І АЛГОРИТМІВ АНАЛІЗУ ДІАГНОСТИЧНИХ ДАНИХ

Аналіз поточного стану та перспектив розвитку систем технічного діагностування, що застосовуються на транспорті, виконаний в Розділі 1, дозволив виявити актуальність задач інтелектуалізації, реалізації експертних функцій і підтримки прийняття рішень оперативного персоналу для запобігання відмов технічних пристроїв. Одним з етапів вирішення цього завдання є побудова моделей залежностей між діагностичними станами, тобто витяг знань про залежності зі сховищ діагностичних даних.

З урахуванням великого обсягу сховищ, а також постійного зростання інтенсивності надходження даних, викликаного збільшенням кількості об'єктів, оснащених програмно-апаратними засобами систем діагностування, очевидна необхідність автоматизації процесу обробки даних і отримання знань. Таким чином, виникає задача розробки системи аналізу даних, що володіє властивістю «інтелектуальності», необхідним для виявлення стійких залежностей між діагностичними станами, не формалізованих раніше експертами.

На основі виконаного в Розділі 2 аналізу існуючих підходів до подання знань в інтелектуальних системах і вимог, що пред'являються до моделі знань про залежності між діагностичними станами, а також з урахуванням специфіки складноформалізованих процесів функціонування пристроїв АТ, було прийнято рішення про розробку моделі залежностей на основі мережевого підходу з використанням елементів теорії ймовірності та нечіткої логіки.

Кафедра АЕМ				НАУ 21 03 86 000 ПЗ			
Виконав	Кузюра А.А.			РОЗРОБКА МОДЕЛЕЙ ПРЕДСТАВЛЕННЯ ЗНАНЬ, МЕТОДІВ І АЛГОРИТМІВ АНАЛІЗУ ДІАГНОСТИЧНИХ ДАНИХ	Літера	Аркуш	Аркушів
Керівник	Товкач С.С.					1	45
Консульт.					АТ-412Б 154Б		
Н-контр.	Товкач С.С.						
Зав. каф.	Захарченко В.П.						

Розділ присвячений розробці моделей подання знань, методів і алгоритмів синтезу моделей за вибіркою діагностичних даних, а також їх застосування в системі інтелектуального аналізу даних.

У підрозділі 3.1 описується мережева модель представлення імовірнісних залежностей між діагностичними станами.

Під діагностичним станом (ДС) будемо розуміти технічний стан контрольованого об'єкта, що формується системою технічного діагностування на основі значень діагностичних параметрів, і характеризує працездатність пристроїв АТ або проведення спеціалізованих технологічних операцій. Для кожного типу об'єктів заданий перелік всіх можливих діагностичних станів [].

Пункт 3.2 присвячений розробці методу і алгоритмів виявлення залежностей у вибірці даних на основі принципу мінімальної довжини опису. У підрозділі 3.3 описується нечітко-темпоральна модель представлення залежностей. Пункт 3.4 присвячений розробці методу і алгоритму виявлення нечітко-темопральних послідовностей.

3.1 Мережева модель представлення залежностей

Одним з важливих напрямків досліджень в галузі ІІІ є теоретичне вивчення властивостей моделей представлення даних і знань з невизначеністю. Серед них виділяється клас моделей, що включають три компоненти:

- систему змінних (показників, параметрів та ін.);
- систему обмежень що накладаються на значення змінних аксіоматикою відповідних математичних об'єктів;
- систему обмежень, що накладаються на значення змінних конкретними особливостями предметної області або об'єктів з неї.

Крім того, сукупність змінних розбивається (декомпозирується) на невеликі набори, що володіють деякою регулярною структурою, в кожному з

яких елементи і зв'язки між ними можна охарактеризувати найбільш точно. Набори можуть перетинатися, вони також організуються в певну регулярну структуру.

Наведемо ряд положень, узагальнюючих зазначені вище характеристики класу і релевантні результати досліджень в області подання та обробки знань з невизначеністю, якому повинна відповідати така модель представлення знань про предметну область.

Знання про предметну область представляються у вигляді сукупності змінних. Ці змінні зіставляються або твердженнями про предметну область, або параметрами і характеристикам об'єктів предметної області. Кожна змінна має свою множину допустимих значень (як правило, множина допустимих значень - це одна з шкал, дискретна або неперервна, з деяким набором допустимих операцій). Найпростіші приклади множин допустимих значень: одиничний інтервал, бінарна множина з елементами «Істина» і «Брехня», і т.д.

Окремим змінним, а також математичним виразами, побудованим на їх основі, виходячи з доступних відомостей про предметну область, приписується або точкове значення, або множина значень (більш вузьке, ніж множина допустимих значень). набір математичних виразів, в які входять змінні, залежить від області допустимих значень змінних, набору допустимих операцій для даної шкали і специфіки предметної області.

На значення змінних накладаються також обмеження, що впливають з визначення або властивостей розподілу значень зазначених змінних. Виходячи з висновків, зроблених фахівцями в галузі когнітивної психології і штучного інтелекту [], набір змінних можна розбити на невеликі фрагменти, побудовані на змінних, зв'язку яких експерту найпростіше описати, або виконати розбиття на фрагменти виходячи з особливостей предметної області. Фрагментів може бути кілька, змінні можуть входити в різні фрагменти.

Також слід взяти до уваги технологічне обмеження, яке, як правило, явно чи неявно присутнє при спробі перейти до обчислювальних експериментів. Важливо врахувати реалізацію операцій над отриманим поданням невизначеності знань в рамках розвинених або розвиваючих методів математики та інформатики, наприклад, лінійного програмування, теорії графів, теорії реляційних баз даних. і ін. Зазначені теоретичні методи повинні бути підтримані відповідними інформаційними технологіями, наприклад, деякою СУБД, програмними бібліотеками функцій і т.д.

Зазначені положення знайшли своє відображення при побудові, теоретичному дослідженні і програмній реалізації таких моделей знань з ймовірнісною невизначеністю, як байєсовські мережі [].

Байєсовську мережу (БС) можна визначити як двійку $\langle G, B \rangle$, в якій перший компонент G є спрямованим ациклічним графом, вершини, якого відповідають можливим, подіям досліджуваної предметної області, а ребра - причинно-наслідковим зв'язкам між цими подіями.

Другий, компонент двійки - B , являє собою множину параметрів; що визначають мережу - набір таблиць умовних, ймовірностей. Він містить параметри $\Theta_{x^i|pa}(X^i) = P(x^i | p_a(X^i))$ для кожного можливого значення x^i із X^i і $p_a(X^i)$ із $P_a(X^i)$, де $P_a(X^i)$ позначає, набір батьків змінної X^i в G . Кожна змінна X^i в графі G представляється у вигляді вершини.

Повна спільна ймовірність БС обчислюється за формулою:

$$P_{BC}(X^1, \dots, X^N) = \prod_{i=1}^N P_B(X^i | P_a(X^i))$$

З математичної точки зору БС - це модель представлення імовірнісних залежностей, а також відсутність цих залежностей. При цьому зв'язок, $A \rightarrow B$ є причинного, коли подія A є причиною виникнення B , тобто, коли є механізм, у відповідності з яким значення, прийняте A , впливає на значення, прийняте B . БС називають причинною (каузальною), коли всі її зв'язки є причинними.

Графічно структура БС може бути, представлена у вигляді дерева, полідерева, решітки, або їх комбінації. Дерево - така структура БС, в якій будь-яка вершина може мати не більше однієї вершини-предка (рис. 3.1, а). Полідерево - структура БС, в якій будь-яка вершина може мати більш ніж одну вершину-предка, але при цьому між будь-якими двома вершинами має бути не більше одного зв'язуючого їх шляху (рис. 3.1, б). Решітка - мережна структура, в якій будь-яка вершина може мати більш ніж одну вершину-предка; при цьому між будь-якими двома вершинами може бути більш одного зв'язуючого їх шляху (рис. 3.1, в).

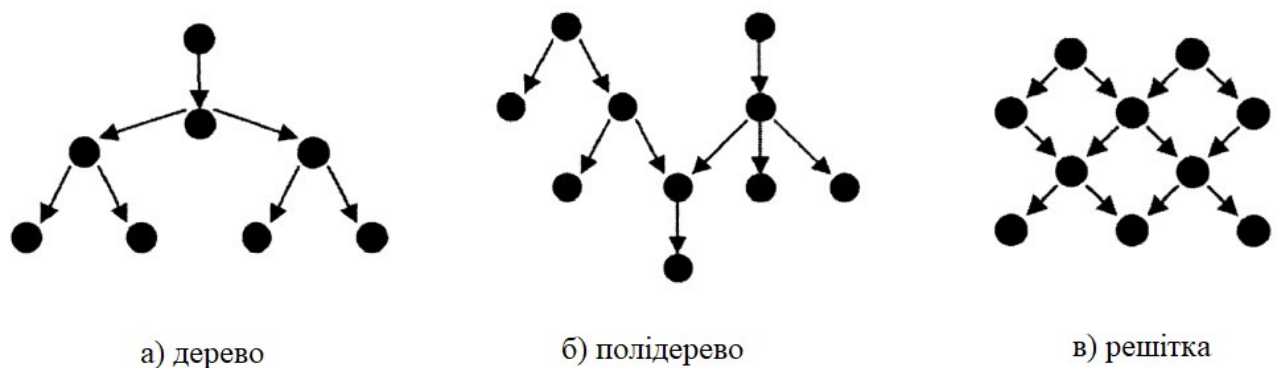


Рис 3.1 Приклади різних структур БС

Таким чином, базу знань про предметну область можна умовно уявити як комбінацію фрагментів, кожен з яких містить модель представлення імовірнісних залежностей.

Формально фрагмент бази знань являє собою набір змінних, і обмеження на їх значення. При цьому передбачається можливість досить повно охарактеризувати значення змінних і зв'язку між ними. У байесовських мережах в якості атрибута фрагмента знань можна розглядати тензор умовних ймовірностей, приписаний до конкретного вузлу мережі.

Вузлу мережі зіставляється пропозиціональна формула x . Визначено, що при відомих істиннісних означувань пропозиціональних формул y_1, \dots, y_m у вузлах, які є безпосередніми попередниками розглянутого, ймовірність істинності x представляється тензором умовних ймовірностей:

$$p(x | y_1 \cap \dots \cap y_m) = p(x | y_1 \dots y_m)$$

Всіх вузлів, що входять в Байєсову мережу, зіставлений подібний тензор умовних ймовірностей. Для кожного істинного означування вузлів-попередників дається точкова оцінка умовної ймовірності істинності пропозиціональної формули x .

Таким чином, кожному істиннісному означуванню зіставляється змінна, що представляє відповідну умовну ймовірність. Значення всіх таких змінних лежать в інтервалі $[0; 1]$.

Слід зазначити деякі особливості байєсовских мереж:

- не допускається інтервальна оцінка умовної ймовірності;
- має місце несиметричність зв'язків між вузлами, тому що попередники вузла вважаються його причиною, а сам вузол - їх наслідком. Для деяких предметних областей застосування причинно-наслідкових зв'язків в поданні знань може бути некоректним;

- при виникненні ненапрямлених циклів в мережі застосовуються тензори спільних ймовірностей;

- використання спрямованих циклів (циклів зворотного зв'язку) не підтримується [].

Фрагмент знань БС представляє сімейство розподілів ймовірностей, проте з урахуванням прийнятої в БС гіпотезі про умовну незалежність d -розділимих вузлів, в сукупності фрагменти знань байєсівські мережі довіри задають єдиний розподіл.

Припущення, що лежить в основі побудови ймовірнісної моделі, полягає в тому, що, в той час як загальна проблема надто складна для застосування звичайних методів обчислення і поновлення ймовірностей, окремі кліки мережі мають порівняно невеликий розмір. Зокрема, припущення полягає в тому, що можливо виконання будь-яких бажаних операцій, таких як маргіналізація або нормування, всередині будь-якої кліки, хоча необов'язково для всієї мережі. В результаті подібну структуру мережі можна використовувати для того, щоб поширити такі обчислення на весь набір змінних.

Як приклад ймовірнісної моделі, на рис. 3.2 приведена структура ФЗ для приводу []. Структура залежностей між станами, а також апріорні і умовні ймовірності отримані в результаті опитування групи експертів в області АТ. Нижче перераховані уніфіковані діагностичні стани стрілочних приводів, які стосуються збоїв.

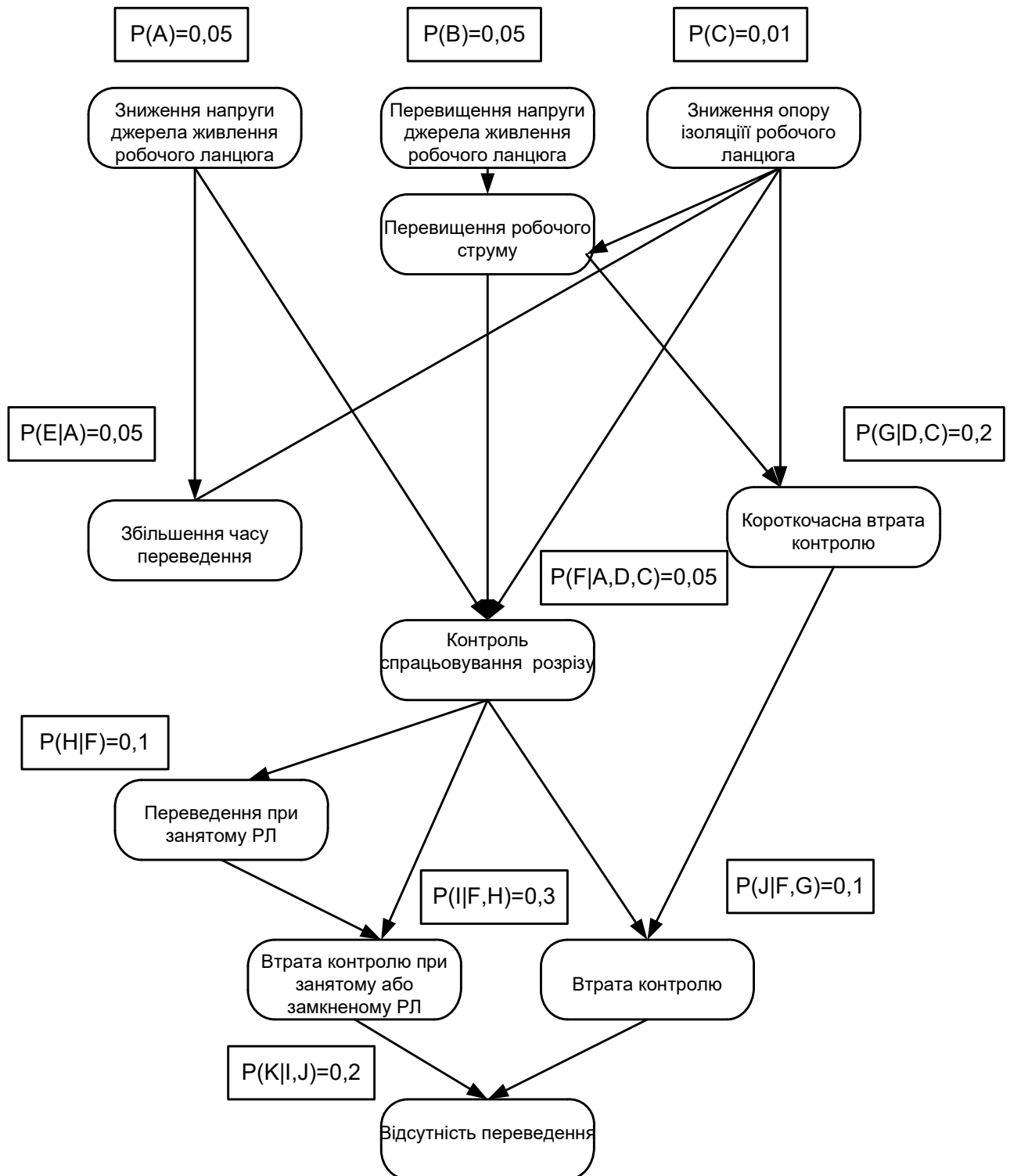


Рис. 3.2 Структура фрагменту знань для уніфікованого типу пристрою

З станів, наведених на рис.3.2, до передвідмов приводів відносяться: короткочасна втрата контролю, збільшений час переведення, контроль спрацьовування сигналу розрізу, зниження опору ізоляції робочої ланцюга, заниження напруги джерела живлення робочого ланцюга, завищення напруги джерела живлення робочого ланцюга, завищення робочого струму переведення. До відмов відносяться: відсутність переведення, втрата контролю, втрата контролю при зайнятому або замкненому РЛ, переведення при зайнятому РЛ.

3.2. Метод і алгоритми синтезу мережевих моделей залежностей

Автоматична побудова моделей навколишнього середовища (тобто вилучення знань з поступаючих ззовні даних) є однією з основних категорій задач ІІІ. При вирішенні таких задач передбачається, що прояви «навколишнього світу» фактично надані вирішальному алгоритму як якийсь набір вхідних даних (наприклад, набір показань сенсорів в послідовні моменти часу), і завданням алгоритму є побудова моделі, що адекватно описує цей набір даних.

У Розділі 2 була приведена загальна ідея комплексного методу асоціативного аналізу даних діагностування технічних пристроїв, що дозволяє виявити залежності між проявами різних діагностичних станів і представити їх в зручній і наочній формі.

В основі запропонованого методу лежить уявлення витягнутих з даних знань у вигляді мережної моделі, структура якої описана вище в розділі 3.1.

Комплексний метод асоціативного аналізу даних і побудови мережевої моделі залежностей між подіями складається з наступних основних етапів:

- вибір часового інтервалу для аналізу;
- завантаження діагностичних даних;
- вибір періоду дискретизації часового інтервалу;

- перетворення даних з одновимірного виду (символьний часовий ряд, кожен символ якого представляє подію деякого типу) до матриці, що складається з ознак $E_{i,j}$ прояви аналізованих подій j -го типу за i -й період дискретизації;

- алгоритмічна обробка матриці даних, що дозволяє на основі заданого критерію визначити структуру мережевої моделі, відповідну діагностичним даними.

Одним з основних елементів методу є критерій оцінювання відповідності структури моделі діагностичним даними. Вибору даного критерію і присвячені наступні пункти даного розділу.

3.2.1 Синтез моделі залежностей на основі критерію мінімальної довжини опису

Можна говорити, що побудована модель адекватно описує потік даних, якщо вона не зустрічає суперечливих їй фактів, тобто, якщо для кожного нового зрізу потоку в даній моделі міг бути отриманий (при заданні належних вільних параметрів моделі) результат, "досить схожий" на спостережуваний.

Умова відсутності протиріч моделі у вхідному потоці даних є необхідною ознакою адекватної моделі, але не є достатнім, адже найпростішою "моделлю", що задовольняє такій умові, буде модель "можливо все", в якій невизначеність, створювана невідомими параметрами, дуже велика і дозволяє узгодити модель не тільки з реальними зрізами потоку, але взагалі з будь-якими наборами величин. Тому природно сформулювати додаткову умову мінімальності невизначеності, що вноситься вільними параметрами моделі. Ця умова, по суті, є формалізацією принципу бритви Оккама: "не слід вводити сутностей понад необхідних".

В цілому, принцип бритви Оккама неявно включений в усі методи вибору моделі. Для формальної реалізації бритви Оккама необхідно оцінити

"економічність" моделі по відношенню до наявних даних. Оскільки вибір моделі можна інтерпретувати як побудова опису спостережуваних даних, то в якості такої оцінки вибирається довжина отриманого опису. Таким чином,

одним із загальних принципів вибору моделі для деякого набору даних є принцип мінімальної довжини опису (МДО), який формулюється так: слід вибрати ту модель, яка дозволяє описати дані найбільш коротко []. Правила побудови опису (і оцінка його довжини) і будуть формалізацією поняття "адекватності опису".

Відповідно до теорії кодування К. Шеннона, при відомому розподілі $P(X)$ випадкової величини X довжина оптимального коду для передачі конкретного значення x по каналу зв'язку прагне до $L(x) = -\log P(x)$. Ентропія джерела $S(P) = -\sum_x P(x) \cdot \log P(x)$ є мінімальною очікуваною довжиною закодованого повідомлення. Будь-який інший код, заснований на неправильному уявленні про джерело повідомлень, призведе до більшої очікуваної довжини повідомлення. Іншими словами, чим краще модель джерела, тим компактніше можуть бути закодовані дані.

У задачі навчання джерелом даних є якась невідома нам справжня функція розподілу $P(D|h_0)$, де $D = \{d_1, \dots, d_N\}$ - набір даних, h - гіпотеза імовірного походження даних, $L(D|h) = -\log P(D|h)$ - емпіричний ризик, адитивний по числу спостережень і пропорційний емпіричній помилці.

Відмінність між $P(D|h_0)$ і модельним розподілом $P(D|h)$ у $[p \setminus h]$ за мірою Кульбака-Лейблера визначається як

$$|P(D|h) - P(D|h_0)| = \sum_D P(D|h_0) \cdot \log \frac{P(D|h_0)}{P(D|h)} = \sum_D P(D|h_0) |L(D|h_0) - L(D|h)| \geq 0,$$

тобто являє собою різницю очікуваної довжини кодування даних за допомогою гіпотези і мінімально можливої. Ця різниця завжди невід'ємна і дорівнює нулю лише при повному збігу двох розподілів.

У загальному вигляді задача визначення МДО виглядає наступним чином. Спочатку задається множина навчальних даних $D = \{d_1, \dots, d_n\}$,

$d_i = \{x_i^{(1)}x_i^{(2)} \dots x_i^{(N)}\}$, (нижній індекс - номер спостереження, а верхній - номер змінної), n - кількість спостережень, кожне спостереження складається з $N(N \geq 2)$ змінних $X^{(1)}, X^{(2)}, \dots, X^{(N)}$, кожна j -я змінна ($j = 1, \dots, N$) має $A^{(j)} = \{0, 1, \dots, a^{(j)} - 1\}$ ($a^{(j)} \geq 2$) станів, кожна структура $g \in G$ БС представляється N множинами предків $(\prod^{(1)}, \dots, \prod^{(N)})$, тобто для кожної вершини $1, \dots, N$, $\prod^{(j)}$ - це множина батьківських вершин, така що $\prod^0 \subseteq \{X^{(1)}, \dots, X^{(N)}\} \setminus \{X^{(j)}\}$ (вершина не може бути предком самої себе, тобто петлі в графі відсутні).

Тоді МДО структури $g \in G$ при заданій послідовності з n спостережень $x^n = d_1 d_2 \dots d_n$ обчислюється за формулою

$$L(g, x^n) = H(g, x^n) + \frac{k(g)}{2} \cdot \log(n),$$

де $k(g)$ - кількість незалежних умовних ймовірностей в мережевій структурі g .

$H(g, x^n)$ - емпірична ентропія: $H(g, x^n) = \sum_{j \in J} H(j, g, x^n) k(g) = \sum_{j \in J} k(j, g)$,

де МДО j -ї вершини обчислюється за формулою

$$L(j, g, x^n) = H(j, g, x^n) + \frac{k(j, g)}{2} \cdot \log(n);$$

$k(j, g)$ - кількість незалежних умовних ймовірностей j -ї вершини

$$k(j, g) = (a^{(j)} - 1) \cdot \prod_{k \in \phi(j)} a^k,$$

де $\phi(j) \subseteq \{1, \dots, j-1, j+1, \dots, N\}$ - це множина, така що $\prod^{(j)} = \{X^k : k \in \phi(j)\}$.

Емпірична ентропія j -ої вершини обчислюється за формулою

$$H(j, g, x^n) = \sum_{s \in S(j, g)} \sum_{q \in A^{(j)}} -n[q, s, j, g] \cdot \log \frac{n[q, s, j, g]}{n[s, j, g]}, \quad (3.1)$$

де $n(s, j, g) = \sum_{i=1}^n I(\pi_i^j = s)$; $n[q, s, j, g] = \sum_{i=1}^n I(x_i = q, \pi_i^j = s)$,

де $\pi^{(j)} = \prod^{(j)}$ означає $X^{(k)} = x^{(k)}$, $\forall k \in \phi^{(1)}$, функція $I(E) = 1$, коли предикат $E = true$, в іншому випадку $I(E) = 0$.

Побудова мережевої моделі залежностей можна виконати за допомогою ітерації (exhaustive search) за множиною всіх можливих

нециклічних моделей, в процесі якої вибирається модель, найбільш адекватно відповідає навчальним даним. Дана задача є *NP*-складною, так як при повному переборі кількість всіх можливих нециклічних моделей (спрямованих нециклічних графів) згідно рекурентної формули [] дорівнює:

$$f(n) = \sum_{i=1}^n (-1)^{i+1} \cdot C_n^i \cdot 2^{i(n-i)} \cdot f(n-i), \quad (3.2)$$

де n - кількість вершин, а $f(0) = 1$.

На рис. 3.3а наведена блок-схема простого алгоритму вибору моделі, найближче, з позицій принципу МДО, що відбиває залежності в експериментальних даних, в порівнянні з іншими можливими варіантами моделі для даного числа вершин [].

На початку роботи алгоритму виконується завантаження даних з результатами спостережень, а також ініціалізація вихідного графа з числом вершин, що дорівнює кількості спостережуваних подій. Потім виконується формування множини графів, що містить всі можливі унікальні комбінації спрямованих ребер, і виключення з множини циклічних графів.

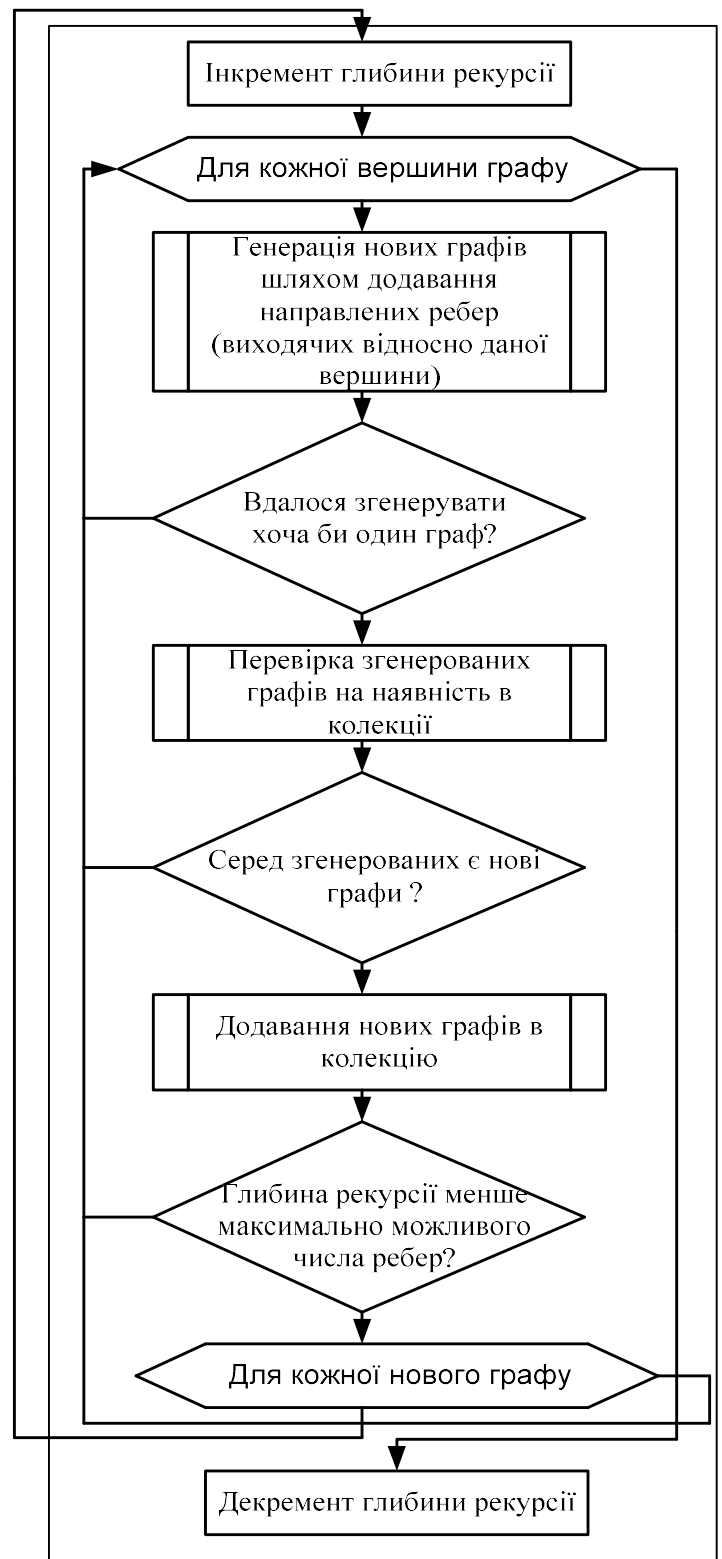
Далі, за формулою (3.1) для кожного графа обчислюється довжина кодового опису емпіричних даних, при використанні даного графа в якості моделі залежностей між подіями. За результатами обчислень визначається граф з мінімальною довжиною опису даних, який і є шуканою моделлю залежностей.

Розглянемо застосування алгоритму вибору структури моделі і принципу МДО на прикладі аналізу залежностей між уніфікованими діагностичними станами (ДС) одного з найбільш поширених пристроїв автоматики та телемеханіки - стрілочного приводу.

Нехай заданий набір даних з 30 спостережень (табл. 3.1). Кожне спостереження містить інформацію про набір поточних діагностичних станів стрілочного приводу.



а) узагальнена блок-схема алгоритму



б) блок-схема рекурсивної процедури формування множини графів

Рис. 3.3. Блок-схема простого алгоритму синтезу мережевої моделі залежностей за емпіричними даними

Таблиця 3.1 Дані про діагностику стрілкового приводу

№ п/п	$X^{(1)}$	$X^{(2)}$	$X^{(3)}$	$X^{(4)}$		№ п/п	$X^{(1)}$	$X^{(2)}$	$X^{(3)}$	$X^{(4)}$
1	1	1	0	1		16	1	0	1	1
2	1	1	0	0		17	1	0	1	1
3	1	0	0	1		18	1	0	1	1
4	1	1	0	1		19	1	0	1	1
5	1	1	0	0		20	0	1	1	1
6	1	1	0	0		21	0	1	1	1
7	1	1	0	1		22	0	1	1	1
8	1	0	0	1		23	1	0	1	1
9	1	0	0	1		24	1	1	0	1
10	0	0	1	1		25	1	0	0	1
11	0	0	1	1		26	1	1	0	0
12	0	0	1	1		27	1	0	0	1
13	0	0	1	1		28	1	0	0	0
14	1	0	1	1		29	1	0	0	0
15	1	0	1	1		30	1	1	0	0

У табл. 3.1 використані такі позначення:

$X^{(1)}$ - контроль спрацьовування сигналу розрізу стрілки (передмов);

$X^{(2)}$ - короткочасна втрата контролю (передвідмова);

$X^{(3)}$ - втрата контролю (відмова);

$X^{(4)}$ - відсутність переведення (відмова);

0/1 - відсутність/наявність ДС в період спостереження.

Розрахуємо за формулі (3.2) кількість можливих спрямованих нециклічних графів (directed acyclic graph, DAG) з числом вершин дорівнює чотирьом:

$$f(n) = \sum_{i=1}^n (-1)^{i+1} \cdot C_n^i \cdot 2^{i(n-i)} \cdot f(n-i) = \sum_{i=1}^n (-1)^{i+1} \cdot \frac{n!}{i!(n-i)!} \cdot 2^{i(n-i)} \cdot f(n-i)$$

$$f(0) = 1, f(1) = \sum_{i=1}^1 (-1)^{i+1} \cdot C_1^i \cdot 2^{i(1-i)}. f(0) = \sum_{i=1}^1 1 \cdot 1 \cdot 1 = 1$$

$$f(2) = \sum_{i=1}^2 (-1)^{i+1} \cdot C_2^i \cdot 2^{i(2-i)} \cdot f(2-i) = 1 \cdot \frac{2!}{1!(2-1)!} \cdot 2 \cdot f(1) - 1 \cdot \frac{2!}{2!(2-2)!} \cdot 1 \cdot f(0) = 4 - 1 = 3$$

$$f(3) = \sum_{i=1}^3 (-1)^{i+1} \cdot C_3^i \cdot 2^{i(3-i)} \cdot f(3-i) = 1 \cdot \frac{3!}{1!(3-1)!} \cdot 2^{1(3-1)} \cdot f(2) - 1 \cdot \frac{3!}{2!(3-2)!} \cdot 2^{2(3-2)} \cdot f(1) + 1 \cdot \frac{3!}{3!(3-3)!} \cdot 2^{3(3-3)} \cdot f(0) = 36 - 12 + 1 = 25.$$

$$f(4) = \sum_{i=1}^4 (-1)^{i+1} \cdot C_4^i \cdot 2^{i(4-i)} \cdot f(4-i) = 1 \cdot \frac{4!}{1!(4-1)!} \cdot 2^{1(4-1)} \cdot f(3) - 1 \cdot \frac{4!}{2!(4-2)!} \cdot 2^{2(4-2)} \cdot f(2) + 1 \cdot \frac{4!}{3!(4-3)!} \cdot 2^{3(4-3)} \cdot f(1) - 1 \cdot \frac{4!}{4!(4-4)!} \cdot 2^{4(4-4)} \cdot f(0) = 800 - 288 + 32 - 1 = 543.$$

Отже, в даному прикладі слід розглянути 543 можливих графів. Деякі з них наведені нижче на рис. 3.4.

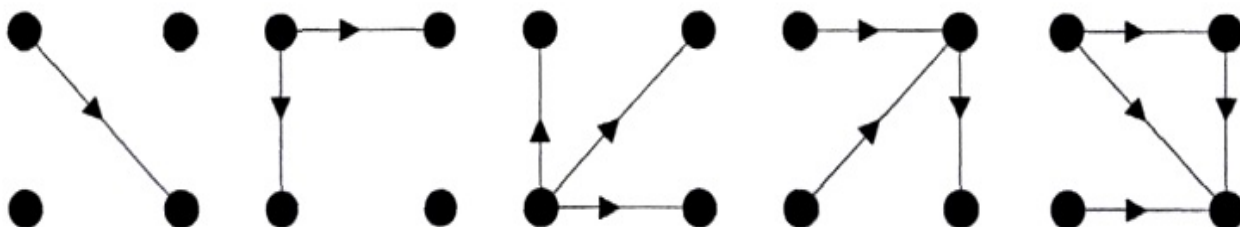


Рис. 3.4. Приклади нециклічних графів з чотирма вершинами

Для автоматизації виконання обчислень була розроблена комп'ютерна програма на об'єктно-орієнтованій мові програмування С#, що реалізує вищеописаний алгоритм.

Інтерфейс програми дозволяє перейти до групи аналізованих станів (в даному випадку - ДС стрілочного приводу), об'єкт і період часу, за який потрібно отримати дані зі сховища.

Після виконання розрахунків, (список графів впорядкований за зростанням довжини кодового опису), найменшою довжиною опису володіє

граф «X1-X3, X2-X3, X3-X4». Структура графа з назвами ДС показана на рис.3.5.

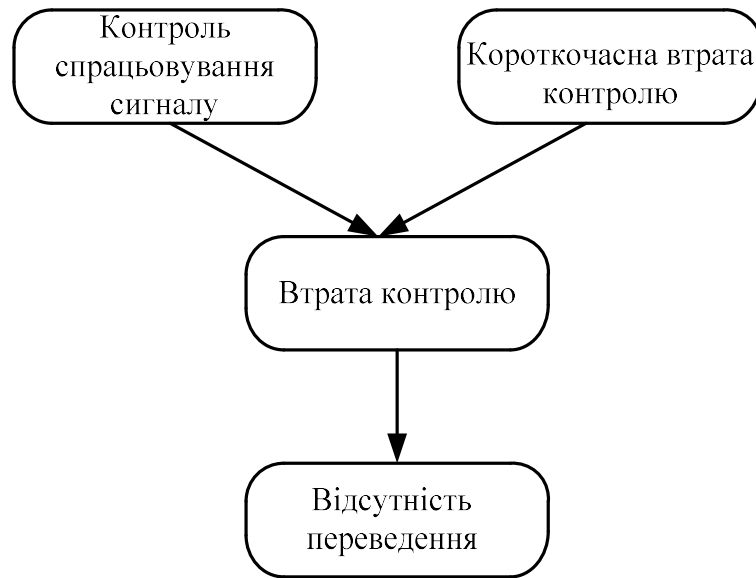


Рис. 3.5. Мережева модель залежностей діагностичних станів (табл.3.1)

Наведемо розрахунок довжини опису оптимальної структури. При обчисленні можна використовувати логарифм з будь-якою основою. В даному прикладі використовується натуральний логарифм.

Вершина $X^{(1)}$ не має предків, тобто $\Pi^{(1)} = \{ \}$. Емпірична ентропія обчислюється як $H(j=1, g) = -7 \cdot \log\left(\frac{7}{30}\right) - 23 \cdot \log\left(\frac{23}{30}\right) = 16,2982$, а кількість незалежних умовних ймовірностей $k(j=1, g) = 2 - 1 = 1$. Отже, довжина опису вершини $X^{(1)}$ дорівнює $L(1, g) = 16,2982 + \frac{1}{2} \cdot \log(30) = 17,9988$.

Таблиця 3.2 Значення параметрів вершин $X^{(1)}$ і $X^{(2)}$

$X^{(1)}$	$n[q, s, j, g]$	$n(s, j, g)$
0	7	30
1	23	

1	12	
---	----	--

$X^{(2)}$	$n[q, s, j, g]$	$n(s, j, g)$
0	18	30

Вершина $X^{(2)}$ не має предків, тобто $\prod^{(2)} = \{ \}$. Емпірична ентропія обчислюється як $H(j=2, g) = -18 \cdot \log\left(\frac{18}{30}\right) - 12 \cdot \log\left(\frac{12}{30}\right) = 20,1904$, а кількість незалежних умовних ймовірностей $k(j=2, g) = 2 - 1 = 1$. Отже, довжина опису вершини $X^{(2)}$ дорівнює $L(2, g) = 20,1904 + \frac{1}{2} \cdot \log(30) = 21,8909$.

Вершина $X^{(3)}$ має двох предків: $X^{(1)}$ і $X^{(2)}$, тобто $\prod^{(3)} = \{X^{(1)}, X^{(2)}\}$.

Емпірична ентропія обчислюється як

$$H(j=3, g) = \left(-0 \cdot \log\left(\frac{0}{4}\right) - 4 \cdot \log\left(\frac{4}{4}\right)\right) + \left(-0 \cdot \log\left(\frac{0}{3}\right) - 3 \cdot \log\left(\frac{3}{3}\right)\right) + \left(-7 \cdot \log\left(\frac{7}{14}\right) - 7 \cdot \log\left(\frac{7}{14}\right)\right) + \left(-9 \cdot \log\left(\frac{9}{9}\right) - 0 \cdot \log\left(\frac{0}{9}\right)\right) = 9,7041$$

а кількість незалежних умовних ймовірностей $k(j=3, g) = (2-1) \cdot 4 = 4$.

Отже, довжина опису вершини $X^{(3)}$ дорівнює

$$L(3, g) = 9,7041 + \frac{4}{2} \cdot \log(30) = 16,5065.$$

Таблиця 3.3 Значення параметрів вершини $X^{(3)}$

$X^{(1)}$	$X^{(2)}$	$X^{(3)}$	$n[q,s,j,g]$	$n(s,j,g)$
0	0	0	0	4
0	0	1	4	
0	1	0	0	3
0	1	1	3	
1	0	0	7	14
1	0	1	7	
1	1	0	9	9
1	1	1	0	

Вершина $X^{(4)}$ має одного предка тобто $\prod^{(4)} = \{X^{(3)}\}$. Емпірична ентропія

обчислюється як

$$H(j=4, g) = \left(7 \cdot \log\left(\frac{7}{16}\right) - 9 \cdot \log\left(\frac{9}{16}\right)\right) + \left(-0 \cdot \log\left(\frac{0}{14}\right) - 14 \cdot \log\left(\frac{14}{14}\right)\right) = 10,965,$$

а кількість незалежних умовних ймовірностей $k(j = 4, g) = (2 - 1) \cdot 2 = 2$.

Отже, довжина опису вершини $X^{(4)}$ дорівнює

$$L(4, g) = 10,965 + \frac{2}{2} \cdot \log(30) = 14,3662.$$

Таблиця 3.4 Значення параметрів вершини $X^{(4)}$

$X^{(2)}$	$X^{(3)}$	$n[q,s,j,g]$	$n(s,j,g)$
0	0	7	16
0	1	9	
1	0	0	14
1	1	14	

В результаті сумарна довжина опису графа g , представленого на рис.

3.5, дорівнює $L(g, x^n) = \sum_{j=1}^3 H(j, g, x^n) = 70,7624$.

Як було зазначено вище, задача побудови мережевої моделі залежностей за допомогою ітерації за множиною всіх можливих нециклічних моделей є *NP*-складною, тому для досягнення можливості практичного застосування алгоритму потрібно змінити його структуру і висунути якесь евристичне допущення, яке з одного боку дозволило б зменшити асимптотичну оцінку складності алгоритму, а з іншого - синтезувати оптимальну, в сенсі стислості опису, модель залежностей.

Розглянемо алгоритм виявлення залежностей на основі принципу опису мінімальної довжини з урахуванням значення взаємної інформації між подіями [].

В першу чергу для зниження складності попереднього алгоритму слід відмовитися від перебору всіх можливих мережевих структур, замість чого виконувати процедуру покрокового синтезу моделі, продовжуючи керуватися принципом МДО. Для реалізації такого підходу і виконання вимоги оптимальності результуючої моделі необхідно вибрати такий порядок аналізу залежностей (зв'язків між вузлами мережевої моделі), щоб в першу чергу розглядалися найбільш «значимі», в певному сенсі, з них.

В якості критерію значущості залежності пропонується використовувати розрахункове значення взаємної інформації між парами змінних x^i і x^j , представлених вузлами мережі.

Взаємна інформація $\{mutual\ information, MI\}$ - це статистична функція $MI(x^i, x^j)$ двох випадкових величин, що описує кількість інформації, що міститься в одній випадковій величині (x^j) відносно іншого (x^i). Взаємна інформація визначається через ентропію і умовну ентропію двох випадкових величин як:

$$MI(X, Y) = H(X) - H(X | Y).$$

Дане визначення проілюстровано на рис. 3.6, де взаємна інформація показана як загальна частина невизначеності вимірювання двох випадкових величин:

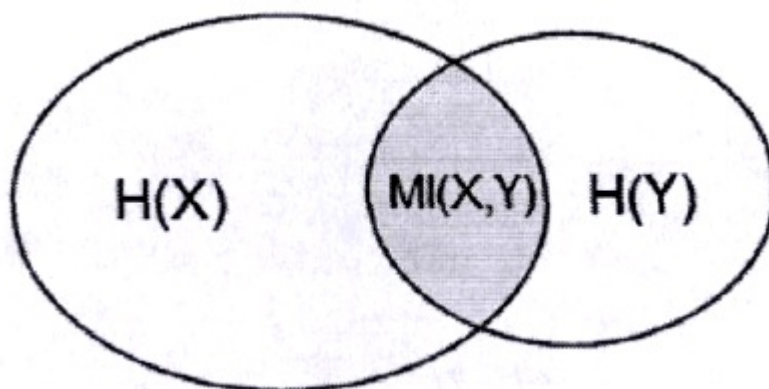


Рис. 3.6. Ілюстрація до поняття взаємної інформації

Виразимо взаємну інформації не через умовну ентропію, а через ентропію об'єднаної системи $H(X, Y)$ і окремих її частин $H(X)$, $H(Y)$:

$$\begin{aligned} H(X, Y) &= H(Y) + H(X | Y), \\ MI(X, Y) &= H(X) + H(Y) - H(X, Y) \end{aligned}$$

Далі, виразимо повну взаємну інформацію через ймовірності станів системи. Для цього запишемо значення ентропії окремих систем через математичне очікування:

$$H(X) = M[-\log P(X)], \quad H(Y) = M[-\log P(Y)], \quad H(X, Y) = M[-\log P(X, Y)].$$

Тоді $MI(X, Y) = M[-\log(P(X)) - \log(P(Y)) + \log(P(X, Y))]$, і далі:

$$MI(X, Y) = M\left[\log\left(\frac{P(X, Y)}{P(X)P(Y)}\right)\right] = \sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j) \log\left(\frac{p(x_i, y_j)}{p(x_i)p(y_j)}\right). \quad (3.3)$$

Значення взаємної інформації є невід'ємним $MI(x_i, x_j) \geq 0$, а у випадку, якщо вершини x_i і x_j повністю незалежні один від одного, то $MI(x_i, x_j) = 0$, так як $p(x_i, x_j) = p(x_i) \cdot p(x_j)$ і, отже,

$$\log\left(\frac{p(x_i, x_j)}{p(x_i) \cdot p(x_j)}\right) = \log\left(\frac{p(x_i) \cdot p(x_j)}{p(x_i) \cdot p(x_j)}\right) = \log(1) = 0.$$

Значення власної $p(x_i)$ та спільної $p(x_i, x_j)$ ймовірностей подій обчислюються за вибіркою емпіричних даних.

У випадку, якщо мережева модель складається з N вершин, для обчислення $MI(x_i, x_j)$ для всіх поєднань x_i і x_j потрібно виконати $\frac{N \cdot (N - 1)}{2}$ обчислень, тобто $MI(x_i, x_j) = MI(x_j, x_i)$.

На рис. 3.7 приведена блок-схема запропонованого евристичного алгоритму синтезу мережевої моделі залежностей за експериментальними даними з використанням принципу МДО та попередніми упорядкуванням зв'язків між подіями, представленими вузлами моделі, за значенням взаємної інформації [].

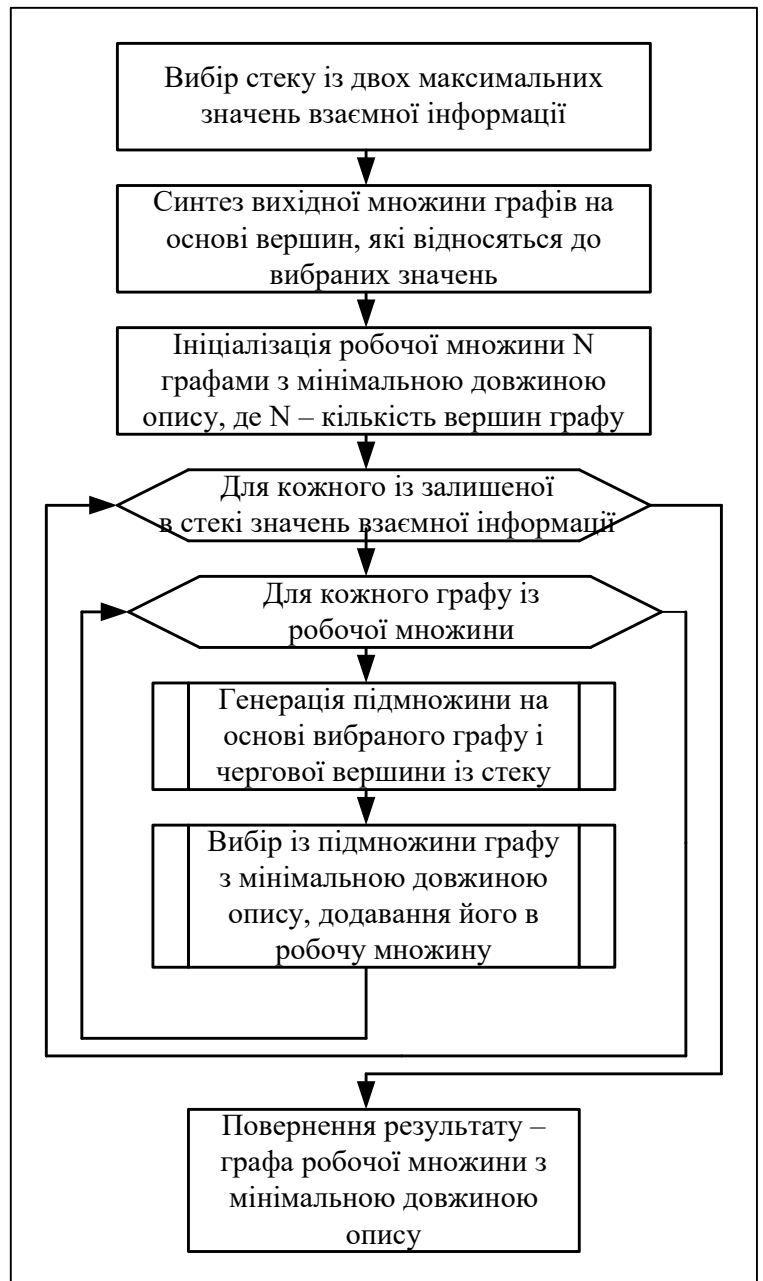
На початку роботи алгоритму виконується завантаження даних з результатами спостережень, а також ініціалізація вихідного графа з числом вершин, що дорівнює кількості спостережуваних подій. Потім виконується розрахунок власних і спільних ймовірностей спостережуваних подій, а також значень взаємної інформації між парами подій за формулою (3.3).

Розрахункові значення MI упорядковано відповідно до зменшення і поміщаються в стек S . Після цього з стека витягуються два найбільших значення MI , і на основі належних до них подій (вершин графа) будується множина графів G виду:

$$G \in \left\{ \begin{array}{l} (m_1 \rightarrow m_2; m_3 \rightarrow m_4), (m_1 \rightarrow m_2; m_3 \leftarrow m_4), (m_1 \leftarrow m_2; m_3 \leftarrow m_4), (m_1 \leftarrow m_2; m_3 \rightarrow m_4), \\ (m_1 \leftarrow m_2; m_3 \leftarrow m_4), (m_1 \rightarrow m_2; m_3 \leftarrow m_4), (m_1 \leftarrow m_2; m_3 \rightarrow m_4), (m_1 \leftarrow m_2; m_3 \leftarrow m_4), \\ (m_1 \leftarrow m_2; m_3 \leftarrow m_4) \end{array} \right\}$$



а) узагальнена блок-схема алгоритму



б) блок-схема евристичного алгоритму синтезу графа залежностей

Рис. 3.7. Блок-схеми алгоритмів синтезу моделі залежностей

Запис виду $m_i \rightarrow m_j$ означає, що вершина x^{m_i} є предком вершини x^{m_j}

"J. Запис виду $m_i \not\rightarrow m_j$ означає, що вершини x^{m_i} хщ і x^{m_j} не пов'язані.

Для кожного з графів множини G обчислюється довжина кодового опису, і N графів з мінімальною довжиною опису поміщаються в робочу множину R .

Далі, виконується подвійний цикл: зовнішній - для кожного значення MI з стека S , і внутрішній - для кожного графа R_i множини R . Графи, побудовані у внутрішньому циклі, на базі R_i і чергового значення MI , поміщаються в тимчасове підмножина графів Rt . По завершенні внутрішнього циклу множина R очищається, а елементи підмножини Rt переміщуються в R . Після завершення зовнішнього циклу множина R сортується, і визначається граф з найменшою довжиною кодового опису, який найбільш точно відображає залежності між подіями, представленими у вибірці емпіричних даних.

3.2.2 Синтез моделі залежностей на основі логарифмічної метрики

Як було зазначено вище, в процесі побудови моделі залежності необхідний критерій, що дозволяє оцінити ступінь відповідності структури моделі даними спостережень. В якості такого критерію можна використовувати оціночну метрику $K2$, запропоновану в роботі [1].

Нехай Z - множина дискретних змінних, де змінна $z_i \in Z$ має r_i можливих значень, D - набір даних, що містить m спостережень, причому кожне спостереження містить значення кожної з змінних $z_i \in Z$ (пропущених значень немає), а g - модель залежностей (бассова мережа), вершини якої відповідають змінним z_i .

Кожній вершині $X_i \in Bs$ відповідає множина батьківських вершин π_i , яке може бути порожнім. Нехай φ_{ij} визначає j -й набір означувань π_i відносно D , і існує q_i таких означувань. Визначимо N_{ijk} як число спостережень серед D , при

яких вершина X_i означувалась, як v_{ij} , а π_i - як φ_{ij} , тоді $N_{i,j} = \sum_{k=1}^{r_i} N_{i,j,k}$, а

розрахункова формула метрики:

$$P(g, x^n) = P(g) \cdot \prod_{i=1}^n \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{i,j} + r_i - 1)!} \cdot \prod_{k=1}^{r_i} N_{i,j,k}! \quad (3.4)$$

Співмножник $P(g)$ в правій частині рівняння (3.4) відображає наявну інформацію про структуру моделі, наприклад в разі, коли частина залежностей свідомо відома. Якщо такої інформації немає, даний співмножник не враховується.

Як впливає з формули (3.4), метрика $K2$ може приймати значення з інтервалу $[0; 1)$.

Недоліком такої метрики є складність обчислення і оперування факторіалами великих чисел, і як один з наслідків - обмеженість обсягу використовуваної вибірки. Для вирішення цієї проблеми введемо поняття модифікованої логарифмічної метрики $LK2$. Прологарифмуємо рівняння (3.4):

$$\begin{aligned} \log(P(g, x^n)) &= \log \left(P(g) \cdot \prod_{i=1}^n \prod_{j=1}^{q_i} \frac{(r_i - 1)! \cdot \prod_{k=1}^{r_i} N_{i,j,k}!}{(N_{i,j} + r_i - 1)!} \right) = \\ &= \log(P(g)) + \sum_{i=1}^n \left(\sum_{j=1}^{q_i} \left(\log((r_i - 1)!) - \log((N_{i,j} + r_i - 1)!) + \sum_{k=1}^{r_i} \log(N_{i,j,k}!) \right) \right) \end{aligned}$$

Отриманий вираз помножимо на -1 (щоб працювати з позитивними числами), будемо вважати апріорні ймовірності можливих структур моделі рівними ($\log(P(g)) = 0$), і замість пошуку структури з максимальним значенням функції $K2$ будемо здійснювати пошук структури з мінімальним значенням функції $LK2$:

$$F(g, x^n) = \sum_{i=1}^n \left(\sum_{j=1}^{q_i} \left(\log(r_i - 1)! - \log((N_{i,j} + r_i - 1)!) + \sum_{k=1}^{r_i} \log(N_{i,j,k}!) \right) \right) \quad (3.5)$$

Для спрощення обчислення факторіалів скористаємося одним з властивостей гамма-функції Ейлера, а саме:

$$\begin{aligned} \Gamma(n + 1) &= n! \\ \log(n!) &= \log(\Gamma(n + 1)) \end{aligned}$$

Для обчислення натурального логарифма гамма-функції існує наближення асимптотичного розкладу (формула Стірлінга):

$$\ln(\Gamma(n)) = \ln \sqrt{2\pi} + \left(n + \frac{1}{2}\right) \ln n - n + \frac{1}{12n}.$$

Отже, використовуючи у формулі (3.5) натуральні логарифми, отримаємо:

$$F(g, x^n) = \sum_{i=1}^n \left(\sum_{j=1}^{q_i} (\ln(\Gamma(r_i)) - \ln(\Gamma(N_{i,j} + r_i))) + \sum_{k=1}^{r_i} \ln(\Gamma(N_{i,j,k} + 1)) \right). \quad (3.6)$$

За результатами обчислювальних експериментів, використання функцій *K2* і *LK2* на одних і тих же вибірках навчальних даних призводить до однакових результатів, але використання *LK2* зменшує обчислювальну складність алгоритму і дозволяє працювати з великими вибірками даних.

На рис. 3.8 приведена блок-схема пропонуваного евристичного алгоритму побудови моделі залежностей між спостережуваними подіями, із застосуванням модифікованої логарифмічної метрики *LK2*, а також розрахункових значень взаємної інформації між цими подіями [].

На початку роботи алгоритму виконується завантаження даних з результатами спостережень, а також упорядкованого особливим чином множини можливих подій. Упорядкування необхідно для коректної роботи алгоритму, і може задаватися різними способами, наприклад, за допомогою експертів.

Після цього генерується граф, вершини якого представляють розглядувані події і не з'єднані ребрами. Створюється колекція *Parents*, в якій будуть зберігатися набори предків кожної з вершин.

Потім, виконується ітерація по вершинах графа, згідно заданому порядку. На початку кожної ітерації виконується ініціалізація значення метрики $LCH_{old} = 0$, і графа $G_{old} = \{V_i\}$. Для кожної вершини V_i , виділяється список *P*-вершин, що мають більший порядковий індекс події.

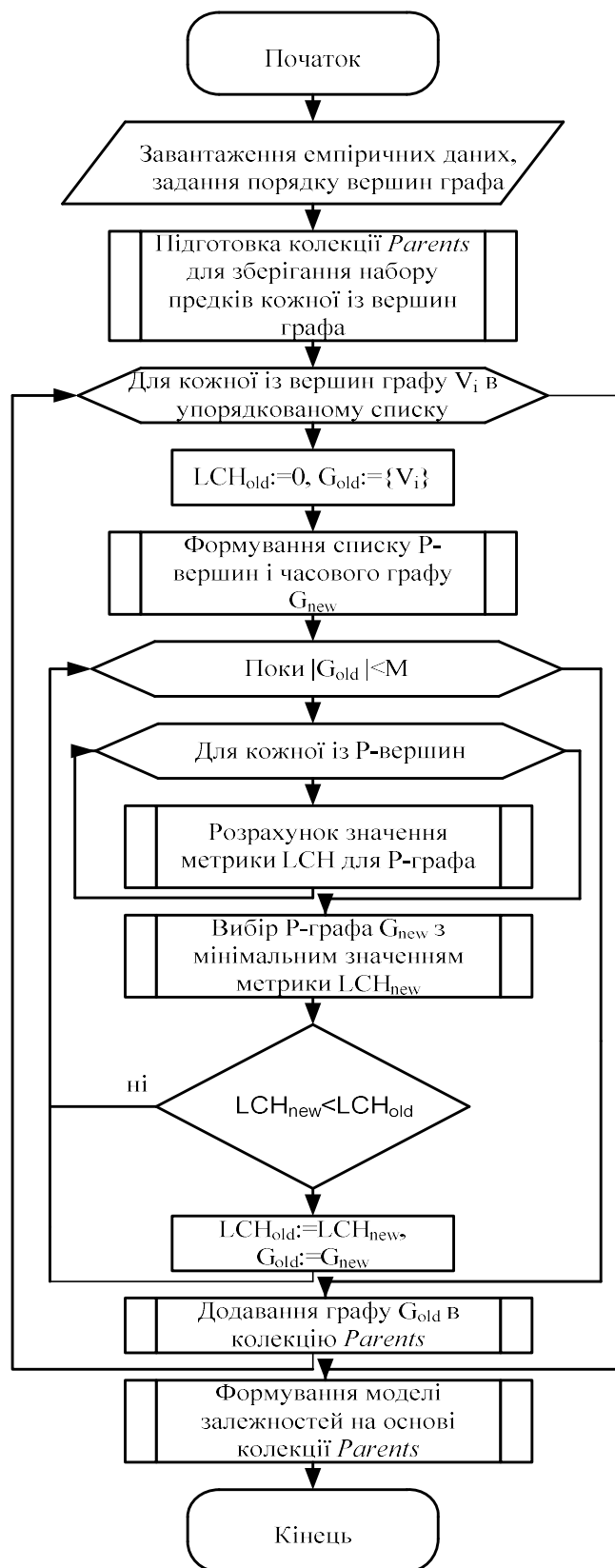


Рис. 3.8. Блок-схеми алгоритму синтезу моделі залежностей за метрикою ЛК2

Далі, виконується пошук P -вершини, додавання якої в граф G_{old} мінімізує значення модифікованої метрики LCH_{old} , розрахованої за формулою (3.6). Якщо вершина знайдена, значення метрики підміняється новим $LCH_{old} = LCH_{new}$, а граф G_{old} доповнюється найденною вершиною, і пошук триває далі, поки кількість знайдених P -вершин не перевищить максимально дозведеного значення M , або розрахункове значення метрики не стане перевищувати поточний. Якщо ж вершина не знайдена, виконується перехід до ітерації за наступною вершиною.

По завершенні ітерації на вихід алгоритму подається модель залежностей - множина подій, для кожного з яких визначено множину причинно-наслідкових зв'язків.

Розглянемо приклад застосування запропонованого алгоритму синтезу моделі залежності на прикладі аналізу вибірки даних про діагностичні стани одного з пристроїв автоматики та телемеханіки .

Нехай заданий набір даних з 30 спостережень (табл. 3.5). Кожне спостереження містить інформацію про набір поточних ДС.

Таблиця 3.5. Дані технічної діагностики

№ п/п	$X^{(1)}$	$X^{(2)}$	$X^{(3)}$	$X^{(4)}$	$X^{(5)}$
1	1	1	1	1	1
2	1	0	0	1	1
3	0	1	1	0	0
4	1	1	1	1	0
5	0	0	0	0	1
6	1	1	1	1	1
7	1	0	0	1	0
8	1	1	1	1	1
9	0	0	0	0	1
10	0	1	1	0	1
11	1	0	0	1	0
12	0	0	1	0	0
13	1	1	0	1	1
14	0	0	0	0	1
15	1	1	1	1	1
16	1	1	0	1	1
17	0	1	1	0	1
18	1	0	0	1	0

19	0	1	1	0	1
20	1	0	0	1	1
21	0	1	1	0	0
22	0	0	0	0	0
23	1	0	0	1	1
24	0	1	1	0	1
25	0	1	1	0	0
26	1	0	1	0	1
27	0	1	1	0	1
28	1	0	0	0	1
29	0	1	1	0	1
30	1	0	0	0	0

У табл. 3.5 використані наступні позначення:

$X^{(1)}$ - несправність кабелю (передвідмова);

$X^{(2)}$ - пробій ізолюючих стиків (передвідмова);

$X^{(3)}$ - відсутність кодування (передвідмова);

$X^{(4)}$ - помилкова зайнятість (відмова);

$X^{(5)}$ - відмова реле;

0/1 - відсутність / наявність ДС в період спостереження.

За результатами виконання алгоритму, отримана модель, представлена на рис.3.9. Таким чином, з вихідних даних слідує, що пробією ізолюючих стиків і відсутності кодування, як правило, передує несправність кабелю, а пробією в свою чергу може привести до помилкової зайнятості або відмови реле.

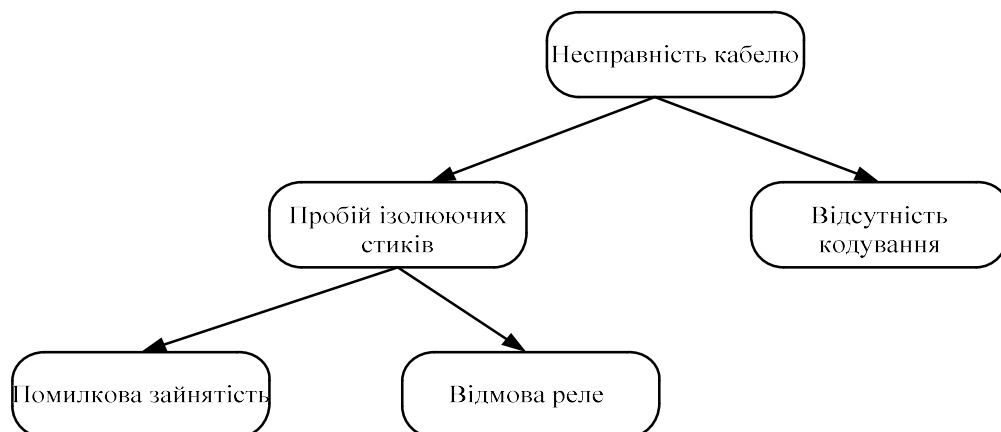


Рис. 3.9. Блок-схема алгоритму синтезу моделі залежності за модифікованою логарифмічною метриці

3.2.3 Аналіз якості роботи алгоритмів за навчальною вибіркою

Для оцінювання якості навчання моделей за наборами даних за допомогою алгоритмів, запропонованих в розділах 3.2.1 і 3.2.2, використовуємо спосіб підрахунку кількості зайвих, відсутніх і реверсованих зв'язків в навченій мережі в порівнянні з оригінальною. В якості запобіжної помилки навчання використовуємо значення структурної різниці {structure difference) між навченою і оригінальною мережевими моделями.

Для обчислення структурної різниці скористаємося формулою симетричної різниці структур []:

$$\begin{aligned}\delta &= \sum_{i=1}^n \delta_i = \sum_{i=1}^n \delta C(\Pi_i(B) \Delta \Pi_i(A)) = \\ &= \sum_{i=1}^n \delta C(\Pi_i(B) | \Pi_i(A) \cup \Pi_i(A) | \Pi_i(B))\end{aligned}\tag{3.7}$$

де B - мережева модель на виході алгоритму (навчена модель),

A - оригінальна модель,

n - кількість вершин мережі,

$\Pi_i(B)$ - множина предків i -ої вершини мережі B ,

$\Pi_i(A)$ - множина предків i -ої вершини оригінальної мережі A ,

$C(\Pi)$ - потужність (cardinality) кінцевої множини Π яка визначається як кількість елементів, що належать множині Π .

В якості оригінальної моделі використана модель залежностей між діагностичними станами приводу (рис.3.2), що містить 11 вершин. На основі даної моделі по методу Монте-Карло згенеровані вибірки навчальних даних розміром 100, 200, 500 і 1000 записів.

Далі ці вибірки подані на вхід алгоритмів МДО та К2, синтезовані моделі залежностей, підраховано кількість зайвих, відсутніх і реверсованих зв'язків в моделях і за формулою (3.7) розраховані значення структурної різниці. Результати зведені в таблицю 3.6.

Таблиця 3.6. Результати порівняння роботи алгоритмів МДО і ЛК2

Об'єм вибірки	100		200		500		1000	
	МДО	К2	МДО	К2	МДО	К2	МДО	К2
Зайві зв'язки	0	2	0	1	0	1	0	0
Відсутні зв'язки	1	1	1	1	0	0	0	0
Реверсовані зв'язки	2	1	1	1	1	1	1	1
Структурна різниця	3	4	2	3	1	2	1	1
Час виконання алгоритму, с	15	19	20	27	34	44	69	90

На основі виконаного порівняння можна зробити висновок про те, що обидва алгоритми можуть бути використані для синтезу моделей залежностей між діагностичними станами в складі системи інтелектуального аналізу даних, причому найкращою швидкістю володіє алгоритм, побудований на основі критерію МДО.

3.3 Метод формування нечітко-темпоральних описів залежностей

Моделі, розглянуті в розділах 3.1 і 3.2, дозволяють показати структуру залежностей між подіями, але не відображають таких деталей, як тривалість інтервалу, що розділяє події, кількість реалізацій події за інтервал дискретизації часу, прийнятий при трансформації протоколу діагностичних станів до матричного вигляду. В окремих випадках вимагає уточнення черговість і відносне положення подій в часі.

Для збільшення інформативності моделі пропонується доповнити її нечітко-темпоральних описом залежностей між подіями. Формування нечітко-темпоральних описів залежностей здійснюється з допомогою наступного методу.

Для опису тривалості інтервалу, що розділяє події, і кількості реалізацій події за цей інтервал, використовуємо лінгвістичні змінні «тривалість» і «частота». Базові терм-множини цих змінних мають вигляд:

Тривалість \in {мала, невелика, середня, велика}

Частота \in {одноразово, кілька, багаторазово}

Черговість і відносе положення подій в часі представимо за допомогою інтервальних відносин темпоральної логіки Аллена [], показаних на рис.3.10.

Прямі відношення

Інверсні відношення

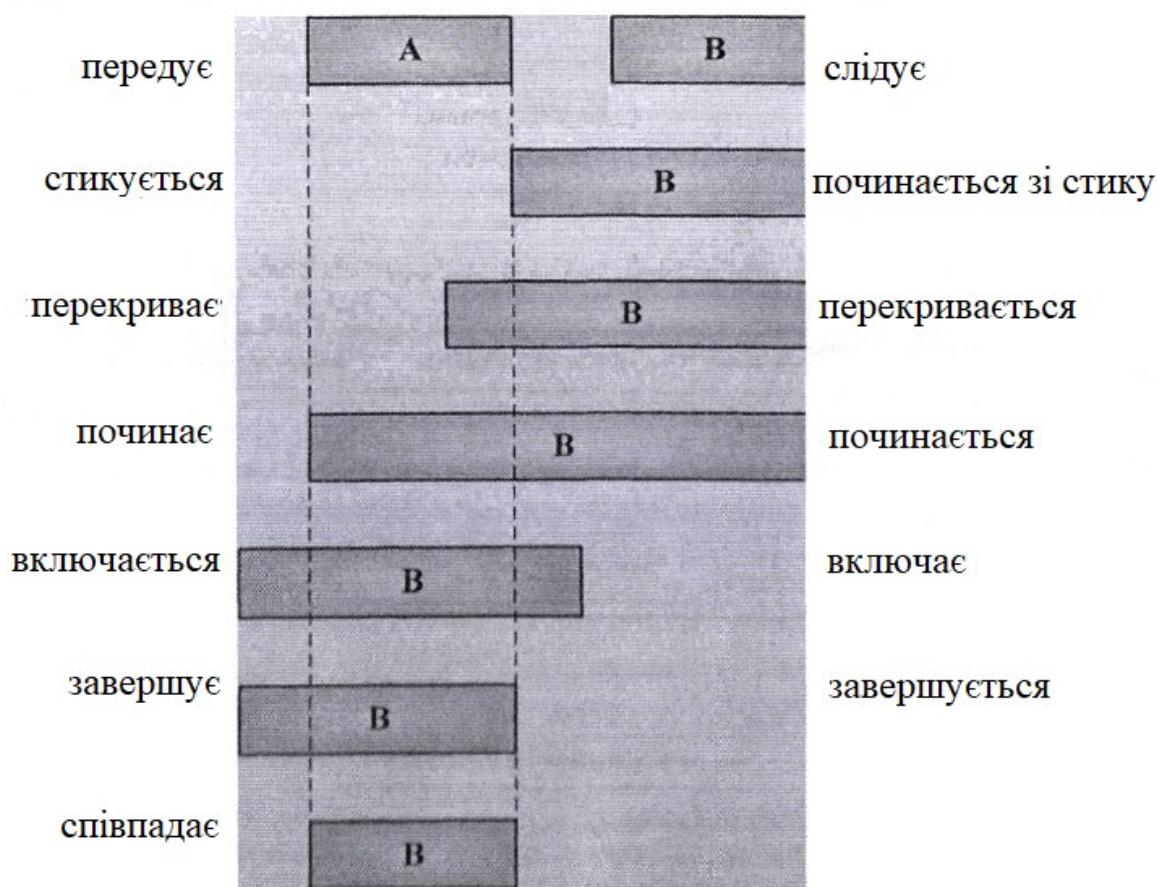


Рис 3.10. Перелік інтервальних відносин темпоральної логіки Аллена

Нечітко-темпоральний опис залежностей між подіями - це набір четвірок $TFDD = \{\text{«тривалість»} - T_i, \text{«частота першої події»} - F1_i, \text{«частота}$

другої події» - $F2_i$, «темпоральні відношення» - R_i } (*TFDD - Temporal Fuzzy Dependency's Description*), наприклад залежність між двома діагностичними станами з моделі на рис.3.9 описується так: $TFDD = \{T_i = \text{«мала»}, F1_i = \text{«багаторазово»}, F2_i = \text{«одноразово»}, R_i = \text{«перекриває»}\}$.

Розроблено алгоритм формування нечітко-темпорального опису залежності між подіями на основі фрагмента моделі, що містить пару подій, і часового ряду реалізації подій (діагностичного протоколу). Функції приналежності значень лінгвістичних змінних задані за допомогою стандартних кусочно-лінійних *Z*-, *P*- і *S*-функції.

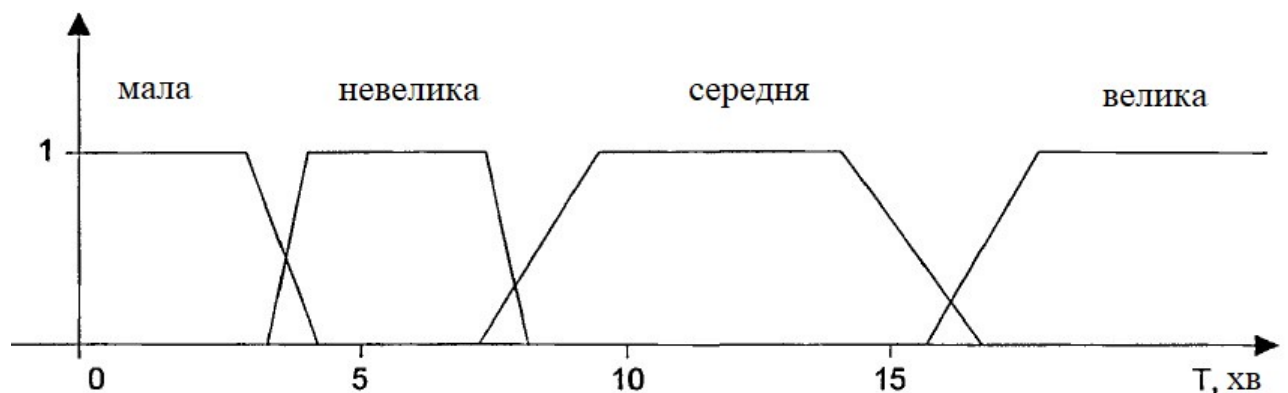


Рис. 3.11. Функція приналежності значень лінгвістичної змінної «тривалість»

Наприклад, функція приналежності «тривалості інтервалу» показана на рис.3.11.

Аналогічним чином визначаються функції належності для інших змінних.

Діагностичний протокол містить набір записів {момент часу; тип ДС}, тому, з урахуванням відомих з фрагмента моделі типів ДС, потрібно розробити правила обробки часових інтервалів протоколу $[t_i; t_{i+1}]$, що містять шукані ДС, оцінки розділяючого їх інтервалу, частот виникнення і темпорального відношення. Припустимо, що часовий інтервал протоколу, що містить обидві ДС визначено. Тоді порядок обробки даних буде наступним.

На першому кроці виконується підрахунок кількостей обох ДС N_1 і N_2 і визначення частот $F1_i$, $F2_i$ шляхом підстановки значень в характеристичну функцію $\mu_F(N)$. На другому кроці порівнюються межі інтервалів появи/усунення ДС. Для спрощення логіки алгоритму розроблено функції порівняння початку інтервалів елементів нечіткого опису *Compare* ($t1$, $t2$) і пошуку накладень інтервалів *CompareOver* ($t1$, $t2$, $t3$, $t4$), а також використаний спосіб представлення темпорального відношення як трійки ознак $\{T_{start}, T_{end}, T_{overlay}\}$, де:

T_{start} - співвідношення між початковими точками часових інтервалів;

T_{end} - співвідношення між кінцевими точками часових інтервалів;

$T_{overlay}$ - ознака накладення/стикування інтервалів.

Кожна ознака може приймати три можливих стани: -1, 0 і 1. Для співвідношень між точками часових інтервалів ці стани позначають - «раніше», «одночасно» і «пізніше», а для ознаки накладення/стикування інтервалів - «не накладаються», «стикуються» і «накладаються».

У випадку наявності кількох ДС одного типу, вони розглядаються як одна ДС з часовим інтервалом ΔT_i від моменту появи першої, до моменту усунення останньої ДС. Наприклад, рівність початкових меж інтервалів $t_{1,1}$ і $t_{2,1}$ в сукупності з нерівністю кінцевих меж $t_{1,2} > t_{2,2}$ дозволяє зробити висновок про наявність відношення $R = \{ \text{«починається»} \}$.

Далі розраховується інтервал між подіями Δt , в загальному випадку рівний часовому проміжку між першими ДС (тому що їх може бути кілька) кожного типу, і потім перетворюється в нечітку оцінку тривалості шляхом підстановки значення Δt в характеристичну функцію

Блок-схема алгоритму формування нечітко-темпорального опису залежності між подіями приведена нарис. 3.12.

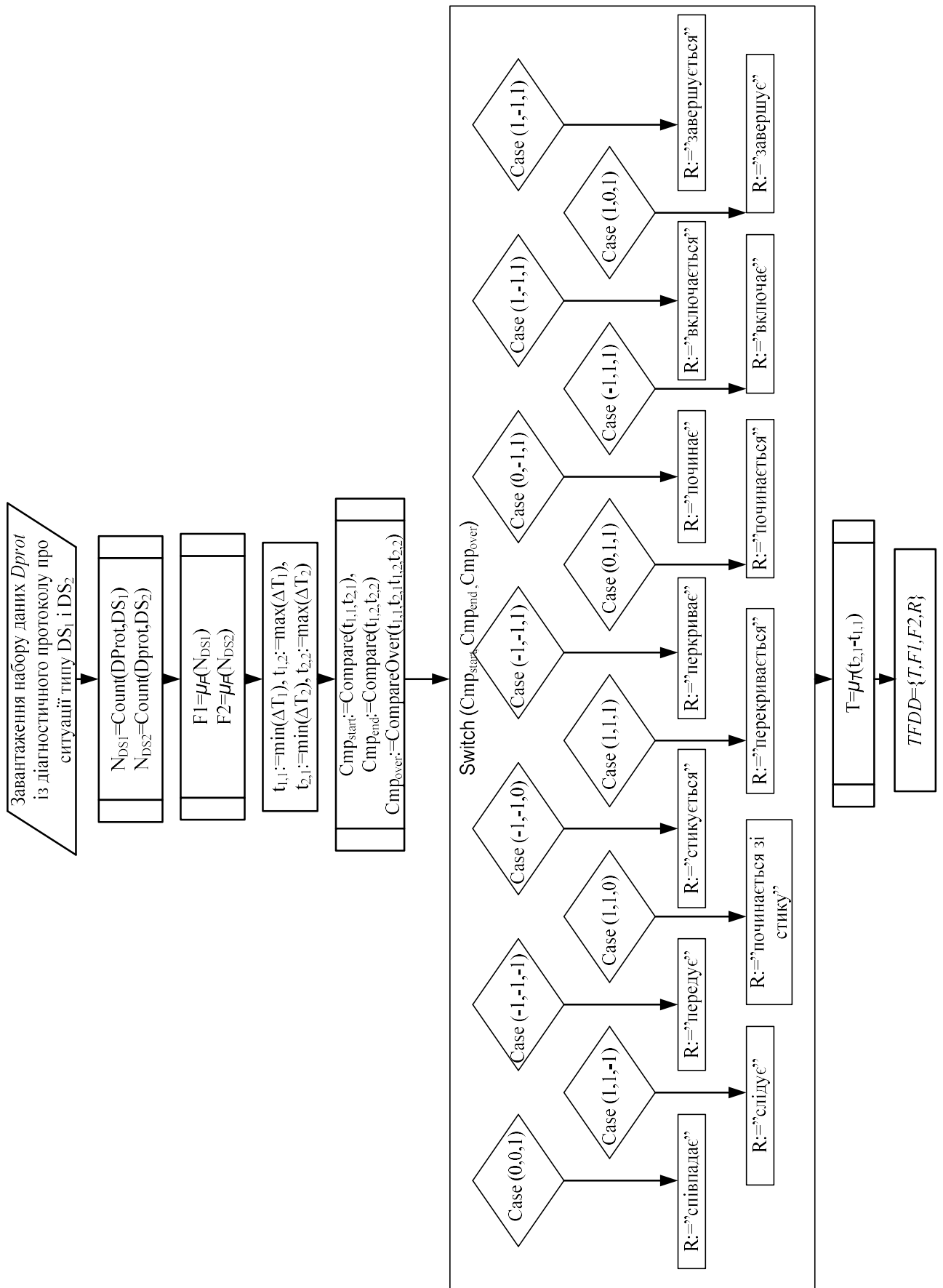


Рис. 3.12. Блок-схема алгоритму формування нечітко-темпорального опису залежності між подіями

3.4 Метод нечітко-темпорального опису змін параметрів контрольованих пристроїв

Як було зазначено раніше, основною метою системи діагностування є формування вірного діагнозу про технічний стан контрольованих пристроїв. Основні проблеми в цій галузі пов'язані зі складною електромагнітною обстановкою, що впливає на точність функціонування засобів вимірювань, і з тим, що діагностичні стани більшості пристроїв слабоформалізовані. Характерні приклади задач діагностування були приведені в розділі 1.1.

Задачі діагностування виконують аналіз працездатності модельованого пристрою, використовуючи стан задач контролю і вимірювання, що діагностується, а також стан задач інших об'єктів. У свою чергу, завдання вимірювання, контролю та формування стану вже взаємодіють безпосередньо з сигналами контролю та вимірюваними параметрами пристроїв.

Таким чином, для докладного вивчення обставин переходу пристрою в той чи інший діагностичний стан, необхідно звернутися до даних протоколу змін сигналів і параметрів, запросивши їх зі сховища. При розборі однієї ситуації можна скористатися відповідними функціями АРМ оператора системи діагностування. Якщо ж потрібно проаналізувати множину ситуацій за певний період часу, або по групі пристроїв, то з'являється необхідність автоматизації цього процесу, необхідність наявності підсистеми, яка виконує аналіз даних протоколу сигналів і вимірювань, формує інформативні узагальнюючі, наприклад нечіткі, описи змін параметрів, придатні для подальшого використання оператором.

Для аналізу даних протоколу змін параметрів пропонується метод формування нечітко-темпорального опису зміни параметра об'єкта діагностування.

Основою методу є нечіткий опис варіантів зміни параметра в часі з урахуванням темпоральних відносин між елементами даного опису параметрів об'єкта.

Для задання нечітких термів введемо чотири лінгвістичних змінних: «початкова величина параметра», «зміна параметра», «тривалість зміни», «вихід за норму». Базові терм-множини цих змінних мають вигляд:

Початкова величина параметру \in {малий, невеликий, середній, великий, значний}

параметру

Зміна параметру \in {значне зменшення, середнє зменшення, незначна зміна, середнє збільшення, значне збільшення}

Тривалість зміни \in {мала, невелика, середня, велика}

Вихід за норму \in {нижче норми, на нижній межі, в межах норми, на верхній межі, вище норми}

Для вираження відносного положення змін параметрів у часі використовуємо часові відносини темпоральної логіки Аллена, перераховані в розділі 3.3.

Нечіткий опис зміни одного параметра об'єкта - це послідовність четвірок $FC = \{\text{«початкова величина параметра»} - V_i, \text{«зміна параметра»} - C_i, \text{«тривалість зміни»} - T_i, \text{«вихід за норму»} - N_i\}$, наприклад зміни, показані на рис.3.13, описується послідовністю, приведеною в табл. 3.7:

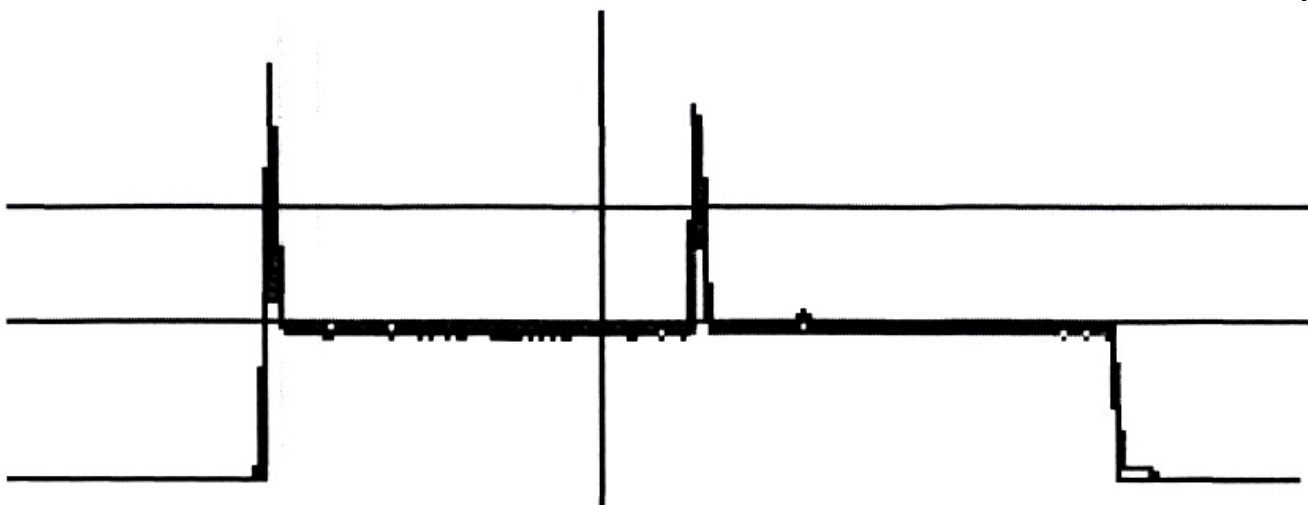


Рис. 3.13. Приклад зміни струму переведення.

Таблиця 3.7 Приклад нечіткого опису зміни параметра об'єкта

№	Початкова величина параметру	Зміна параметру	Тривалість зміни	Вихід за норму
1	Від невеликого	Незначна зміна	Невелика	В межах норми
2	Від невеликого	Значне збільшення	Мала	Вище норми
3	Від великого	Середнє зменшення	Мала	В межах норми
4	Від середнього	Мала зміна	Середня	В межах норми
5	Від середнього	Середнє збільшення	Мала	Вище норми
6	Від великого	Середнє зменшення	Мала	В межах норми
7	Від середнього	Мала зміна	Середня	В межах норми
8	Від середнього	Середнє зменшення	Мала	В межах норми
9	Від невеликого	Мала зміна	середня	В межах норми

Нечітко-темпоральний опис зміни декількох параметрів об'єкта - це мережева структура, вузли якої являють четвірки нечітких описів *FC* (Fuzzy Change), а ребра - темпоральні відносини між ними.

Наприклад, нечітко-темпоральний опис змін параметрів, показаних на рис.3.14,



Рис. 3.14. Приклад зміни параметрів струму переведення.

описується структурою:



Рис. 3.15. Приклад структури НТО зміни параметрів.

Розроблено алгоритм формування нечіткої послідовності з часового ряду вимірювань параметра. Першим кроком є задання функцій приналежності значень лінгвістичних змінних. Використовуються стандартні кусочно-лінійні *Z*-, *P*- і *S*-функції. Наприклад, для «зміни параметра» функції приналежності мають вигляд:

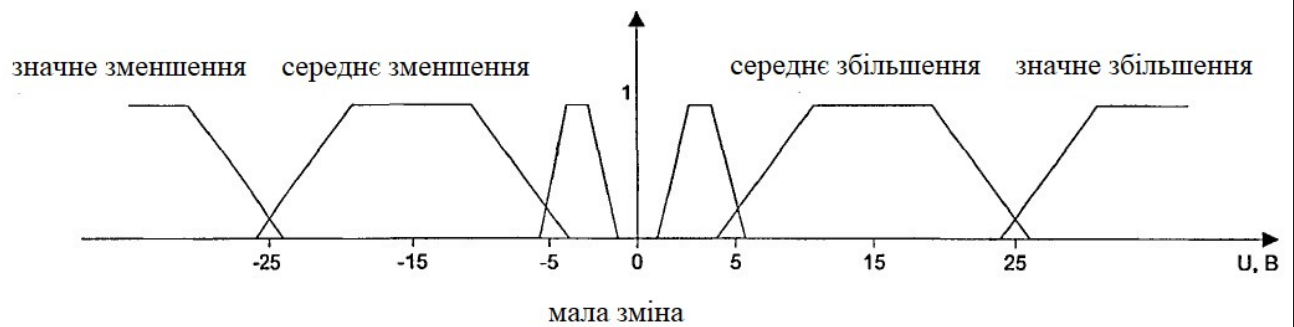


Рис. 3.16. Функція приналежності значень лінгвістичної змінної «зміна параметра» на прикладі напруги

Аналогічним чином визначаються функції належності для інших змінних.

Часовий ряд (ВР) задається набором точок: {момент часу; значення}, тому потрібно розробити правило виділення і перетворення інтервалу $[t_i; t_{i+1}]$ ряду в один елемент нечіткої послідовності. При виділенні інтервалу будемо розширювати його до тих пір, поки різниця значень на межах не перевищить нечітке значення «незначна зміна», або поки не буде досягнутий кінець ВР.

Перетворення обраного інтервалу $[t_i; t_{i+1}]$ в елемент нечіткої послідовності здійснюється підстановкою початкового значення, величини зміни, довжини інтервалу і кінцевого значення у функції приналежності базових терм-множин і визначенням найбільш достовірних нечітких значень. Кінцеве значення використовується для порівняння з нормами параметра і формування ознаки «вихід за норму».

Далі описані дії повторюються для решти часового ряду. Блок-схема алгоритму формування послідовності нечітких описів змін параметрів наведена нижче.

Тут $MAP_1, MAP_2, MAP_3, MAP_4$ - відображення в базові терми множини відповідних лінгвістичних змінних. Збереження чітких значень часу для кожного елемента послідовності дозволить провести додатковий аналіз з

метою встановлення темпоральних зв'язків між нечіткими змінами параметрів.

Отримані послідовності нечітких описів класифікуються експертами, зберігаються в базі знань і можуть використовуватися для порівняння з часовими рядами параметрів, отриманими при даному діагностичному стані, і формування груп станів зі схожими описами.

Порівняння може проводитися або між отриманою послідовністю і набором, що зберігаються в базі знань, або зіставленням безпосередньо часового ряду з усіма послідовностями бази знань. Другий спосіб, на відміну від першого, дозволяє оцінити достовірність приналежності ВР до кожного з класів.

Розроблено алгоритм порівняння послідовності нечітких описів з часовим рядом вимірюваних значень. Для кожного з елементів послідовності нечітких описів (ПНО) оцінюється ступінь його подібності зі зміною параметра P на інтервалі $[t_i; t_{i+1}]$. В першу чергу оцінюється достовірність збігу початкових значень ВР і ПНО - $Dv(V_i, \mu_{P(t_j)})$, де μ - функція приналежності нечіткого значення. Якщо достовірність дорівнює нулю, то порівняння переривається. В іншому випадку розглядається достовірність збігу зміни ВР на інтервалі $Dc(C_i, \mu_{P(t_{j+k})-P(t_j)})$, довжин інтервалів $Dt(T_i, \mu_{t_{j+k}-t_o})$ і ознак виходу вимірюваного параметру за норми $Dn(N_i, \mu_{P(t_{j+k})N})$

Достовірність збігу зміни параметра з відповідним елементом ПНО від значення V_i на величину C_i за час T_i на інтервалі $(t_j; t_{j+k})$ і з ознакою виходу за норми N_i ; розраховується як нечітка кон'юнкція: $D_{i,j} = \min(Dv, Dc, Dt, Dn)$.

Виконується покрокове розширення порівнюваного інтервалу ВР до тих пір, поки розрахункова достовірність $D_{i,j}$ не стане позитивною, а потім рівною нулю. Максимальне значення і номер відліку k_m , на якому воно було досягнуто, фіксуються, а інтервалу $(t_j; t_{j+k_m})$ присвоюється достовірність D_i . Подальше порівняння триває починаючи з k_m -го відліку ВР з $(i+1)$ -м елементом ПНО. Якщо черговий крок порівняння не виявляє інтервалу з

позитивною достовірністю, виконання алгоритму переривається і робиться висновок про схожість ВР з нечіткою послідовністю тільки по $i-1$ першим інтервалах. Результуюча достовірність обчислюється як нечітка кон'юнкція достовірностей усіх подій послідовності: $D = \min\{D_i\}$.

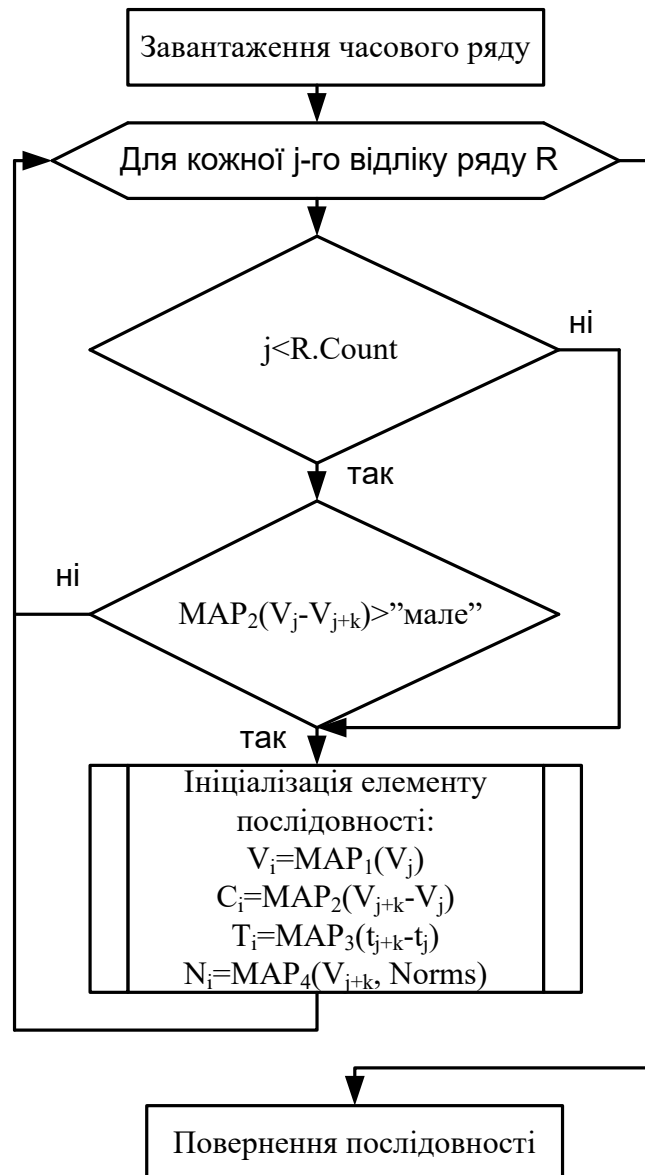


Рис. 3.17. Блок-схема алгоритму формування послідовності нечітких описів

Блок-схема алгоритму порівняння часового ряду вимірювань параметра з послідовністю нечітких описів i , розрахунку достовірності їх відповідності наведена на рис. 3.18.

Нечітко-темпоральний опис змін параметрів (НТО ЗП) діагностуючого пристрою доцільно використовувати для комплексного аналізу їх відносної динаміки. Для визначення темпоральних зв'язків між нечіткими змінами і формування НТО ЗП розроблений наступний алгоритм.

На першому кроці з бази знань завантажуються послідовності нечітких описів, а з протоколу вимірювань - збережені значення часу початку t_{start} елементів послідовності, що відносяться до аналізованого періоду часу роботи конкретного пристрою.

Далі виконується попарне порівняння послідовностей SFC_i (Sequence of Fuzzy Changes) і SFC_j з урахуванням симетричності темпоральних відносин Аллена, що дозволяє вдвічі скоротити число виконуваних операцій. Процедура порівняння являє собою поелементний аналіз початкових меж $t_{i,m}$ і $t_{j,n}$ часових інтервалів елементів послідовності, меж $t_{i,m+1}$ і $t_{j,n+1}$ наступних інтервалів, а також нечітких оцінок $T_{i,m}$ і $T_{j,n}$ їх тривалості.

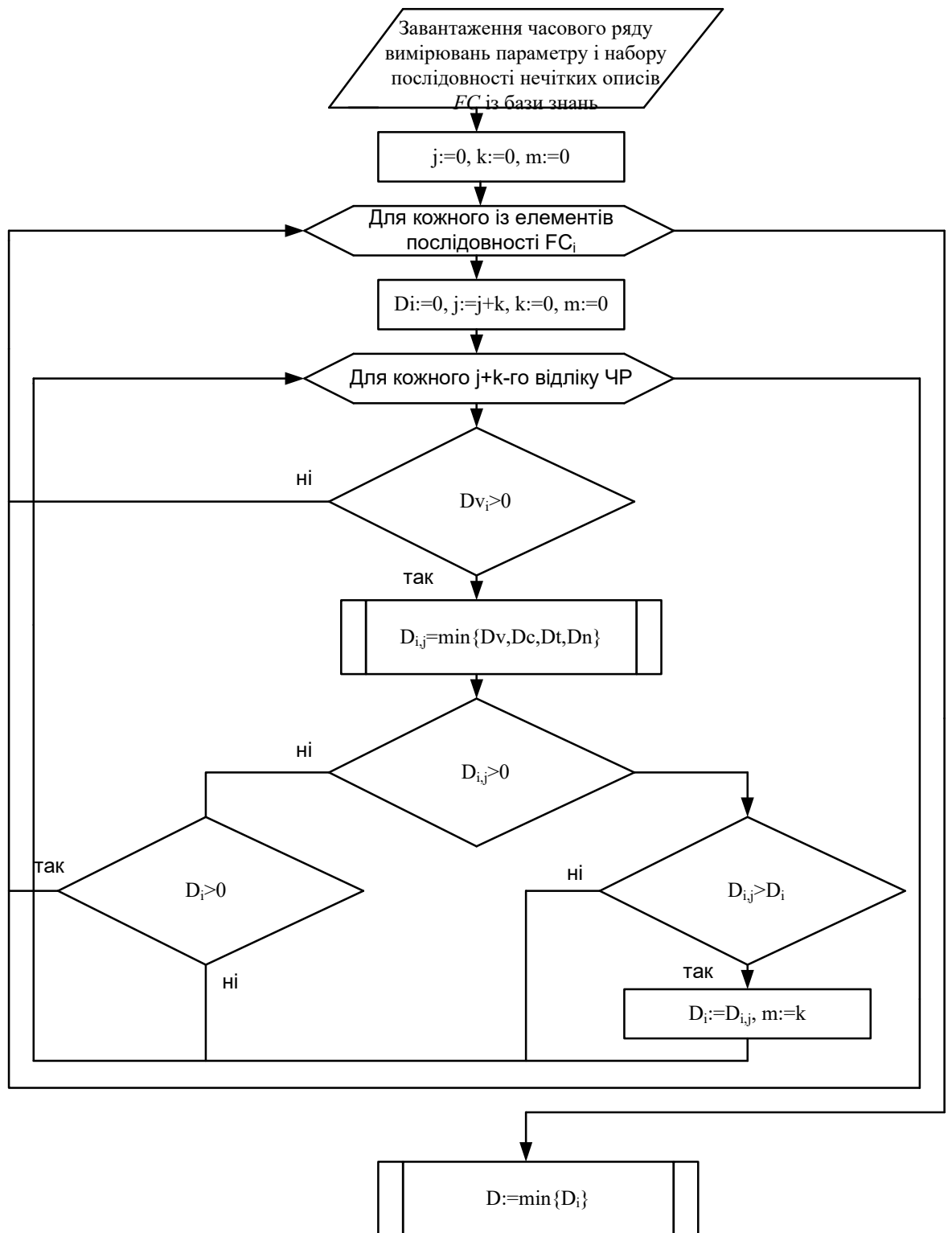


Рис 3.18. Блок-схема алгоритму порівняння часового ряду з послідовністю нечітких описів зміни параметра

Для спрощення логіки роботи алгоритму використовуються запропоновані в розділі 3.3 функції порівняння початку інтервалів елементів нечіткого опису *Compare* (tl , $t2$) і пошуку накладень інтервалів *Compare* (tl ,

t_2, t_3, t_4). Темпоральні відношення представляються як трійки ознак $\{T_{start}, T_{end}, T_{overlay}\}$. На відміну від алгоритму, описаного в розділі 3.3, темпоральні відношення визначаються не між двома інтервалами, а між набором послідовностей інтервалів.

Наприклад, рівність початкових меж інтервалів $t_{i,m}$ і $t_{j,n}$ дозволяє зробити висновок про наявність відношень $R_{i,j,m,n} = \{ \text{«збігається» або «починає» («починається») } \}$. Уточнення типу темпорального відношення $R_{i,j,m,n}$ виконується за допомогою аналізу меж $t_{i,m+1}$ і $t_{j,n+1}$ наступних елементів ПНО, а також (у випадку аналізу останніх елементів послідовності) порівняння нечіткої оцінки різниці T_i між початковими межами інтервалів з нечіткою оцінкою тривалості $T_{i,m}$ поточного елемента.

Таким чином, результуюча множина темпоральних відношень між i -ою і j -ою послідовностями $R_{i,j}$ містить відношення $R_{i,j,m,n}$ між такими m -ми і n -ми елементами, які стикаються, або повністю або частково накладаються один на одного.

Сукупність послідовностей нечітких описів SFC і множини темпоральних відносин між ними і складають нечітко-темпоральні опис зміни декількох параметрів пристрою.

Блок-схема алгоритму формування нечітко-темпорального опису змін параметрів пристрою приведена на рис. 3.19.

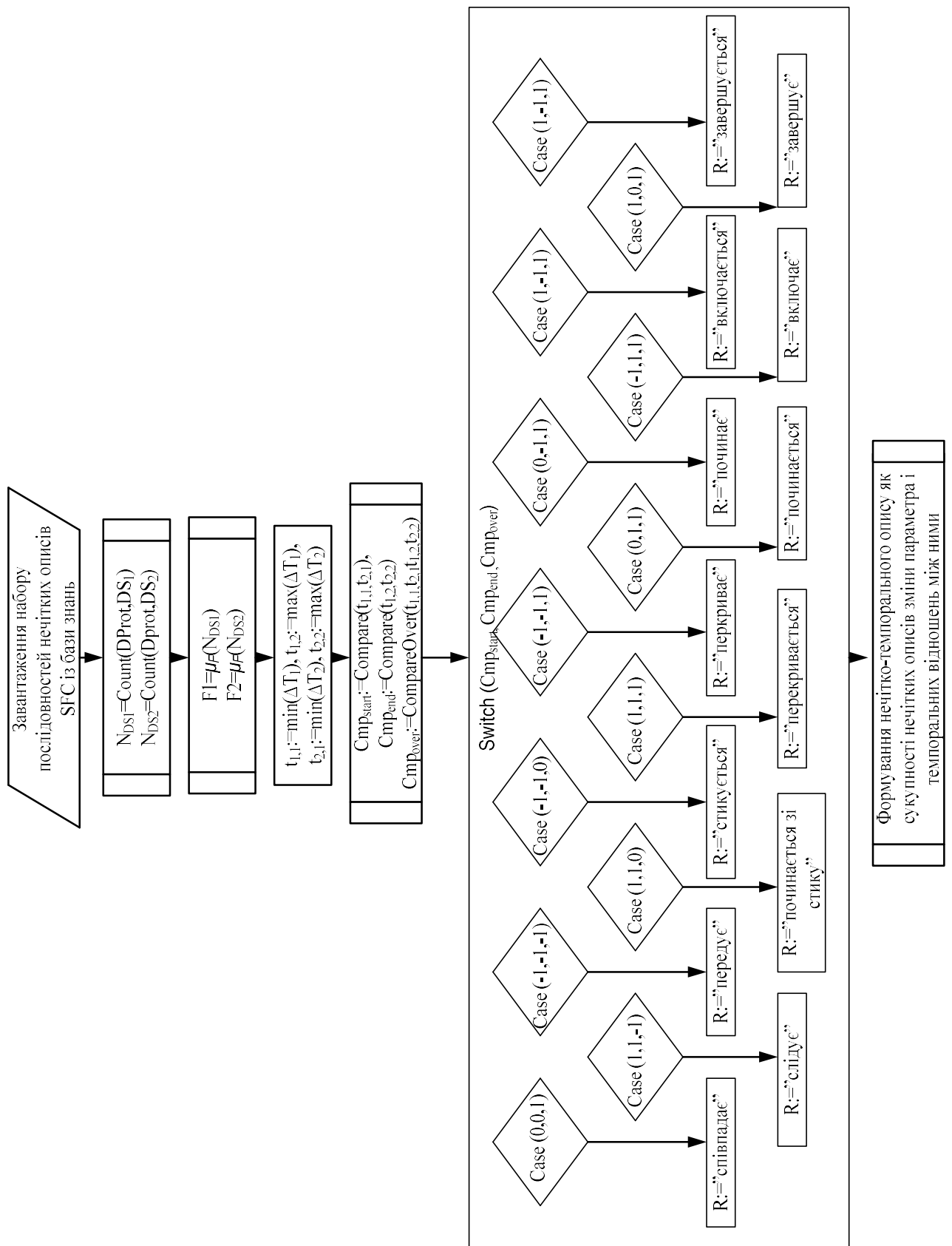


Рис 3.19. Блок-схема алгоритму формування нечітко-темпорального опису зміни параметра

3.5. Висновки до розділу 3

1. На основі підходу, обґрунтованого в Розділі 2, розроблена мережева модель представлення залежностей між діагностичними станами пристроїв, доповнена такими атрибутами, як таблиці умовних ймовірностей і нечітко-темпоральні опису змін діагностичних станів і змін вимірюваних параметрів контрольованих пристроїв, що дозволяють розширити спектр застосовуваних підходів при використанні бази знань виявлених залежностей. Наведено приклад мережевої моделі залежностей між діагностичними станами стрілочного приводу.

2. Розроблено комплексний метод асоціативного аналізу даних і побудови мережевої моделі залежностей між подіями, розроблені алгоритми побудови моделі залежностей на основі принципу мінімальної довжини опису і модифікованої логарифмічною метрики, наведені приклади використання алгоритмів для обробки вибірок діагностичних даних, виконано порівняльний аналіз ефективності роботи алгоритмів.

3. Розроблено метод нечітко-темпорального опису залежностей між діагностичними станами пристроїв АТ, розроблений алгоритм його формування за діагностичним даними, визначений вид функцій приналежності нечітких лінгвістичних змінних, наведено приклад нечітко-темпоральної залежності між парою діагностичних станів.

4. Розроблено метод нечітко-темпорального опису зміни параметрів контрольованого об'єкта, алгоритми формування нечітко-темпорального опису залежності між змінами параметрів на основі аналізу протоколів вимірювань.

ВИСНОВОК

В результаті виконаних теоретичних і практичних досліджень по темі дипломного проекту досягнуті наступні результати:

1. Проведений аналіз стану та актуальних проблем технічного діагностування пристроїв автоматики та телемеханіки дозволив виявити потребу в застосуванні моделей і методів, що володіють низьким ступенем помилкових спрацьовувань і допускають некритичне відхилення вимірюваного параметра від заданих норм, а також важливість для системи діагностування таких параметрів, як наявність розподіленої структури, властивість розширюваності і функцій інформаційної взаємодії з суміжними системами.

2. Розгляд сучасних систем діагностування технічних пристроїв, застосовуваних на транспорті і в інших областях, дозволило виявити відставання систем діагностування пристроїв від систем схожого призначення інших галузей за рівнем інтелектуалізації та реалізації експертних функцій, а саме, виявити відсутність функціональності формування експертних оцінок стану технічних пристроїв, виявлення прихованих закономірностей і тенденцій за допомогою методів інтелектуального аналізу даних, прогнозування стану контрольованих пристроїв.

3. Виконаний аналіз стану, проблем та принципів реалізації сучасних систем зберігання і аналізу даних дозволив зробити висновок про актуальність розробки нових моделей уявлення і методів аналізу даних.

4. Доведено необхідність застосування інтелектуальних методів аналізу даних і експертних систем, що дозволить усунути відставання в рівні автоматизації і підвищить ефективність роботи персоналу центрів діагностування та моніторингу, а також забезпечить підтримку прийняття рішень при обробці позаштатних діагностичних станів контрольованих об'єктів.

5. На основі виконаного аналізу існуючих підходів до подання знань в інтелектуальних системах, і з урахуванням вимог, що пред'являються до моделі знань про залежності між діагностичними станами, зроблено

висновок про доцільність застосування мережевого підходу до подання знань.

6. Розгляд методів інтелектуального аналізу даних з урахуванням основних завдань діагностики технічних пристроїв дозволило визначити групу методів для вирішення завдання виявлення залежностей, запропонувати спосіб поліпшення характеристик даної групи і сформулювати загальну ідею комплексного методу асоціативного аналізу даних.

7. Проведений аналіз методів формування міркувань на основі неповних даних з урахуванням специфіки трудноформалізуємих процесів функціонування пристроїв показав необхідність застосування засобів імовірнісного і нечітко-логічного підходів на етапі використання виявлених залежностей для підтримки прийняття рішень персоналу центру діагностування та моніторингу, і як наслідок - необхідність доповнення моделі залежностей діагностичних станів відповідними атрибутами.

8. На основі обгрунтованого в Розділі 2 підходу-розроблена мережева модель представлення залежностей між діагностичними станами, доповнена такими атрибутами, як таблиці умовних ймовірностей і нечітко-темпоральні опису змін діагностичних станів і змін вимірюваних параметрів контрольованих пристроїв, що дозволяє розширити спектр застосовуваних підходів при використанні бази знань виявлених залежностей. Наведено приклад мережевої моделі залежностей між діагностичними станами визначеного приводу.

9. Розроблено комплексний метод асоціативного аналізу даних і побудови мережевої моделі залежностей між подіями, алгоритми побудови моделі залежностей на основі принципу мінімальної довжини опису і модифікованої логарифмічною метрики, наведені приклади обробки вибірок діагностичних даних, виконано порівняльний аналіз ефективності роботи алгоритмів.

10. Розроблено метод нечітко-темпорального опису залежностей між діагностичними станами пристроїв, алгоритм формування НТО залежності по діагностичним даними, визначений вид функцій приналежності нечітких лінгвістичних змінних, наведено приклад НТО залежності між парою діагностичних станів.

11. Розроблено метод нечітко-темпорального опису зміни параметрів контрольованого об'єкта, розроблені алгоритми формування нечітко-темпорального опису залежності між змінами параметрів на основі аналізу протоколів вимірювань.

12. Розроблено структуру системи інтелектуального аналізу даних, що включає підсистеми аналізу діагностичних даних і формування моделей залежностей між ДС, формування нечітко-темпоральних описів залежностей і змін параметрів пристроїв, підсистему призначеного для користувача інтерфейсу і базу знань.

13. Виконано розробку структури і реалізація на основі СУБД MS SQL Server бази даних системи діагностування (включаючи дані про діагностичні та динамічних станах контрольованих пристроїв, протоколи вимірювань параметрів, нормативно-довідкову інформацію, проектні і конфігураційні дані) і бази знань системи інтелектуального аналізу даних (включаючи дані про параметри моделей залежностей, нечітко-темпоральні опису залежностей і змін параметрів, які супроводжують зміну діагностичного стану).

14. Розроблено підсистеми аналізу діагностичних даних (в складі модулів взаємодії з базою даних, трансформації протоколу даних в матричну форму, аналізу даних, взаємодії з підсистемами формування нечітко-темпоральних описів, взаємодії з базою знань) і формування нечітко-темпоральних описів змін стану і змін вимірювальних параметрів (в складі служб оперативного аналізу протоколів), виявлення фактів виконання робіт з ТОіР.

СПИСОК БІБЛІОГРАФІЧНИХ ПОСИЛАНЬ ВИКОРИСТАНИХ

ДЖЕРЕЛ

1. *Kazuo Tanaka*. Fuzzy Control Systems Design and Analysis: A Linear Matrix Inequality Approach / Nanaka Kazuo, O.Wang Hua. – Wiley-Interscience; 1st edition, 2008. – 320 p. ISBN-13: 978-0471323242.
2. *Steven Schockaert*. Reasoning about fuzzy temporal and spatial information from the web (intelligent information systems) / Schockaert Steven. – World Scientific Publishing Company; 3rd edition, 2010. – 608 p. ISBN-13: 978-9814307895.
3. *Witold Pedrycz*. An Introduction to Computing with Fuzzy Sets: Analysis, Design, and Applications (Intelligent Systems Reference Library, 190) / Pedrycz Witold. – Springer, 2020. – 297 p. ISBN-13: 978-3030527990.
4. *Jeff Hawkins*. On Intelligence: How a new understanding of the brain will lead to the creation of truly intelligent machines / Hawkins Jeff, Blakeslee Sandra. – St. Martin's Griffin; Reprint edition, 2005. – 272 p. ISBN-13: 978-08050078534.
5. *M Hadjiski*. Intuitionistic fuzziness and other intelligent theories and their applications (studies in computational intelligence 757) / Hadjiski M, Atanassov K T. – Springer; 1st edition 2019 edition, 2018. – 203 p. ISBN-13: 978-3319789309.
6. *Bing-Yuan Cao*. Fuzzy information and engineering (Advances in Intelligent and Soft Computing) / Cao Bing-Yaun, Wang Guojun. – Springer; 2010th edition, 2010. – 1200 p. ISBN-13: 978-3642148798.
7. *Bai Lu Yi Zhu*. Fuzzy temporal data modeling and XML query: Concepts and Techniques Electrical Principles and Practices / Zhu Yi Lu Bai. – Northeastern University Press; Chinese edition, 2014. – 181 p. ISBN-13: 978-7551708425.
8. *Черняк О.І.*. Інтелектуальний аналіз даних: підручник / О.І. Черняк, П.В. Захарченко. – К.: МОН України, 2010. – 837 с.

9. *Ланде Д.В.* Основи теорії і практики інтелектуального аналізу даних у сфері кібербезпеки / Д.В. Ланде, І.Ю. Субач, Ю.Є. Биков: навчальний посібник. – К.: ІСЗЗ КПІ ім. Ігоря Сікорського, 2018. – 297с. ISBN 978-966-2577-5.
10. *Бурдаков С. Ф.* Робастное управление нелинейными механическими системами с помощью линейных обратных связей / С.Ф. Бурдаков, А.А. Первозванский, Л.Б. Фрейдович// Автоматика и телемеханика. – № 11. – 1999. – с. 69-80.
11. *Прокопенко Т.О.* Теорія систем і системний аналіз: навчальний посібник / Т.О. Прокопенко. – Черкаси: ЧДТУ, 2019. – 139 с.
12. Інтелектуальний аналіз даних. Режим доступу: https://thedigital.gov.ua/lms_ai/intelektualnij-analiz-danih
13. *Барсегян А.А.* Анализ данных и процессов: учеб. пособие / А.А. Барсегян, М.С. Куприянов, И.И. Холод. – СПб.: БХВ-Петербург, 2009. – 512 с. ISBN 978-5-9775-0368-6.
14. *Афанасьева Т.В.* АПрименене методов интеллектуального анализа данных и процессов: практикум / Т.В. Афанасьева. – Ульяновск: УлГТУ, 2018. – 51 с.
15. Темпоральная логика. Режим доступа: <https://ru.wikipedia.org/wiki/>